Helmet-Mounted Displays:
Sensation, Perception and Cognition Issues

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Helmet-Mounted Displays: Sensation, Perception and Cognition Issues**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Aeromedical Research Laboratory,P.O. Box 620577,Fort Rucker,AL,36362-0577**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**970**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

CONTRIBUTORS

Bruce E. Amrein, Research Engineer
Visual and Auditory Processes Branch
Human Research and Engineering Directorate
U.S. Army Research Laboratory
AMSRD-ARL-HD / Bldg. 520
Aberdeen Proving Ground, MD 21005-5425

Michael M. Bayer, Display Systems Consultant
1624 The Alameda #41
San Jose, CA 95126

Mary S. Binseel, Research Engineer
Visual and Auditory Processes Branch
Human Research and Engineering Directorate
U.S. Army Research Laboratory
AMSRD-ARL-HD / Bldg. 520
Aberdeen Proving Ground, MD 21005-5425

James H. Brindle, Display Systems/Electronics Engineer
LME, Inc.
P.O. Box 2303
Warminster, PA 18974

Frederick T. Brozoski, General Engineer
Injury Biomechanics Branch
Warfighter Protection Division
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577, Bldg. 6901
Fort Rucker, AL 36362-0577

John A. Caldwell, Research Psychologist
Senior Scientist
Archinoetics, LLC
700 Bishop Street, Suite 2000
Honolulu, HI 96813

J. Lynn Caldwell, Senior Research Psychologist
Human Performance Directorate
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH 45433

Kara D. Cave, Military Audiologist
Hearing Program Section
Soldier Family Medical Center
Ricker Road / Bldg. 2596
Fort Bliss, TX 79916

Ian P. Curry, Flight Surgeon
Royal Army Medical Corps
Joint Helicopter Command
Wilton, UK

Diana C. Emanuel, Professor
Department of Audiology, Speech-Language Pathology and Deaf Studies
Towson University
8000 York Road
Towson, MD 21252-0001

Gregory S. Francis, Professor
Psychological Sciences
Purdue University
Lafayette, IN 47907

Thomas G. Ghirardelli, Assistant Professor
Department of Psychology
Goucher College
1021 Dulaney Valley Road
Baltimore, MD 21204-2794

Thomas H. Harding, Neurophysiologist
Director, Sensory Research Division
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577, Bldg. 6901
Fort Rucker, AL 36362-0577

Keith L. Hiatt, Flight Surgeon
Medical Director
U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760

Ronald P. King, Flight Surgeon
Director, Warfighter Protection Division
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577, Bldg. 6901
Fort Rucker, AL 36362-0577

Gregory T. Lang, Senior Flight Surgeon
Diving Medical Officer
Medical Corps, United States Army

Tomasz R. Letowski, Senior Research Scientist
Visual and Auditory Processes Branch
Human Research and Engineering Directorate
U.S. Army Research Laboratory
AMSRD-ARL-HD / Bldg. 520
Aberdeen Proving Ground, MD 21005-5425

Sumalai Maroonroge, Associate Professor
Communication Disorders Department
Lamar University
Beaumont, TX 77710

William E. McLean, Research Optometrist
Visual Sciences Branch
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577, Bldg. 6901
Fort Rucker, AL 36362-0577

James E. Melzer, Manager of Research and Technology
Rockwell Collins Optronics
2752 Loker Avenue West
Carlsbad, CA 92010

Kimberly P. Myles, Human Factors Engineer
Visual and Auditory Processes Branch
Human Research and Engineering Directorate
U.S. Army Research Laboratory
AMSRD-ARL-HD / Bldg. 520
Aberdeen Proving Ground, MD 21005-5425

Robert K. Noback, Flight Surgeon
Commander/Dean
U.S. Army School of Aviation Medicine
301 Dustoff Avenue
Fort Rucker, AL 36362

Allan. J. Pantle, Professor Emeritus
Department of Psychology
Miami University
Oxford, OH 45056

Alan R. Pinkus, Research Psychologist
Warfighter Interface Division
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH 45433

Clarence E. Rash, Research Physicist
Visual Sciences Branch
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577, Bldg. 6901
Fort Rucker, AL 36362-0577

Michael B. Russo, Neurologist/Sleep Disorders Specialist
Tripler Army Medical Center
1 Jarrett White Road
Tripler Army Medical Center, HI 96859

Thomas O. Salmon, Associate Professor of Optometry
Northeastern State University
1001 N. Grand Avenue
Tahlequah, OK 74464

Angélique A. Scharine, Research Psychologist
Visual and Auditory Processes Branch
Human Research and Engineering Directorate
U.S. Army Research Laboratory
AMSRD-ARL-HD / Bldg. 520
Aberdeen Proving Ground, MD 21005-5425

Elmar T. Schmeisser, Research Physiologist
Program Manager, LS/Neurosciences
U.S. Army Research Office,
Research Triangle Park, NC 27709

H. Lee Task, Optical Physicist
Task Consulting
Tuscon, AZ 85745

Leonard A. Temme, Research Psychologist
Chief, Visual Sciences Branch
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577, Bldg. 6901
Fort Rucker, AL 36362-0577

Phuong K. Tran, Research Engineer
Visual and Auditory Processes Branch
Human Research and Engineering Directorate
U.S. Army Research Laboratory
AMSRD-ARL-HD / Bldg. 520
Aberdeen Proving Ground, MD 21005-5425

Robert M. Wildzunas, Research Psychologist
U.S. Army School of Aviation Medicine
301 Dustoff Avenue
Fort Rucker, AL 36362

Helmet-Mounted Displays:
Sensation, Perception and Cognition Issues

Edited by

Clarence E. Rash
Michael B. Russo
Tomasz R. Letowski
Elmar T. Schmeisser

U.S. Army Aeromedical Research Laboratory
Fort Rucker, Alabama

U.S. Army Research Laboratory
Aberdeen Proving Ground, Maryland

With a foreword by

CW4 J. Kevin Heinecke
U.S. Army Master Aviator
AH-64/AH-1/UH-72/UH-60/UH-1/OH-58/OH-6 and C-12 Pilot

Library of Congress Cataloging-in-Publication Data

Helmet-mounted displays: Sensation, perception and cognition issues/ edited by Clarence E. Rash, Michael B. Russo, Tomasz R. Letowski and Elmar T. Schmeisser

ISBN 978-0-615-28375-3

Published by

U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, Alabama 36362-0577

Printed in the United States of America

CONTENTS

Part Five: Meeting the HMD Design Challenge

FOREWORD

All our knowledge has its origins in our perceptions. - Leonardo Da Vinci

As a U.S. Army attack helicopter pilot and veteran user of night vision goggles (NVG) and forward-looking infrared (FLIR) pilotage systems since 1989, I have an ingrained appreciation for the technology pilots use to enhance situation awareness (SA) on the battlefield and in training. SA is defined as knowing where you are in 3-dimensional space, in particular knowing where you are with respect to other military assets (e.g., planes, tanks, troops) and the surrounding terrain/man-made objects. One of the most anxiety-ridden moments faced by any military pilot is a loss of SA while navigating from one place to another. Coupled with a low fuel situation, I can see where a sense of urgency in recognizing familiar terrain features could be heightened. This situation becomes even more stressful when you, the pilot, compound your misplaced aircraft problem with doing so at night. Here is when you need your night vision system to translate the most recognizable rendition of the actual outside world to your brain. At these moments you either thank, or curse, the technology gods for your respective night vision imaging systems and their displays.

The current helmet-mounted display (HMD) systems have a rich history steeped in military needs met by engineering advancements in optics coupled with enhanced understanding of the human night vision dilemma. NVG technology came online in the early 1970's as the U.S. Army was attempting to expand its night warfighting capability. Ground assets were already using goggles mounted to their headgear when Army aviation began using a 2nd generation version. NVGs provide the viewer with an "enhanced" scene of an otherwise darkened landscape through light amplification. The scene is presented in green or orange based upon the phosphor color used in the goggles. The major problem with the older NVG systems was poor visual acuity and depth perception. Where the system lacked in providing visual cues, the pilot made up for through diligence in general piloting skills (e.g., altitude awareness, constant scanning). By the mid- to late-1970s, thermal (infrared) technology (e.g., FLIR) became available as a sensor technology integrated *into the airframe* in the form of the AH-64 Apache Advanced Attack Helicopter. FLIR used variances in temperature to present an object otherwise obscured in darkness. Objects (e.g., a tank, truck, water tower) may not have looked anything like their daytime images, but they were discernable all the same. The image was presented to the pilot through a single helmet-mounted eyepiece that incorporated the use of a cathode-ray-tube (similar to those used in the old TV sets). Both the NVG and FLIR systems in use today have undergone multiple advancements allowing for improved visual acuity and diminished pilot workload. Until recently applications were limited to these two systems, but as new technological advances are realized, new systems are emerging.

This book will provide insight for pilots, educators, academics, and the general public who are interested in the field of human factors engineering, military night flight operations, and the visual and auditory science behind the improvements in advanced aviation (and other Warfighter) sensor systems. From the explanation of the human-machine interaction dilemma, through the detailing of visual and auditory display systems, this book provides the reader a thorough understanding of the issues related to military operations with respect to our senses, how we perceive what is represented, and ultimately how we assimilate and react to this information.

CW5 J. Kevin Heinecke
U.S. Army Master Aviator
AH-1/AH-64/C-12/ OH-6/OH-58/UH-1/UH-60/UH-72 Pilot

PREFACE

In 1999, the U.S. Army Aeromedical Research Laboratory (USAARL), Fort Rucker, Alabama, published a book that addresses issues for the design of helmet-mounted displays (HMDs) for use in helicopters, *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft* (USAARL, 1999).[1] While primarily an engineering overview of image sources, image quality, optical design approaches, communication systems, hearing protection, and helmet head-supported mass and center-of-mass, the book also addresses such human factors issues as visual and auditory performance, head and face anthropometry, and hearing and vision protection.

In the years since the 1999 book was conceived, HMD applications have greatly expanded, not only within the military but also within the manufacturing and simulation training communities. Significant progress has been made in the development of image source technologies, especially miniature displays. This continuing image source development, coupled with advances in power source engineering - smaller size and greater efficiency, has greatly expanded the number of HMD applications. Within the U.S. Army, HMDs are being designed for use by dismounted and mounted Warfighters as well as for aviators.

As advanced technology penetrates the battlespace, the modern Warfighter is being provided with an ever-increasing stream of information. The motivation of this growing flow of information is the Army's objective to "*See First, Understand First, Act First, & Finish Decisively*." Whether it is a field commander or a lower echelon soldier, every Warfighter will have greater access to both tactical and strategic data and imagery. The vast majority of this information will be presented to the Warfighter in visual and auditory forms via HMDs. For this reason, the design and implementation of HMDs must be optimized to ensure optimal user performance, both visual and auditory.

Paramount in achieving this optimization is attaining a thorough understanding of the relationship between HMDs and the human concepts of perception and cognition. An excellent beginning to acquiring this understanding can be found in *Tactical Display for Soldiers-Human Factors Considerations* (National Academy Press, 1997). Presenting the results of the *Panel on Human Factors in the Design of Tactical Displays for the Individual Soldier* (established by the National Research Council at the request of the U.S. Army Natick Research, Development, and Engineering Center, Natick, Massachusetts), this book discusses critical human factors issues associated with the development of the Army's proposed Land Warrior System, an individual Warfighter monocular HMD. The overall goal of the panel was to identify critical characteristics of HMDs and the capabilities and limitations of the target user (i.e., the Warfighter). One major finding of the panel was the presence of "a lack of understanding of the impact of advanced HMD visual and auditory presentations on Warfighter workload, situational awareness and overall performance." This finding is well-known within the HMD community of researchers and often has been expressed as an important issue.

The work presented here is the second in a series of HMD books. Where the first book focused on engineering design issues, this book focuses on filling the National Research Council's identified gap in understanding the relationship between the HMD hardware design and user perception and cognition of the visual and auditory displays.

Structure of the book

This book is divided into five parts: *Part One – Identifying the Challenges; Part Two – Helmet-Mounted Displays; Part Three – The Human Visual and Auditory Sensory Systems; Part Four – Perception, Cognition and Performance*; and *Part Five – Meeting the HMD Design Challenge*.

---

[1] In 2000, SPIE Press, Bellingham, WA, republished this book under the same title.

In *Part One – Identifying the Challenge*, Chapter 1, *The Military Operational Environment*, discusses the diverse operational requirements of the modern Warfighter, to include taking on such diverse roles as combatant, peacekeeper, and disaster relief worker; operating in dissimilar physical environments where heat, cold, fog, rain, smoke, etc., degrade system and human performance; and enduring such performance stressors as fatigue, interruption of circadian rhythm, working under severe time constraints, etc. The chapter also describes the ongoing "transformation" of the U.S. Army and its effect on the individual Warfighter. Chapter 2, *The Human-Machine Interface Challenge*, examines the age-old problem of the human-machine interface, describes the visual tasks encountered by Warfighters in both training and combat; briefly introduces the visual and auditory sensory inputs and the concepts of human perception and cognition; explains the roles of stimuli, sensors and displays in HMDs; discusses future HMD systems and trends; and concludes with a statement of the challenge facing HMD designers in ensuring that newly developed systems optimize user performance by taking into consideration the performance characteristics and limitations of the human brain and visual and auditory senses.

In *Part Two – Helmet-Mounted Displays*, Chapter 3, *Introduction to Helmet-Mounted Displays*, defines an HMD; describes one method of classifying HMDs by optical approach; reviews the history of HMD development; discusses HMD applications; lists the advantages and limitations of HMDs; and provides synopses of current and future HMD programs. Chapter 4, *Visual Helmet-Mounted Displays,* addresses design considerations for *visual* HMDs, to include the importance of image quality, image source technologies, and design parameters (e.g., field-of-view, magnification, exit pupil, etc.). Chapter 5, *Audio Helmet-Mounted Displays*, provides a parallel discussion of *audio* HMD design considerations, to include noise attenuation and communication speech intelligibility.

Part Three – The Human Visual and Auditory Sensory Systems - introduces the human sense organs for vision and audition. Chapters 6, *Basic Anatomy and Physiology of the Human Eye,* and 8, *Basic Anatomy of the Hearing System*, review the basic anatomy, structure, and physiology of the human eye and ear, respectively. These chapters are intended to provide the reader with a fundamental understanding of these two critical sensory systems. Chapter 7, *Visual Function*, discusses the vision process, starting with light originating from an object or source and the formation of an image on the retina. This chapter continues with explanations of various visual functions, e.g., color vision, accommodation, the ocular-motor function, etc. In an analogous manner, Chapter 9, *Auditory Function,* discusses the hearing process, starting with the production of sound by stimuli and continues with explanations of theories of hearing, neural coding and the processing of sound in the brain.

Having discussed vision and audition from a sensory perspective, *Part Four – Perception, Cognition and Performance* – addresses the major impetus of this book – perceptual and cognitive issues associated with HMD design. Chapters 10, *Visual Perception and Cognitive Performance*, and 11, *Auditory Perception and Cognitive Performance*, discuss visual and auditory perception and performance, respectively. Visual factors discussed include brightness perception, pattern recognition, motion and depth perception, and 2- vs. 3-dimensional presentations. Auditory factors include loudness and pitch perception, speech recognition, sound localization, and hearing deficits. Chapters 12, *Visual Perceptual Conflicts and Illusions*, and 13, *Auditory Conflicts and Illusions*, describe perceptual conflicts and illusions (visual and auditory, respectively) that Warfighters may encounter and must overcome. Visual conflicts and illusions discussed include static and dynamic illusions, masking, binocular rivalry, spatial disorientation, and special issues such as hyperstereopsis and luning. Auditory conflicts and illusions discussed include masking, spatial hearing, binaural rivalry and the issue of auditory channel capacity. In Chapter 14, *Auditory-Visual Interactions*, the issues of multisensory perception, including synergy, redundancy and synchrony are explored. The cognitive factors of attention, memory and decision making are discussed in Chapter 15, *Cognitive Factors*. This section concludes with an in-depth overview of performance effects in the presence of mechanical, physiological, sensory, and cognitive adverse operational factors in Chapter 16, *Performance Effects Due to Adverse Operational Factors*. Such factors include vibration, fatigue, stress, workload level, and extreme environmental conditions.

The book concludes with *Part Five – Meeting the HMD Design Challenge*. The first chapter, Chapter 17, *Guidelines for HMD Design,* provides summary guidelines and recommendations for creating an optimal design for an HMD for a defined application based upon the various optical/visual, acoustic/auditory, perceptual/cognitive, and user adjustment topics and concerns discussed in earlier chapters. It discusses tradeoffs in design parameter values and the impact of such tradeoffs on system and user performance. Included in this chapter is a brief, but essential, reminder of other design issues not covered in previous chapters, e.g., the biodynamic issues of head-supported weight and center-of-mass offsets. Chapter 18, *Exploring the Tactile Modality for HMDs*, goes beyond current optical and acoustic HMD designs and explores the potential of adding a haptic modality to HMD designs by introducing tactile information flow and force feedback. The final chapter, Chapter 19, *The Potential of an Interactive HMD*, looks further to the future of HMDs. The concept of the HMD as an interactive system is explored through the implementation of neuro-physiological monitoring technologies, such as electro-cortical, evoked potentials, and ocular-motor measures.

Limitations of the book

This book is intended to address the issues of HMDs as they pertain to the processes of human sensation, perception and cognition. However, the enormous scope of these subject areas precludes this work from being all-inclusive. The emphasis is placed on the military environment. Nonmilitary HMD applications, especially in the fields of virtual reality and simulation, are not explored to their fullest extent. While the authors liberally draw upon data derived from research supporting such applications, the data are presented in a military context, and even then with greater emphasis on Army applications. This is an unapologetic consequence of the areas of experience and expertise of most of the book's contributors. Fortunately, this does not preclude the information presented here from being useful in tri-service military and nonmilitary applications. While not explicit, much of the technical data is derived from research and development from around the world. Indeed, many nations have contributed and to continue to contribute (and many cases, lead) to the design, production and fielding experience of HMDs.

In addition, the material in Chapter 17 only superficially discusses the equally important biodynamic design issues that remind us that the HMD is not a stand-alone component, but instead is an integrated part of the Warfighter, vehicle or aircraft system.

Acknowledgments

The editors would like to extend their sincere appreciation to all of the authors who took an immense amount of time out of their personal lives to make this book possible. The dedication and expertise of these authors are indicative of the professionalism of these individuals. Drawn from the various military services, federal laboratories and academic world, these individuals are all highly respected subject-matter-experts within their fields of knowledge.

As important as the role of the authors have been to the completion of this book, it would never have come to fruition without the editorial, graphics and production contributions of many other individuals. Hence, the editors and authors would like to thank the following individuals for their assistance in the development of this book:

- John S. Martin, Scott C. Childress, and John D. Sowell, U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL, are acknowledged for their invaluable assistance in creating many of the figures and tables.
- John D. Sowell and Scott C. Childress, U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL, deserve special recognition for the cover design and artwork.
- Alan R. Pinkus, Air Force Research Laboratory, Wright-Patterson Air Force Base, OH, for his tremendous personal and institutional support that includes checking for scientific accuracy,

development of chapter outlines, and recruitment of leading experts in vision as contributors to this effort.

- Sharon D. Manning, Bruce Amrein, and Szymon Letowski for their meticulous and often repetitive review of the many drafts of various chapters.
- Loraine Parish St. Onge, editor, U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL, for her review of the final chapter versions.

Clarence. E. Rash
Michael B. Russo
Tomasz R. Letowski
Elmar T. Schmeisser

Fort Rucker, AL
Aberdeen Proving Ground, MD
Research Triangle Park, NC

Helmet-Mounted Displays:
Sensation, Perception and Cognition Issues

Part One

Identifying the Challenges

The role of the Warfighter in the modern world has changed tremendously over the past two decades. While the primary job remains defeating the enemy, the Warfighter's role has been expanded to include peacekeeping, disaster relief, humanitarian aid, and anti-terrorism. To more effectively perform these tasks, the U.S. military is transforming itself into a more responsive and agile force that leverages advanced technologies. These advanced systems can expand the operational environment and multiply individual and unit capabilities. However, achieving optimal performance with these systems requires matching the engineering design characteristics of the system with the characteristics of the human user. Nowhere is this truer than for head- or helmet-mounted displays (HMDs), because such systems are intimately mated to the human senses of vision and audition. Failure to understand the human-machine interface can result in degraded performance, which for the Warfighter can mean the difference between mission success and failure or between a safe return and becoming a casualty. The issues of the human-machine interface encompass human anatomy and anthropometry, ergonomics, and human factors. Embedded in these issues is the important requirement to understand the roles of sensation, perception and cognition in the optimization of human performance with these advanced systems.

1 THE MILITARY OPERATIONAL ENVIRONMENT

Keith L. Hiatt
Clarence E. Rash

Helmet- (and head-) mounted displays (HMDs) are but one of an array of technologies proliferating on the modern battlefield, now referred to as the "battlespace," thereby recognizing the true x, y, z three-dimensional (3-D) nature of today's military engagements. Some would argue that the battlespace has become multi-front and very fluid in nature and should be thought of as four-dimensional, adding time as an additional dimension.

The intent of these new technologies, especially HMDs, is to increase individual and unit performance to ensure mission success when operating in such complex scenarios. Advances in technology have successfully decreased the physical demands of many occupations but at the expense of increasing the mental or cognitive demands (Cheung, Westwood, and Knox, in press). Paradoxically, this increase in cognitive demand is paralleled and exacerbated by an increase in the availability of information needed to be processed in today's military setting.

In today's battlespace, information is considered as important as any weapon system. More and more, HMDs are becoming the mode of choice for presenting this information. HMDs provide Warfighters[1] with the capability of head-up presentation of the vast amount of tactical and strategic information becoming increasingly available at the individual Warfighter level. While long a mainstay of the aviation community, HMDs are rapidly expanding across all military applications, being fielded by infantry, mechanized, aviation and shipboard Warfighters alike.

An HMD can be described as a compact optical projection system, mounted on or built into a helmet, and used to project a scene and/or data directly into the eye(s) of the user (Laurin Publishing, 2005). In many applications, it is also referred to as a visually coupled system (VCS). While a basic HMD design may only consist of an image source with display delivery optics (attached to a helmet or other head mount), the concept of a visually coupled display includes some mechanism for head/eye tracking. An example of an HMD is provided in Figure 1-1, which depicts the Integrated Helmet and Display Sighting System (IHADSS) HMD used on the U.S. Army's AH-64 Apache helicopter (see Chapter 3, *Introduction to Helmet-Mounted Displays*).

To recognize the advantages, limitations and constraints of HMDs and their associated technologies for the Warfighter, and ultimately their impact on perceptual and cognitive performance, it is necessary to understand the military environment and how this environment and the role of the Warfighter itself have changed over the past few decades, as well as how they will change in the coming decades.

This chapter will attempt to present the multiple roles the modern Warfighter plays and the complex circumstances he/she faces. The diversity of Warfighter demographics, missions, working environment, and the tremendous physical, physiological, and psychological factors that are encountered are introduced and briefly described and will be further explored in the chapters that follow.

Current and Changing Roles

Whether an infantryman, helicopter pilot, tank mechanic, computer specialist, photographer, or cook, each member of the U.S. military is, first and foremost, a Warfighter. Historically, the primary job of every Warfighter has been to fight and defeat the enemy. However, the ending of the Cold War, the ever-changing role of the U.S. in world affairs, and the aftermath of September 11, 2001, have each expanded and added to the duties, tasks, and functions of today's Warfighter.

---

[1] Warfighter is a term used to describe all military personnel trained to engage in combat operations.

Figure 1-1. A representative HMD, the Integrated Helmet and Display Sighting System (IHADSS), used on the U.S. Army's AH-64 Apache attack helicopter.

In addition to being a combatant, today's Warfighter is at times called upon to be a peacekeeper, counter-drug specialist, anti-terrorist operative, humanitarian assistant, and disaster relief worker (Murray, 2001). In addition to these expanding roles, the Warfighter has a specific occupational area of expertise (i.e., job classification). Within the Army, these are referred to as Military Occupational Specialties (MOSs). Examples are Combat Engineer, Radar Repairer, Artillery Mechanic, and Accounting Specialist. Each of these specialties requires a certain knowledge base and mastery of a set of occupational skills (U.S. Department of the Army, 1999). For virtually all MOSs, the Warfighter may be required to assume the role of teacher/trainer, passing on accumulated knowledge and skills to other Warfighters. The other military services employ similar nomenclatures.

Role as a peacekeeper

No role for the Warfighter appears more opposite to the primary role as combatant than the role of peacekeeper. The U.S. has invested significant military, political, and economic resources in conducting operations following worldwide conflicts and civil unrest (Dobbins et al., 2003). While this role often is thought of as a recent phenomenon, the U.S. military has embarked on a number of such missions, spanning over 60 years. From post WW-II stabilization and reconstruction in Germany and Japan; through Korea in 1950; in Bosnia and Herzegovina (Bosnia) in 1995; and to Afghanistan and Iraq in 2002, the U.S. military, in the most recent century, has spent more time in the role of peacekeeping than that of combatant (until the most recent engagements in Afghanistan and Iraq) (North Atlantic Treaty Organization, 2001).

However, it was not until President Clinton established the U.S. Army Peacekeeping Institute (PKI) in 1993 within the U.S. Army War College (USAWC), located in Carlisle, Pennsylvania, following the catastrophe in Somalia, which the U.S. military sought to study and understand how Warfighters performed when sent in to carry out non-combat operations. (In 2003, the U.S. Army PKI was transformed into the U.S. Army Peacekeeping and Stability Operations Institute [PKSOI].)

The peacekeeping role borders that of law enforcement in many of its tasks. As such, it abounds with applications for HMDs. HMDs can provide tactical information and communications and can enhance situation awareness.

Role as a counter-drugs enforcer

The use of the military to counter illicit drug trafficking has been in effect at least since the mid to late 1980s, operating under the authority of the Posse Comitatus Act (1981 amendment) (Ahart and Stiles, 1991; Cathcart, 1989; Dickens, 1989; Simpson 1992; U.S. Army War College, 1988; U.S. Congress, 1989). Such efforts have been focused within, but not limited to, the South and Central Americas. In addition, units of the National Guard have been conducting surveillance missions in Latin America since the early 1990s (Haskell, 2004). National Guard counter-drugs task forces, such as in the California and Tennessee National Guard, are comprised of members of both the Army and Air National Guard. These guardsmen conduct observation missions within the United States in remote, rural, and semi-rural areas for long periods of time and depend heavily on HMDs with integrated night vision sensors that allow effective nighttime operation.

The use of military troops in anti-drug operations has not been without criticism, with both military and civilian leaders expressing concern about blurring the distinction between military and police authority (Marshall, 1988; Murray, 2001). However, it is more than likely that this role for the military will increase, not decrease, in the future.

Role as an anti-terrorist operative

Since the attack on the World Trade Center, September 11, 2001, military operations, primarily in Afghanistan and Iraq, have greatly expanded the role of the Warfighter into the area of antiterrorism. Of all of the expanded roles discussed, that of antiterrorist operative is the closest to that of the Warfighter’s fundamental role of combatant.

The pursuit of Al Qaeda and other terrorist networks by U.S. military forces is worldwide. In addition to the highly publicized search for terrorist cells and insurgent personnel in Afghanistan and Iraq, the U.S. military is establishing programs throughout the world for the purpose of training local troops in methods to prevent the emergence of Al Qaeda in poor, rural areas. In one such program, the Pentagon is planning to train thousands of African troops as battalions equipped for extended desert and border operations and to link the militaries of different countries with secure satellite communications. This initiative, with proposed funding of $500 million over seven years, encompasses the countries of Algeria, Chad, Mali, Mauritania, Niger, Senegal, Nigeria, Morocco, and Tunisia (Tyson, 2005). The Pentagon also is assigning more military officers to U.S. embassies around the world, thereby hoping to increase intelligence gathering capabilities.

Special Forces units, as well as other Army units, can employ HMDs in search-and-destroy surveillance and search-and-rescue operations. As in other applications, HMDs can present tactical information, provide for communications, and increase situation awareness.

Role as a humanitarian aid and disaster relief provider

The military has long been involved in providing humanitarian assistance, both at home and abroad. Such assistance ensures the delivery of life saving and life sustaining aid to civilian populations. Humanitarian operations encompass a wide-range of missions, including sea search and rescue; refugee assistance and disaster relief; and the provision of food, medical supplies, and services (Juda, 1993).

Recent worldwide disasters, such as the December 2004 earthquake and resulting tsunami in the Indian Ocean region, have brought to the forefront the role that the U.S. Warfighter plays as a provider of humanitarian aid. Approximately 13,000 U.S. Navy, Marine Corps, Army, Air Force, and Coast Guard personnel were involved in the relief efforts following this disaster. By January 2005, military relief operations had flown over 400 missions and delivered 316,664 pounds of water, 135,102 pounds of food, and 8,246 pounds of medical supplies (U.S. Department of State, 2005).

Within the U.S., there were massive efforts by both the Active military and National Guard in response to the 2005 hurricanes Katrina and Rita – massive humanitarian assistance logistics (food, water and medical supplies) as well as all the search and rescue via rotary wing platforms from the Army, Navy and Coast Guard. Over 70,000 Active-duty and National Guard personnel were deployed either on the ground, in the air, or aboard ships supporting relief operations. Twenty U.S. Navy ships, 346 helicopters, and 68 fixed-wing aircraft were deployed to the area (Hiatt 2006).

The U.S. military has had a continuous worldwide presence, from Afghanistan to Uzbekistan, in humanitarian endeavors. These Warfighters, turned humanitarians, have delivered food and clothing, rebuilt infrastructure (e.g., orphanages, schools, and bridges), donated money for supplies and equipment, and worked side-by-side with local civilians to rebuild communities devastated by war or natural disaster (Barnes, 1989; Covey, 1992; Foster, 1983; Harrison, 1992; Jones, 1991; Kelly, 1992; Miles, 1991; Nalepa, 1993; Shotwell, 1992; Stackpole et al., 1993; Sutton, 1992).

It may be in this humanitarian role that the application of HMDs seems most out of place. However, when delivering food or rebuilding a school, the military relies on organization, planning, and communicating. Presentation of information via HMDs can assist in the performance of these functions, both at the command and control level as well as in the field.

The Demands of Combat

In spite of these ever-expanding roles, the primary purpose of a Warfighter is to engage in combat. Combat is defined as an engagement fought between two military forces. However, when such engagements are considered in the most personal manner (e.g., hand-to-hand fighting), this description falls far short of truly defining the essence of combat. The so-called “rigors of combat” are broad in scope - being physical, mental, and psychological in nature. It has been well recognized that the added uncertainty and stress of combat have a major effect on both physical and cognitive performance (Lieberman et al., 2002; Nindl, 2002; U.S. Army Center for Health Promotion and Preventive Medicine, 2005).

The physical rigors of combat, which include physical exertion, endurance, and overcoming the effects of extreme temperature, fatigue, and dehydration, are intended to be mitigated by intense physical training. Some common demands that combat places on Warfighters include marching long distances bearing heavy loads and still being able to function effectively; moving quickly and evasively under fire; carrying wounded to safety; setting up heavy weaponry; handling large-caliber ammunition for extended periods; climbing walls, cliffs, and other high obstacles; operating in physically confined spaces; and performing field maintenance on aircraft or heavy equipment (United States Marine Corps, 1998).

Just as critical to combat readiness are the mental and emotional states of the Warfighter. The competitive and combative spirit of the Warfighter has a tremendous impact on mission performance. Natural physical fear directly leads to cognitive degradation as well as physical fatigue, and these effects must be lessened by instilling confidence in the Warfighter — confidence in his performance, his command structure, and his equipment. The modern Warfighter is the most technologically advanced in the history of warfare. To make the most of this technology, equipment provided to the Warfighter must be reliable and useful and must enhance, not degrade performance. Through design and training, the operation of equipment must be second nature. The equipment must become a natural extension of the Warfighter.

The conditions under which military missions are or will be conducted will continue to vary with respect to the physical environment, the number of tasks, and the task complexity (National Research Council, 1997). In all situations, the Warfighter must be able to move, communicate, engage the enemy, and survive. It is the purpose of systems such as HMDs to offer the possibility of increasing individual Warfighter and unit performance.

Uniqueness of the Tri-Service Military Communities

While certain similarities exist, each of the four branches of the military has a distinctive operational environment and role. This individuality is defined by distinctiveness in mission, personnel and vehicles, and operating environments.

The U.S. Air Force's mission statement is to fly and fight in "air and space." The main component of the Air Force's arsenal is fixed-wing aircraft. With over 7,000 aircraft in service (Figure 1-2), the Air Force provides six distinctive core capabilities:

- Air and space superiority
- Global attack
- Rapid global mobility
- Precision engagement
- Agile combat support
- Information superiority

Figure 1-2. U.S. Air Force aircraft: C-17 Globemaster Tactical Transport (top center), F-16 Falcon Fighter (bottom left), and B-1B Lancer Bomber (bottom right). (Source: U.S. Combat Camera)

The latter capability emphasizes the ability of commanders and airmen to keep pace with information and to incorporate it into evolving plans of action.

The U.S. Navy operates in excess of 280 ships and 4,000 aircraft and is responsible for naval operations on the Earth's seas and oceans (Figure 1-3). As of January 2004, ship classes of the U.S. naval fleet included: Aircraft carriers, amphibious assault ships, amphibious transport docks, dock landing ships, submarines, cruisers, destroyers, frigates, and battleships. Naval aircraft include both fixed- and rotary-wing (helicopters) aircraft. These aircraft operate from the land as well as from ocean-going ships. Navy Warfighters have the most diverse operating environments, having to perform tasks on land, in the air, and on and beneath the water. The Navy also has expanded its harbor defense forces in response to the war on terrorism. The main components of Naval Harbor Defense include:

- Inshore Boat Units (IBUs)
- Mobile Inshore Undersea Warfare Units (MIUWUs)
- Special Boat Units (SBUs)

The Navy also has special warfare operatives, the "Navy Seals." Their primary purpose is to engage in "special activities other than war."

The U.S. Army is the branch of the U.S. armed forces that has primary responsibility for land-based military operations. The Army is highly focused on mobility and, therefore, maintains a diverse inventory of vehicles. Vehicle types include armored, transport and supply, and rotary-wing aircraft (Figure 1-4). Component-wise, the Army possesses the greatest proportion of combat personnel within the U.S. military forces. Within the Army, infantry Warfighters make up the largest contingent of combat personnel.

The U.S. Marine Corps serves as a versatile combat element and is adapted to a wide variety of combat operations. The Marine Corps possesses ground and air combat elements but relies upon the U.S. Navy to provide sea combat elements. A major mission of the Marine Corps is amphibious assault, the attack of an objective located on land by a force attacking from the sea. Landing craft are used to transport troops from ships to land. It is perhaps the most complex military maneuver in the history of warfare. Marines consistently use air, ground, and sea elements of combat together. Vehicles used by the Marines include fixed- (AV-8B Harrier), rotary-wing (AH-1Z Super Cobra and CH-53E Super Stallion), and hybrid (MV-22 Osprey) aircraft, plus assault amphibian vehicles (AAVP7A1) (Figure 1-5). Marines sometimes are employed to enter and hold an area until a larger military force can be mobilized.

Warfighter Demographics

From a human factors engineering (HFE) perspective, it is important to have an understanding of the users of a technological system or device. Previously, when only physical attributes of the user were considered, user anthropometry was most important. In aircraft design, arm and leg reach, torso height, etc., have been and still are important parameters. During the introduction of HMDs, a number of head and facial anthropometry measures were added to the list (Rash, 2000). These include the bizygomatic breadth (the maximum horizontal breadth of the face, between the zygomatic arches), eye inset (the distance between the supraorbital notch [eyebrow] and the cornea of the eye), the disparity between the two eyes, etc.

Now, as we wish to bring to the forefront perceptual and cognitive issues, it is important to expand our knowledge of the user population. In this section, the demographics of the Army user community are explored as a subset of the military user population, with comparisons to the other U.S. military services where available.

After the fall of the former USSR and the end of the Cold War, both the Federal Government and the Army Leadership realized that the structure of the 1960s' Big Army, designed to fight a protracted land war based in Europe, was no longer required, and a major reduction in active-duty forces was undertaken. This downsizing, occurring over the years 1992-1999, is reflected in Figure 1-6, which depicts U.S. Department of Defense active-duty military personnel strength levels for fiscal years (FYs) 1950-2002 (U.S. Department of Defense, 2005).

Although the size of the active-duty component of the Army has decreased since the mid 1980s from around 775,000 to about 490,000 today (approximately a 35% decrease), the distribution of ranks (officers, warrants, and enlisted) has remained fairly stable. However, changes in gender and ethnicity distributions have occurred.

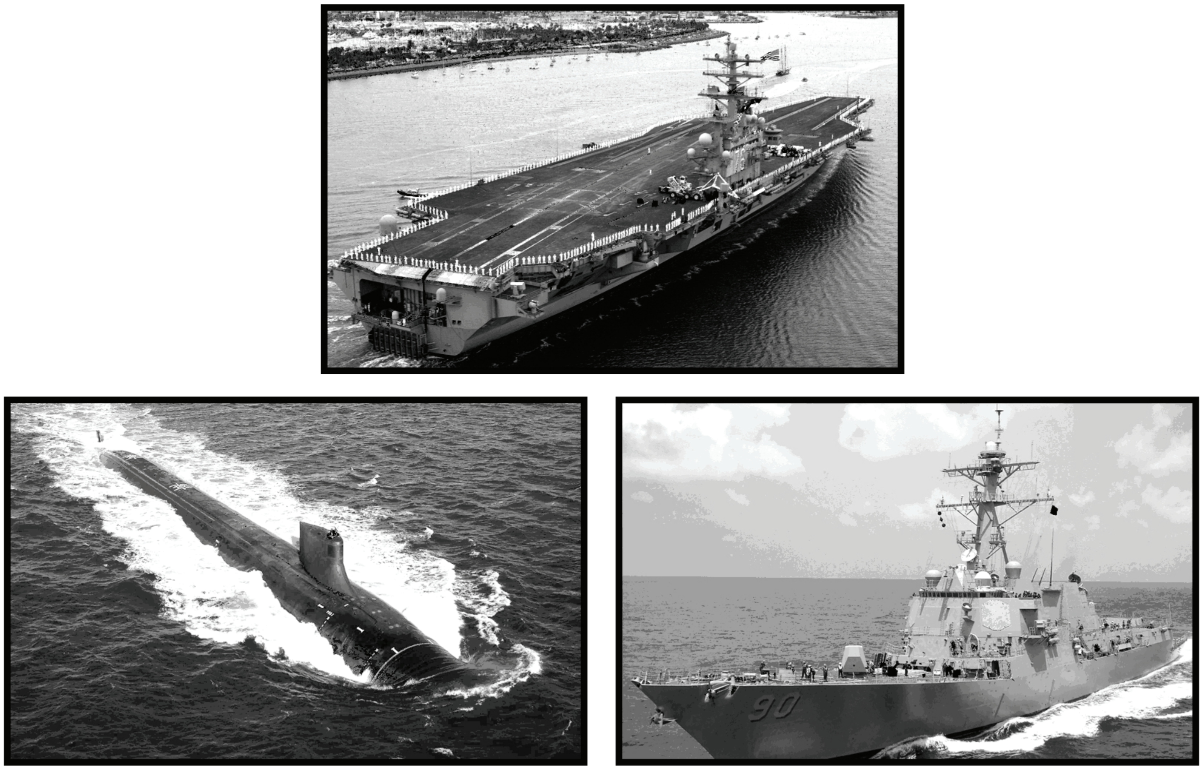


Figure 1-3. U.S. Navy vehicles: CVN-76 Ronald Reagan Aircraft Carrier (top center), SSN-23 Jimmy Carter Submarine (bottom left), and DD-356 Destroyer (bottom right). (Source: U.S. Combat Camera)

Figure 1-4. U.S. Army vehicles: AH-64D Apache Attack Helicopter (top center), M1A1 Abrams Main Battle Tank (bottom left), and M2 Bradley Infantry Fighting Vehicle (bottom right). (U.S. Combat Camera)

Figure 1-5. U.S. Marine Corps vehicles: MV-22 Ospry (top center), CH-53E Super Stallion (bottom left), and AAVP7A1 Amphibious Assault Vehicle (bottom right). (Source: U.S. Combat Camera)

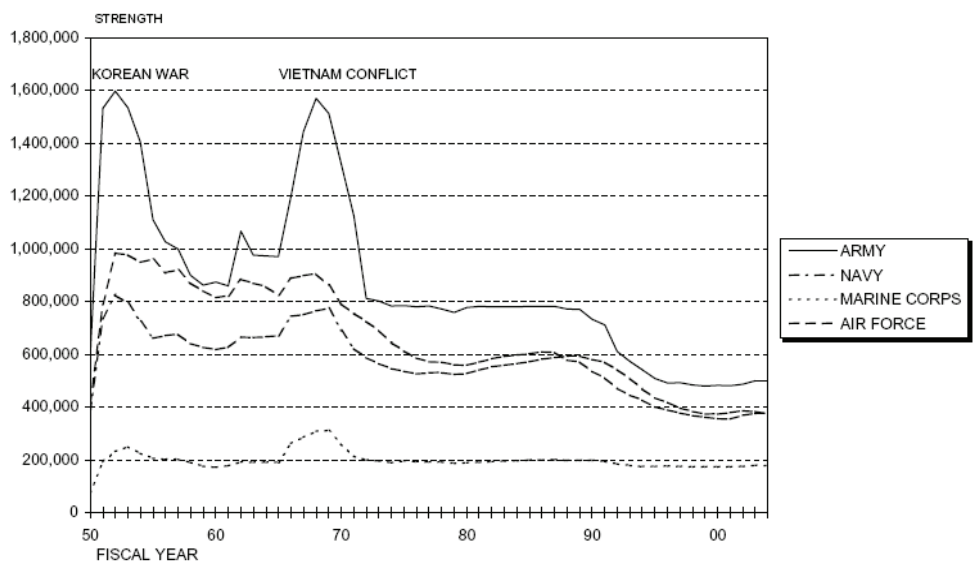


Figure 1-6. U.S. Department of Defense Active-Duty military personnel strength levels for fiscal years (FY) 1950-2002 (Source: Department of Defense, 2005).

Gender

It has been suggested that men and women have some basic behavioral differences, differences that may be based on dissimilarities between the male and female brain. As an example, it has been suggested that women are superior in certain language abilities, while men are superior in certain spatial abilities. Studies have documented an "array of structural, chemical and functional variations" between the brains for the two genders (Cahill, 2005). These studies have, in turn, highlighted gender differences, or biases, in cognition and behavior. Areas in which these differences exist include memory, vision, audition (hearing), and response to stress, all of which are factors that influence performance with HMDs. While within gender performance differences may exceed across gender differences, it still may be useful to define the gender breakdown within the potential user population.

The overall gender makeup of the active-duty Army has undergone significant changes since the end of the Cold War. While the total number of officer and enlisted women on active-duty has been rather constant, statistics show that the proportion of women has increased from approximately 10% to 15% over the period just preceding the end of the Cold War to 2003. This increase has been reflected in all ranks (Table 1-1) (U.S. Department of the Army, 2005).

Analogous data for the ten-year period only from 1993-2003 also show increases in the proportion of women from approximately 5% to 7% for the U.S. Marine Corps, from 16% to 20% for the U.S. Air Force, and from 12% to 15% for the U.S. Navy.

Table 1-1.
Proportion of U.S. Army active-duty women (1983-2003).
(U.S. Department of the Army, 2005).

Females Active Army	**1983 (75,548)**	**1993 (70,797)**	**2003 (74,907)**
TOTAL	**9.8%**	**12.5%**	**15.2%**
Officers	10.2%	14.2%	16.4%
Enlisted	9.9%	12.4%	15.2%
Warrants	1.3%	3.8%	7.1%

Race/Ethnicity

As with gender, there is reason to suspect that cognitive performance may differ with ethnicity. Such differences may have social, cultural, and economic causes (Rushton and Jensen, 2005). Therefore, as with gender, it may be useful to define the ethnic breakdown within the potential user population.

Over the same period considered for gender makeup, there has been a gradual shift in the racial/ethnic makeup of the active-duty military. For the Army, there has been a decreasing trend in the proportion of White and Black Warfighters over the period FY83-03; in contrast, there has been an increasing trend for Hispanic and Asian Warfighters over the same period (Table 1-2).

Within the U.S. Air Force, the proportion of Black Warfighters has been rather constant at approximately 15%, while the Hispanic proportion has increased only slightly from approximately 3% to 6%. For the U.S. Navy, both the Black and Hispanic Warfighter proportions have increased, from approximately 16% to 19% and from 7% to 9%, respectively. For the U.S. Marine Corps, Black Warfighter proportions have decreased from 17% to 13%, while Hispanic proportions have increased from 8% to 13%.

Table 1-2.
Ethnicity proportions of U.S. Army active-duty Warfighters (FY83-03).
(U.S. Department of the Army, 2005).

	FY83	FY93	FY03
White	64.0%	62.4%	59.3%
Black	28.3%	27.6%	24.0%
Hispanic	3.8%	4.7%	9.9%
Asian	1.3%	2.0%	3.5%
Other	2.6%	3.3%	3.3%

Education level

The modern Warfighter, by general standards, is well educated (Table 1-3). Almost 98% are high school graduates, and at least 96% of each of the U.S. military service's commissioned officers have earned college degrees (U.S. Department of Defense, 2004). For enlistment purposes, the military breaks education into three overall categories: Tier 1, Tier 2, and Tier 3. Tier 1 includes high school graduates or equivalent. Tier 2, known as Alternative Credential Holders, must achieve a minimum set score on Armed Forces Qualification Test (AFQT). The final tier (Tier 3) includes non-high school graduates, i.e., individuals who are not attending high school and are neither high school graduates nor alternative credential holders. However, the military services rarely accept a Tier 3 candidate for enlistment.

Table 1-3.
Educational level of U.S. active-duty military personnel.
(U.S. Department of Defense, 2004)

Education	U.S. Army	U.S. Navy	U.S. Marine Corps	U.S. Air Force
	Commissioned Officers			
College graduate[2]	98.7%	96.0%	97.2%	97.3%
High school graduate	100.0%	99.5%	97.9%	97.6%
	Warrant Officer			
College graduate[2]	31.5%	23.1%	14.4%	----
High school graduate	100.0%	100.0%	100.0%	----
	Enlisted			
College graduate[2]	5.3%	2.8%	1.3%	5.0%
High school graduate	98.5%	98.2%	99.2%	99.9%

While no direct correlation between education and cognitive skills is claimed, a higher level of education is considered to be an attribute that is advantageous in the use of technically complex systems.

[2] A 4 -year degree.

Age

The age of active-duty personnel can range from 17 [3] to 60. An age distribution based on 2004 data is provided, by gender, in Table 1.4. There is a relatively high correlation between the male and female distributions. The median age (based on reported data only) is 26 and 25, for males and females, respectively.

The U.S. military is relatively young. Approximately 47% of female and 41% of male active-duty personnel is under 25 years of age. Age has been shown to be a factor in cognitive performance in complex and simultaneous task environments. Becker and Milke (1998) cite that for the air traffic control occupation, where the ability to handle simultaneous visual and auditory input is critical to success, there is a strong positive relationship between age and job performance.

Table 1-4.
Age distribution of U.S. active-duty military personnel.
(Expressed in percent)
(U.S. Department of Defense, 2004)

Age/ Gender	**19 or under[3]**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50+**
Female	8.96	37.74	21.70	12.27	9.45	5.66	2.36	0.94
Male	7.74	33.28	20.35	14.33	12.43	7.58	2.47	0.82

Note: 0.99% male and 0.94% female not reported.

Service components

All of the demographic statistics presented have been for active-duty personnel. However, in addition to the Active component of the military branches, there also is the Reserve component. The U.S. Army also has the National Guard component; the U.S. Air Force has the Air National Guard. For the Army, the total Army personnel strength at the end of FY04 was 1,041,340, with the active component (494,291) representing 47%, the reserve component (204,131) representing 20%, and the National Guard component (342,918) representing 33%. In peacetime, Reserve and National Guard personnel are generally confined to training operations. However, with Operation Enduring Freedom and Operation Iraqi Freedom, the Department of Defense has been relying heavily on the fielding of these components in combat operations. Demographic statistics for these components, for all military branches, are prepared annually by the Department of Defense's Washington Headquarters Services, Information Technology Management Directorate, Arlington, Virginia, and can be accessed via their website, http://www.dior.whs.mil/

Army Transformation Plan

As the U.S. military moves into the 21st century, it is adopting a new vision and a new model for its structure and operation. The form of warfare envisioned during the Cold War and the type of Armed Forces previously built to fight that war have been determined to be outdated, cost prohibitive, and ineffective. Today's and tomorrow's Armed Forces must be leaner and more responsive. A major principle of the plan to achieve this "transformation" is to depend more heavily on technology as a "force multiplier." HMDs are one of the many technologies being employed to achieve this goal.

Since the Army represents a major portion of the U.S. military personnel (approximately 35%, compared to 26% each for the Navy and the Air Force, and 12% for the Marine Corps), it may be instructive to look at the

[3] Age of 17 is the youngest enlistment age (with parental consent).

Army's ongoing program to restructure itself into a leaner, more technology-based organization. This restructuring is currently referred to as the "Army Transformation Plan."

For the latter half of the 20th century, the U.S. Army has been organized and equipped in preparation of fighting the large armies of the Soviet block. With the collapse of the Soviet empire and the end of the Cold War, the new challenges became multiple flashpoints scattered around the globe, e.g., Haiti, Somalia, Bosnia, Kosovo (Steele, 2001). To meet the changing demands on the future Warfighter, the Army is redefining itself via a transformation process that will bridge two decades.

The basic tenets of the transformation, while subject to modification, include (Murray, 2001; Steele, 2001):

- The future Army must become more responsive.
- A deployment capability plan must be able to put a combat-ready brigade anywhere in the world within 96 hours, a full division within 120 hours, and five divisions within 30 days.
- Equipment designated for the new Army will have increased capabilities and do much of the routine processing of data.
- The planned transformation must produce an Army that is more strategically, operationally, and tactilely mobile than current forces.

The Army Transformation plan provides for three forces: the Legacy force, the Interim force, and the Future force (formerly Objective force) (Murray, 2001). These three forces follow separate paths during the first decade of the transformation, finally merging into the new Army sometime near 2020 (Figure 1-7). The Legacy force consists of the Army's current heavy and light forces, e.g., the M1 Abrams tanks and the M2/M3 Bradley fighting vehicles. The Interim force improves on the capability of the Legacy force. It will consist of re-equipped heavy and light brigades that will be capable of faster deployment. These new units will be referred to as the Interim Brigade Combat Teams. The Future force will be the culmination of two decades of research and development. This force will possess a greater responsiveness, deployability, agility, and versatility than the current force.

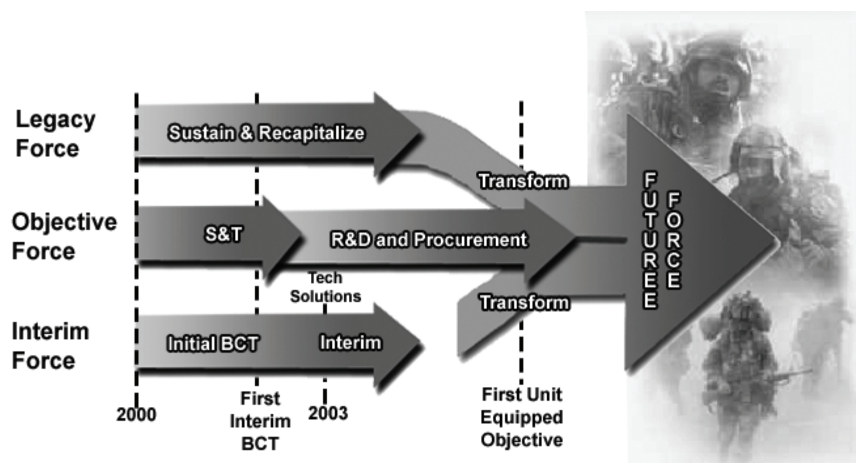


Figure 1-7. The Army Transformation Campaign Plan (Adapted from Murray, 2001).

The Future Force concept exploits the vast opportunities made possible by the expected capacity to quickly collect, organize, and distribute battlespace information. Data from multiple sources, e.g., sensors and databases, will be available to the Warfighter. It is critical that technologies such as HMDs employed in systems inherent to the Future Force concept be Warfighter-centered so as to enhance cognitive functions. Science and technology (S&T) and research and development (R&D) will play major roles in the Future Force design.

Under the Army's Future Force concept, the Warfighter is referred to as the Future Force Warrior (FFW). The concept seeks to create a lightweight, overwhelmingly lethal, fully integrated individual combat system, including weapon, head-to-toe individual protection, networked communications, Warfighter-worn power sources, and enhanced human performance, demonstrating "optimized cognitive and physical fightability." The Warfighter will become a "system-of-systems." One integral system is a "head-borne vision enhancement" system (an HMD) that provides fused I^2/IR sensor imagery (U.S. Army Natick Soldier Center, 2004).

An important challenge for the military planners of the Interim and Future Forces will be the development of Warfighter training that emphasizes intellectual development, flexibility, pragmatism, and cognitive decision making. Decision making is an inseparable component of all cognitive activities. The 21st century Warfighter must be able to select the critical information from a host of data made available by new technology and must be technically competent in the operation of such technology, such as the HMD, which will play a major role in the presentation of the information (Murray, 2001).

Battlespace Information, Information Superiority, and Network Centric Warfare (NCW)

With the explosion of imaging sensor technologies, the arrival of unmanned aerial systems (UASs) in the battlespace, the development of miniature, low-power displays, and the ability to link high-quality data and images via high-speed communication systems, the battlespace is saturated with information, with virtually all of it being made available to the individual Warfighter. It is crucial that Warfighters and their commanders receive this continual flow of information in order to achieve information superiority (Matson and DeLoach, 2003).

Information superiority in the battlespace means having an advantage in acquiring, processing, and distributing information on the status and location of your Warfighters and the enemy's Warfighters. This superiority also results in an uninterrupted flow of battle information while denying an enemy's ability to have the same information (Cohen, 1999).

Garstka (2000) suggests that to understand how information impacts our ability to conduct military operations it is necessary to consider three domains—the physical domain, the information domain, and the cognitive domain (Figure 1-8). The physical domain consists of the material battlespace where the intent is to exert influence or control in the situation. It encompasses the environments of land, sea, air, and space and is where the physical platforms and the communications networks that connect them are located. This is where the information resides; it is where information is created, shaped, and shared. It is the domain that makes possible the distribution of information among Warfighters.

The information domain is the domain where the data exists. It is where the data are created, manipulated, and shared. It is the domain that facilitates the distribution of information between Warfighters. It is the domain where the command and control of modern military forces is exercised, where commander's intent resides. In the key battle for information superiority, the information domain is "ground zero." Information Superiority is a condition in the information domain, a condition that is created when one adversary is able to establish the superior information state (Garstka, 2000).

The third domain is the cognitive domain, which is in the brain of the Warfighter. This is where perceptions, awareness, understanding, beliefs, and values reside and where decisions are made. Most importantly, this is the domain where most battles and wars are won and lost. The cognitive domain is where concepts of leadership, morale, unit cohesion, level of training and experience, situational awareness, and public opinion are found. This is the domain where an understanding of the battle plan, doctrine, tactics, techniques, and procedures influences decision-making (Garstka, 2000).

All of the contents of the cognitive domain are filtered by human perception. This filtering is defined by the Warfighter's individual worldview, the body of personal knowledge the Warfighter brings to the situation, experience, training, values, and individual capabilities (e.g., intelligence, personal style, perceptual capabilities, and cultural background). While there is one reality (physical domain), which is transformed into selective data, information, and knowledge by the various sensor and imaging systems in the battlespace, each Warfighter has his/her own perception of reality. The military, through training and shared experiences, strives to mold these individual perceptions and resulting cognitive behavior into a similar collective perception of reality (Garstka, 2000).

An important aspect of this reality is situation awareness. Situation awareness refers to the Warfighter having a global awareness of the tactical situation and of his/her status within the situation. The components of the situation include mission purpose, mission constraints, environmental factors, available resources, and interaction with other Warfighters and Warfighter elements. Alberts et al. (2001) discuss situation awareness within the context of the cognitive domain. Maintaining situation awareness in the presence of the high information flow in the modern battlespace requires considerable cognitive function. If cognitive function is compromised due to any of a host of factors, situation awareness also becomes compromised. Further, this awareness must be shared and yet must avoid both cognitive illusions and "groupthink."[4]

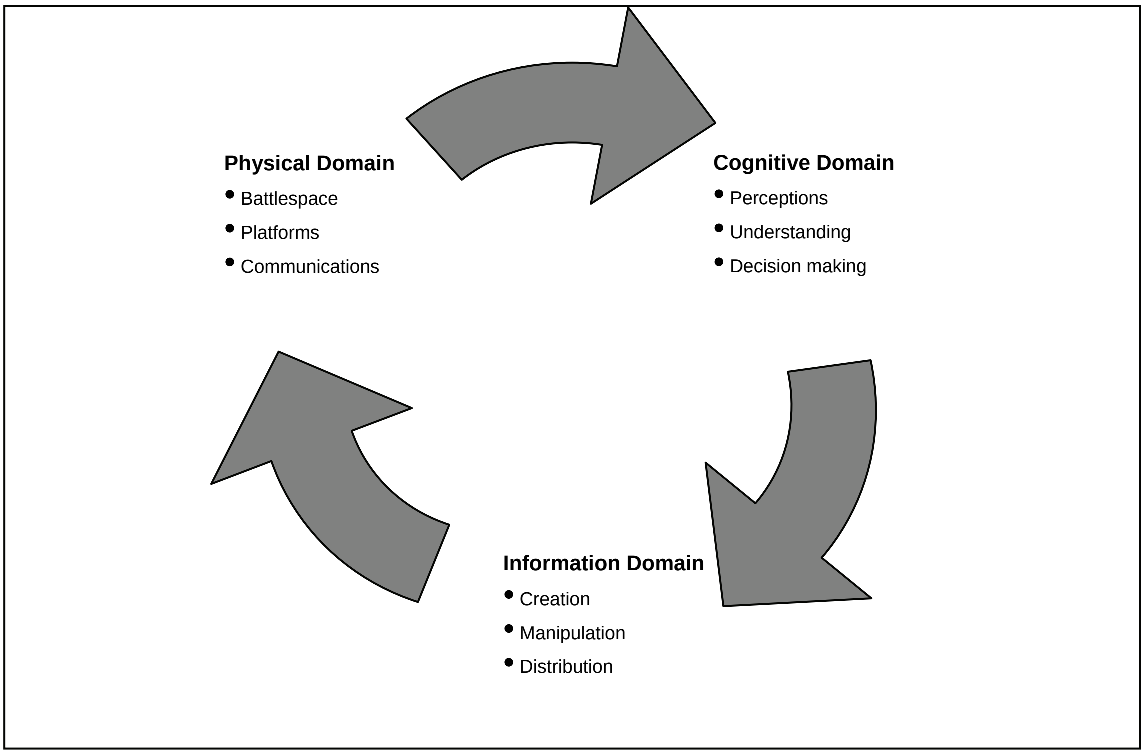


Figure 1-8. The three domains that aid in the understanding of how information impacts the conduct of military operations. (Suggested by Garstka, 2000)

The relatively modern concept of having the ability for geographically dispersed Warfighter forces (individuals, teams, and higher order structured units) to create and maintain a high level of shared battlespace awareness that can be exploited via self-synchronization and other operations to achieve strategic success is known as Network Centric Warfare (NCW) (Alberts et al., 1999). NCW is about human and organizational behavior in the battlespace. It mandates adopting a new way of thinking—network-centric thinking—and applying it to military

[4] The term "groupthink" was suggested by the psychologist Irving Janis in 1972 to describe a process by which a group can make poor decisions that result from each member of the group attempting to conform to what they believe to be the group consensus.

operations. NCW focuses on the combat power that can be generated from the effective linking or networking of the warfighting elements, converting the position of information superiority into military action.

The Physical Environment

Most professions and occupations have a single working environment whose characteristics define a set of typical surround conditions within which the worker performs tasks or duties. This is not the case for the Warfighter. The Warfighter's physical environment runs the gamut from benign to severe. Physical factors that could affect both human and equipment performance must therefore be taken in consideration include operation in confined spaces, at high altitudes, in reduced illumination levels, in adverse weather conditions (e.g., rain, sleet, snow, fog, etc.), in smoke and other battlespace obscurants, in regions of extreme heat or cold, etc.

Until the recent trend to shift to performance specifications and to adopt more off-the-shelf technology, the military was well-known for establishing rigid specifications. Referred to as military specifications (MIL-SPECs) and military standards (MIL-STDs), these specifications precisely defined the operational environmental requirements of newly developed devices and systems. These publications are still widely used and routinely referenced in many performance specification documents. These specifications and standards typically address such environmental factors as temperature, altitude, solar radiation, humidity, rain, sand, dust, vibration, shock, salt, fog, fungus, etc.

The only dedicated military specification or standard for HMDs is MIL-A-49425 (U.S. Department of Defense, 1989) for the Aviator's Night Vision Imaging System (ANVIS), an image intensifier (I^2) tube-based HMD. However, there are a number of such documents which are directly or indirectly applicable. These include, but are not limited to, MIL-STD-461E (Electromagnetic Emission and Susceptibility Requirements for the Control of Electromagnetic Interference), MIL-STD-1295 (Human Factors Engineering Criteria for Design Criteria for Helicopter Cockpit Electro-Optical Displays Symbology) (U.S. Department of Defense, 1999), and MIL-STD-1472 (Human Engineering Design Criteria for Military Systems, Equipment and Facilities) (U.S. Department of Defense, 1981). The latter two examples are specifically cited for their guidance in addressing human factors issues. The U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL, also has published a performance-based design guide, *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft* (Rash, 2000).

However, while these specifications and standards have been very effective in ensuring that systems be designed to operate properly in harsh military environments, they do not guarantee that operational user performance in these environments will not be affected. For example, today's Warfighters deploy worldwide. They can be required to operate in regions of extreme heat or cold for long periods. Such temperature extremes can be encountered both in the outside, exposed environments, as well as in the inside, enclosed spaces of ships, tanks, aircraft, etc. Working temperatures in excess of 100°F (38°C) have been recorded in tank cabins and aircraft cockpits. Both heat and cold temperature extremes impact not just system performance but user performance as well. In such regions, the physiological conditions of heat and cold stress may be present. In extreme conditions, injuries of heat exhaustion or heat stroke and frostbite or hypothermia can result.

The physiological effects of heat stress can include fatigue, nausea, headache, and fainting. But, heat stress also can reduce mental performance. Even moderate heat environments take a toll on performance. Tasks that require attention to detail, concentration, and short-term memory will become more difficult. Heat stress slows reaction and decision time. Routine tasks are done more slowly. Vigilant task performance is degraded (U.S. Departments of the Army and Air Force, 2003). It has been suggested that impairment of cognitive performance by heat stress is a function of the resulting internal body temperature during exposure (Hancock, 1981).

An excellent summary list of guidelines for the impact on Warfighter performance is provided by Johnson and Kobrick (2001) and includes:

- Although not directly affected by heat, vision likely will be impaired by secondary factors such as sweat running into the eyes and moisture obscuring optics and lens surfaces.

- Visual distortions due to heat, such as mirages, optical illusions, shimmer and glare, can reduce spatial vision.
- Performance of some visual tasks, such as rifle aiming and distance judgment, can be degraded.
- Equipment controls can interfere with efficient manual operation when they become too hot to handle comfortably.
- Sweating can cause headgear and headphones to become unstable and slide on the head, compromising hearing, vision, and the performance of other tasks.
- Tasks requiring sustained attention, such as sentry duty, watch keeping, and instrument monitoring, will be aversely affected.
- Complex mental tasks, such as mathematical reasoning and decoding of messages, can deteriorate in heat above 90ºF (32ºC) after about 3 hours.
- Continuing heat exposure causes progressive motor instability, leading to impaired steadiness and manual dexterity.
- Target tracking, in which the Warfighter must judge differences in continuous target alignment, can degrade.
- Simple tasks are less affected by heat than are highly complex tasks. Moderately complex tasks tend to be the most resistant to heat effects because they tend to sustain attention while placing only moderate demands on the Warfighter's overall performance.
- Multiple tasks (i.e., two or more tasks being performed concurrently) are more affected by heat than any of the same tasks performed individually.
- Discomfort reactions are widely different among individuals, and heat acclimatization and experience greatly influence degrees of discomfort. High humidity in tandem with conditions of heat compound discomfort.
- Symptoms of heat illness seriously degrade Warfighter performance, and symptom intensity varies widely among individuals.

Cold stress can have equally degrading effects on performance. Physiological effects can include uncontrollable shivering, slow and irregular heartbeat, low blood pressure, fatigue and drowsiness, and pain in the extremities. Cognitive effects for cold stress have been much less investigated than effects for heat stress but include memory lapses and incoherence. In the battlespace, individuals working with computers or other skills requiring fine motor control and good decision-making skills have been shown to be especially vulnerable to the effects of even moderate cold stress (Pozos and Danzl, 2001).

A number of studies have documented decreases in visual vigilance task performance (Hoffman, 2001). However, substantial decrements are likely to be present only during rapid changes in core body temperature, such as with sudden water immersion.

While simple reaction time seems to be relatively unaffected, decrements in cognitive function due to cold increase with task complexity. To offset the effects of cold stress, it might be necessary to pre-divide complex tasks into multiple subtasks.

Additional Operational Factors

Physical factors of the environment are not the only ones that must be considered in the design, development, and fielding of a new device or system. A number of additional factors must be addressed to ensure that user performance with the device or system is optimized. Rash (2004) developed a list of adverse operational factors that should be considered for their possible impact on operational performance with advanced display concepts, to include HMDs. The list contains 19 generalized factors categorized as physical/environmental, mechanical, physiological, sensory, and psychological in nature (Table 1-5). This list of factors should not be considered exhaustive.

Within the mechanical category, the HMD may be worn in conjunction with corrective eyewear and/or some type of chemical, nuclear, biological (NBC) mask. Rash et al. (2002) states that limited mechanical clearance (referred to physical eye relief) between the optics of an HMD and add-on devices, such as corrective spectacles, oxygen masks, NBC masks, etc., can impact fit and the ability to achieve the full field-of-view of the HMD imagery.

Optical alignment problems that can affect targeting tasks and associated decisions can be introduced when the HMD is worn in combination with one or more of these devices. This effect arises from the induced prismatic deviation caused by the presence of multiple optical surfaces.

Table 1-5.
Adverse operational factors to be considered for impact on operational performance with advanced display concepts.

Category	Factor
Physical/Environmental	Temperature (Heat/Cold) Presence of obscurants (Smoke, fog, etc.) Precipitation Sun effects (Sunlight readability)
Mechanical	Interface with NBC and oxygen mask Eyewear (Glasses/Contacts) Vibration and shock
Physiological	Fatigue Hypoxia Sleep deprivation G-loading Existing medical conditions Physiological state (Electrolyte balance, hydration level, etc.) Use of prescribed drugs and over-the-counter (OTC) medications
Sensory	Glare Luminance transients (Flashblindness) No/Low illumination Noise (Impulse/Steady-state)
Psychological	Mental/Emotional state (Stress) Fear/Anxiety Workload

Mechanical factors

In aviation and ground vehicular applications, the HMD must be able to operate satisfactorily in the presence of vibration and mechanical shock (Rash, 2000). Helicopters and ground vehicles produce high levels of vibration. This vibration affects both the vehicle and the operator. Human response to this vibration has been a more difficult problem to understand and solve than that with the aircraft (Hart, 1988). The effects of vibration manifest themselves in retinal blur, which degrades visual performance, and in physiological effects, the resulting degradation of which is not fully understood (Biberman and Tsou, 1991). The problem of the presence of vibration is exacerbated by the fact that all vehicle types differ in their vibration frequencies and amplitudes. Achieving full field-of-view of HMD imagery depends on maintaining proper alignment of the HMD optics, which is a difficult task in the presence of vibration.

Physiological factors

Fatigue, hypoxia, G-loading, sleep deprivation, and the use of drugs/medications are physiological factors that will degrade performance. Fatigue, sleep deprivation, and disruption of circadian rhythm are natural consequences of today's military operational planning where rapid force deployment across multiple time zones is expected, often followed immediately by a high operational tempo. Besides high operational tempos, uncomfortable working and sleeping environments, sustained operations, and insufficient staffing make fatigue a growing concern (Caldwell and Caldwell, 2005). Loss of sleep degrades attention, cognitive speed and accuracy, working memory, reaction time, and overall behavioral capability, often without the sleep-deprived person being aware of the deficits (van Dongen and Dinges, 2000).

Primarily a high altitude aviation problem, hypoxia, a decrease in ambient oxygen level, has significant effects on cognitive function. In mild cases, hypoxia causes only inattentiveness, poor judgment, and reduced motor coordination. Severe cases result in a state of complete loss of awareness and unresponsiveness where brain stem reflexes, including pupillary response to light and breathing reflex, cease.

Hypoxia also can be an issue for mountain operations (Cymerman and Rock, 1994). Warfighters deployed to high mountain terrain can experience a number of effects in vision, cognitive function, psychomotor function, mood, and personality. These effects are directly related to altitude and are much more common over 10,000 feet (3,048 meters). Both cognitive and psychomotor performance degradation occurs at altitudes greater than 10,000 feet (3,048 meters). The effects are most noticeable at extreme altitudes (>18,000 feet [>5,486 meters]) where degradation in perception, memory, judgment, attention, and other mental activity can occur (Cymerman and Rock, 1994).

Another physiological factor, generally confined to the high-performance aviation community, is G-loading. Under G-loading, a pilot's body is subjected to forces many times that of normal gravity (G). A pilot in an aircraft experiencing 4-Gs will be subjected to a force four times that of the force due to gravity. An F-16 fighter jet can pull in excess of 9 Gs during maneuvers.

Without appropriate countermeasures (e.g., wearing of a G-suit, a specialized garment worn by pilots subject to high levels of acceleration in order to prevent loss of consciousness), the effects of excessive G-loading can range from grayout to blackout to loss of consciousness (Harvey, 2006). Grayout is a reduction in visual capacity (often reported as a graying of vision) due to diminished blood flow to the eyes. This can result in a loss of peripheral vision (i.e., tunnel vision) and a loss of color perception and scene contrast but no loss of consciousness. The pilot still has auditory, tactile, and cognitive functions. Full vision can be recovered in two to three seconds after removal of the G-loading. In blackout, the oxygen supply to the eyes' retinas is severely reduced. A complete loss of vision occurs but still no loss of consciousness. Again, the pilot still can hear, feel, and think. Recovery time is a matter of two to three seconds after removal of G-loading. Most severe is loss of consciousness. The subject can no longer hear, feel, or think. Recovery does not occur for 15 to 20 seconds after the G-loading is removed. The time required to return to consciousness may vary from 9 to 20 seconds, and the pilot does not regain full, normal function for several minutes (Beaudette, 1984).

A final physiological factor to be mentioned here is the possible influence of prescribed drugs or over-the-counter (OTC) medications, used either as temporary or long-term medical condition management or as an operational necessity (e.g., countermeasures for fatigue during critical sustained operations). In the aviation community, pilots are routinely grounded if a medical condition warrants the use of prescription drugs. One major future exception to this fact is when approved drugs are administered for operational reasons in extreme situations, e.g., as fatigue countermeasures. The Air Force and the Army have been researching this possibility

(Caldwell et al., 2002a; Caldwell and Brown, 2003). Analogous research has investigated the use of short-acting hypnotics to improve daytime sleep and nighttime performance due to night shift work (Caldwell et al., 2002b).

Even OTC medications that are generally considered harmless can affect performance, both physical and cognitive. Users frequently ignore the ever-present warning against the operation of equipment and machinery during use. Even cold and allergy medications labeled as “non-drowsy” still list sleepiness as a possible side effect.

Perhaps the most used drug in the world is caffeine, present in coffee, tea, and many cola drinks. These high-level caffeine beverages are consumed innocently in large doses over long time periods. Military lore touts the advantages of coffee and tea for keeping Warfighters awake and alert, both on and off the battlefield. Caffeine is a drug that stimulates the central nervous system. Caffeine works on the body by increasing the heart rate, digestive secretions, respiratory rate, metabolic rate, and urine output. Low doses (~ 3 cups of coffee per day) increase alertness, while also increasing urination frequency and stomach acid levels. Higher doses can produce headache, irritability, insomnia, diarrhea, depression, and hyperactivity. Performance enhancement and side effects vary greatly among individuals. Sudden termination of caffeine consumption can result in withdrawal symptoms such as headache, lethargy, difficulty in concentration, and mild nausea.

A 1993 cross-sectional survey of over 9000 Britons investigated the relationship of habitual coffee and tea consumption to cognitive performance (Jarvis, 1993). Subjects completed tests of simple reaction time, choice reaction time, incidental verbal memory, and visuo-spatial reasoning, in addition to providing self-reports of usual coffee and tea intake. The study concluded that overall caffeine consumption showed a dose-response relationship to improved cognitive performance for each cognitive test. Older subjects appeared to be more susceptible to the performance-improving effects of caffeine than were younger subjects.

Sensory factors

Warfighter performance also can be impacted by sensory-related factors, such as the presence of loud and/or constant noise, sudden transients in luminance, glare sources, and operation in periods of no or reduced illumination (which can include operating at night, in foul weather, in caves, and in darkened ship interiors).

Sound provides important, useful information to the Warfighter. It can denote the presence of the enemy, contain strategic or tactical communication, provide information about the status of the local environment or vehicle being used, etc. Sound that is considered non-useful or distracting is identified as noise. A formal definition of acoustical noise is random occurrences of energy spikes varying in both amplitude and frequency (formally having a flat power spectrum across a significant portion of the human auditory response spectrum). Noise is generally characterized as either continuous (steady state) or impulse. As the noise level increases, it can progress from simply being annoying to being painful and damaging. At any level, noise can degrade communication, thereby increasing the potential for error.

Steady state noise technically is defined as lasting one second or longer but more commonly is continuous over the time period of concern. Common examples of steady state noise include road navigation noise, engine and generator noise, aerodynamic noise associated with wind or water rushing over vehicle exteriors, and electronic static. Steady state noise can mask important sounds that contain information. While low-level steady state noise exposure (less than 85 decibels) has not been thought to create adverse health effects, recent troop deployments to Bosnia and Kosovo have shown that low-level noise near military airports significantly impacted individual sleep habits and other noise-sensitive tasks (Luz et al., 2004).

Studies investigating the effects of steady-state noise on cognitive function have shown degradation in reading acquisition, time reaction to perceptual stimuli, attention, both intentional and incidental memory, and complex task performance (Dudek at al., 1991; Lercher, Evans, and Meis, 2003). Noise interferes in complex task performance, modifies social behavior, and causes annoyance. Noise exposure also has been shown to have adverse health effects. Studies of occupational and environmental noise exposure suggest an association with hypertension, whereas community studies show only weak relationships between noise and cardiovascular disease (Stansfeld and Matheson, 2003).

Impulse noise is defined as very intense sounds of short duration, abrupt onset and decay, and high intensity. Impulse noise describes the kinds of sound made by explosions, aircraft breaking the sound barrier, and the discharge of firearms. Exposure to impulse noise may result in temporary and permanent shifts in the threshold of hearing (Hodge and Price, 1978). Intermittent impulse noises will mask speech in varying degrees. Impulse noise in isolated one-second bursts is unlikely to disrupt much speech communication due to the redundancy of speech. However, as the frequency and duration of the noise bursts increase, so does the masking effect (U.S. Environmental Protection Agency, 1973).

Many sources of potentially distracting and damaging noise exist in the military environment, including weapons systems, wheeled and tracked vehicles, fixed- and rotary-wing aircraft, ships, and communications devices. Warfighters encounter noise through training, standard military operations, and combat. Warfighters also may be exposed to noise through activities that are present but not unique to military service, including engineering, industrial, construction, and maintenance tasks (Durch and Humes, 2006).

Studies have determined that individuals exposed to steady state sound levels of 85 decibels (A) (dBA) for an 8-hour period or longer are in danger of losing their hearing. Likewise, individuals exposed to impulse noise of 140 decibels (P) (dBP) or greater also are in danger of hearing loss (U.S. Army Center for Health Promotion and Preventive Medicine, 2006). Studies have shown that many Warfighters operate with hearing decrements (Humes et al., 2005; Shaw and Trost, 2005).

Military vehicles generally are not sound insulated, and weapons, by virtue of their operation, are sources of higher noise levels. Typical noise environments associated with the operation of military vehicles and weapons include the Army's M2 Bradley Fighting Vehicle (74-95 dBA at idle), the Army's UH-60A Black Hawk helicopter (106 dBA in cockpit), the Air Force's F-16 fighter (103 dBA in cockpit), and the Navy's coastal patrol craft (112 dBA in engine room).

A new source of impulse noise has arisen in the U.S. Army as the inadvertent result of an effort to introduce airbags into Army helicopters to reduce impact injuries during crashes (Ahroon et al., 2002). Deployment tests of airbag systems in the Army's UH-60 Black Hawk helicopter measured impulse noise levels from 144.8 to 162.4 dBP sound pressure level. Similarly, in Navy and Air Force aircraft, ejection seat operation can generate impulse noise levels in excess of 165 dBP (Naval Air Test Center, 1981). Of course, in both environments, pilots wear protective helmets with integrated noise attenuation, as well as supplemental noise protection in the form of earplugs.

No/Low illumination

Modern military operations are all-weather, day and night in nature. Combat operations no longer are confined to daytime or illuminated battlefields. Modern sensors expand the Warfighter's capability to fight in rain, fog, and even total darkness. Using microwave, radar, I^2, infrared (IR), and other technology-based imaging sensors, the "seeing" range of the human eye is extended into the darkest of nights and the gloomiest of weathers. However, this capability does not come without cost. Warfighters are expected to view, interpret, and make decisions on these "altered" representations of the outside world. Targets and backgrounds in these altered images are not presented to the eye and brain in the same mode (i.e., with the same spatial content) as when viewed by the natural unaided eye. Time-tested perceptions of objects are no longer fully usable when viewing images acquired from spectral ranges that extend beyond and may not include the normal visual range. As an illustration, Figure 1-9 depicts three presentations of the same scene, one as acquired by the unaided human eye, one as an IR sensor, and one as a radio frequency sensor (Wang, Wang and Peng, 2003).

Psychological Factors

The final category of adverse operational factors, one that is too frequently overlooked, is cognitive factors. Workload and the mental/emotional state of the user (defined by such conditions as stress level, presence of fear and anxiety, etc.) are factors that affect the user's level of attention to and retention of information presented via the HMD.

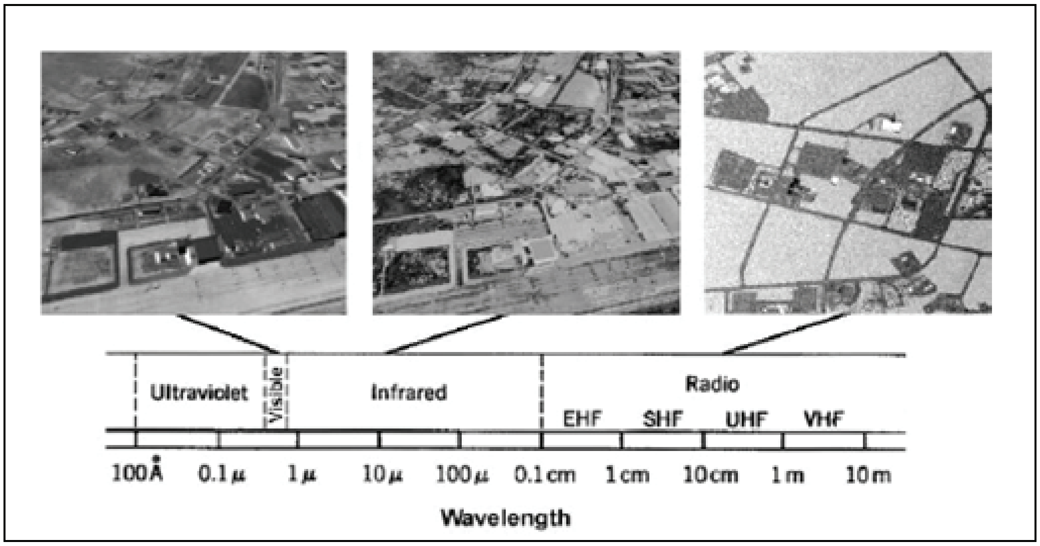


Figure 1-9. Three views of a scene as acquired by the unaided human eye (left), an IR sensor (center), and a radio frequency sensor (right) (Wang et al., 2003).

An undeniable consequence of the use of HMDs in NCW is increased workload. Workload can be defined as the combination of task demands and human response to these demands (Mouloua et al., 2001). In general, workload can be categorized as physical or cognitive. From the perspective of NCW and HMDs, the workload is cognitive in nature. Cognitive workload (or cognitive demand) is not well studied nor well understood, especially in scenarios where both physical and cognitive workload coexist or where multiple, simultaneous cognitive tasks are present (National Research Council, 1997). In addition, effects of and response to workload level differ for excessive and low workload scenarios.

In the benign environment of the development and testing laboratory, devices and systems based on advanced technologies may demonstrate superior performance; in a training environment, performance is often reduced. However, it is only when actual combat conditions and stressors are present that a true evaluation of system and user performance can be realized.

Combat stressors can be both physical and psychological. Physical stressors have a direct effect on the body. They may be both external and internal in origin. External physical stressors usually reflect the external environmental conditions, e.g. heat, cold, noise, and have been introduced previously in this chapter. Internal physical (or physiological) stressors, which include fatigue, hypoxia, sleep deprivation, G-loading, existing medical conditions, physiological state, and the use of prescribed drugs and OTC medications, also have been discussed in a cursory manner in this chapter and will be expanded upon in Chapter 16, *Performance Effects Due to Adverse Operational Factors*.

All of these physical (external and internal) stressors also place demand on the human cognitive and emotional systems, manifesting themselves as slow thought processing, memory lapses, anger, and/or fear (U.S. Army Center for Health Promotion and Preventive Medicine, 2005). Actual individual performance, with or without the use of advanced technology devices (e.g., imaging sensors and HMDs), is determined by the human response to these stressors. Humans respond with either physiological and mental reactions or reflexes designed to counteract these stressors. Responses may include decreased blood flow to the brain, muscles, and the heart; increased sweating; adrenaline release for energy and alertness; and muscle tension. These responses are intended to keep individuals within the range of physiological, emotional, and cognitive performance levels that optimize performance for survival.

The Warfighter's specific emotional and psychological reactions to combat have been referred to as "battle fatigue." Battle fatigue is described as a temporary response to the stress of combat capable of reducing combat performance by 10 to 50 percent. It is considered an inevitable consequence of military conflict (Hazen and Llewellyn, 1991). In modern times, this condition has been recognized as a distinct diagnostic phenomenon, referred to as Posttraumatic Stress Disorder (PTSD) (American Psychiatric Association, 1980). It was categorized as an anxiety disorder because of the presence of persistent anxiety, hypervigilance, exaggerated startle response, and phobic-like avoidance behaviors (Meichenbaum, 1994). Arguably, while often studied in war veteran populations, this disorder is not limited to veterans but can be *in situ* in the battlespace.

The Warfighter, the HMD and Cognition

An argument that the current trend in the military to use advanced technology to reduce manpower requirements and to overcome the vast physical demands of military training and combat has been presented. This argument has further stated that today's military environment is information intensive, and that this information is increasingly being presented in a head-up approach using head- and helmet-mounted displays. This deluge of information places a tremendous cognitive workload on the Warfighter. It is imperative, that, if HMDs are to indeed become a functional and useful technology, their design and execution be accomplished through a comprehensive understanding of their sensory, perceptual, and cognitive implications. Without a doubt, modern Warfighters, whether on land, under the oceans, in the air, or in space, are a special group who operate in environments unforgiving of human error, where cognitive degradation or failure can lead to, at best, an incomplete mission and, at worst, catastrophic consequences (Westerman et al., 2001).

References

Ahart, J., and Stiles, G. (1991). *The military's entry into air interdiction of drug trafficking from South America.* Santa Monica, CA: Rand Corporation.

Ahroon, W.A., Gordon, E., and Brozoski, F.T. (2002). Impulse noise levels of the UH-6OA/L Black Hawk helicopter cockpit air bag system - enhanced lateral air bags. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2002-16.

Alberts, D., Garstka, J., Hayes, R., and Signori, D. (2001). *Understanding Information Age Warfare*. Washington, DC: Command and Control Research Program (CCRP).

Alberts, D., Garstka, J., and Stein, F. (1999). Network centric warfare, developing and leveraging information superiority. Washington, DC: Command and Control Research Program (CCRP).

American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition*. Washington, DC: American Psychiatric Association.

Barnes, R.C., Jr. (1989). Civic action, humanitarian and civic assistance, and disaster relief: Military priorities in low-intensity conflict. *Special Warfare*, 34-41.

Beaudette, D.C. (1984). A hazard in aerobatics: Effects of G-forces on pilots. Federal Aviation Administration No. 91-61.

Becker, J.T., and Milke, R.M. (1998). Cognition and aging in a complex work environment: relationships with performance among air traffic control specialists. *Aviation and Space Environmental Medicine*, 69, 944-9 to 951-9.

Biberman, L.M., and Tsou, B. (1991). Image display technology and problems with emphasis on airborne systems. Alexandria, VA: Institute for Defense Analysis. IDA Paper P-2448.

Cahill, L. (2005). His brain, her brain. *Scientific American*, 292(5), 40-47.

Caldwell, J., and Brown, L. (2003). Running on empty?.Go pills, *Flying Safety Magazine*, March, 4-11.

Caldwell, J.A., and Caldwell, J.L. (2005). Fatigue in military aviation: an overview of U.S. military-approved pharmacological countermeasures. *Aviation Space and Environmental Medicine*, 76(7 Suppl), C39-C51.

Caldwell, J.A., Caldwell, J.L., and Roberts, K.A. (2002a). A comparison between the countermeasures Modafinil and napping for maintaining performance and alertness using a quasi-experimental analysis. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2002-14.

Caldwell, J.L., Hall. K.K., Prazinko, B.F., Norman, D.N., Rowe, T., Erickson, B.S., Estrada, A., and Caldwell, J.A. (2002b). The efficacy of Temazepam for improving daytime sleep and nighttime performance in Army aviators. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2002-05.

Cathcart, J. (1989). And that goes for domestic war to! *The Weinberger Doctrine and Domestic Use of the Military*. Carlisle Barracks, PA: U.S. Army War College.

Cheung, S.S**.**, Westwood, D.A., and Knox, M.K. (2007). Cold strain impairs attention via distraction from skin cooling. *Ergonomics*, 50, 275-288.

Cohen, W.S. (1999). Annual report to the president and the congress. Washington, DC: Department of Defense.

Covey, D.C. (1992). Offering a helping hand in Iraq. *U.S. Naval Institute Proceedings,* 109, 106-107.

Cymerman, A., and Rock, P.B. (1994). Medical problems in high mountain environments. Natick, MA: U.S. Army Research Institute of Environmental Medicine. Technical Note No. 94-2.

Dickens, H. (1989). The role of the reserve components in the war on drugs. Carlisle Barracks, PA: U.S. Army War College.

Dobbins, J., McGinn, J.G., Crane, K., Jones, S.G., Lal, R., Rathmell, A., Swanger, R., and Timilsina, A. (2003). *America's Role in Nation-Building from Germany to Iraq*. Alexandria, VA: RAND Corporation.

Dudek B., Marszal-Wisniewska, M., Merecz-Kot, D., Sulkowski, W., and Bortkiewicz, A. (1991). Effects of noise on cognitive processes of individuals in a laboratory experiment. *Polish Journal of Occupational Medicine and Environmental Health*, 4(3), 269-279.

Durch, J.S., and Humes, L.E. (2006). Too much noise. *Military Medical Technology*, 9, 8.

Foster, G.M. (1983). *The Demands of Humanity: Army Medical Disaster Relief.* Washington: U.S. Army. Center of Military History.

Garstka, J. (2000). Network Centric Warfare: An Overview of Emerging Theory, *PHALANX*.

Hancock, P.A. (1981). Heat stress impairment of mental performance: A revision of tolerance limits. *Aviation Space and Environmental Medicine*, 52, 177-180.

Harrison, T.G. (1992). Peacetime Employment of the Military: The Army's Role in Domestic Disaster *Relief*. Carlisle Barracks, PA: U.S. Army War College.

Harvey, E. (2006). Physiological effects of positive G forces. Retrieved January 13, 2007 from http://www.aerobatics.org.uk/repeats/ physiological_effects.htm.

Hart, S.G. (1988). Helicopter human factors. *Human Factors in Aviation*. San Diego, CA: Academic Press, 591-638.

Haskell, B. (2004). Guard continues to support counter drug operations, *The On Guard*, 33, 9.

Hazen, S., and Llewellyn, C. (1991). Battle fatigue identification and management for military medical students. *Military Medicine*, 156(6), 263-267.

Hiatt, K., and Horosko, S. (2006) The Department of Defense Response to Hurricane Katrina. *Proceedings of the NATO Conference: Strengthening National Public Health Preparedness and Response for Chemical, Biological, and Radiological Agents Threats, Skopje, Republic of Macedonia*.

Hodge, D.C., and Price, G.R. (1978). Hearing damage risk criteria. In Lipscomb, D.M. (Editor), *Noise and Audiology*. Baltimore: University Park Press, 167-191.

Hoffman, R.G. (2001). Psychological aspects of military performance in hot environments. In Pandolf, K., and Burr, R.E. (Editors), *Medical aspects of harsh environments.* Washington, DC: Borden Institute, 394-399.

Humes, L.E., Joellenbeck, L.M., and Durch, J.S. (Editors). (2005). *Noise and Military Service: Implications for Hearing Loss and Tinnitus*. Washington, DC: Institute of Medicine.

Jarvis, M.J. (1993). Does caffeine intake enhance absolute levels of cognitive performance? *Psychopharmacology (Berl*), 110(1-2), 45-52.

Jones, J.L. (1991). Operation Provide Comfort: Humanitarian and Security Assistance in Northern Iraq. *Marine Corps Gazette*, 98-107.

Johnson, R.J., and Kobrick, J.L. (2001). Human psychological performance in cold environments. In Pandolf, K., and Burr, R.E. (Editors), *Medical Aspects of Harsh Environments.* Washington, DC: Borden Institute,149.

Juda, A. (1993). The Military and Humanitarian Operations. *Library Notes*, Navy War College, 21(6), Retrieved from http://www.nwc.navy.mil/library/3Publications/NWCLibraryPublications/LibNotes/ libhuman.htm

Kelly, E.M., Jr. (1992). *The U.S. Military and Disaster Response*. Carlisle Barracks, PA: U.S. Army War College.

Laurin Publishing. (2005). *The Photonics Dictionary*. Retrieved March 16, 2006 from http://www.photonics.com/dictionary.

Lercher, P., Evans, G.W., and Meis, M. (2003). Ambient noise and cognitive processes among primary schoolchildren. *Environment and Behavior*, 35(6), 725-735.

Lieberman, H., Bathalon, G., Falco, C., Georgelis, J., Morgan, C., Niro, P., and Tharion, W. (2002). The “fog of war”: Documenting cognitive decrements associated with the stress of combat. *Proceedings of the 23rd Army Science Conference*.

Luz, G.A., Whiteford, W.D., and Broska, K.A. (2004). Protecting military forces from unhealthy levels of noise during deployment. *U.S. Army Medical Department Journal*. Fort Sam Houston: U.S. Army Medical Department and School. January-March, PB-8-04-1/2/3, 44-48.

Marshall, E. (1988). A war on drugs with real troops? *Science*, 241, 13-15.

Matson, E., and DeLoach, S. (2003). An organization-based adaptive information system for battlefield situational analysis. *Proceedings of the International Conference on Integration of Knowledge Intensive Multi-Agent Systems.* Waltham, MA.

Meichenbaum, D. (1994). *A Clinical Handbook/Practical Therapist Manual for Assessing and Treating Adults with Post-Traumatic stress Disorder.* Ontario, Canada: Institute Press.

Miles, D. (1991). Helping the Kurds. *Soldiers*, July, 13-20.

Mouloua, M., Gilson, R., Kring, J., and Hancock, P. (2001). Workload, situation awareness and teaming issues for UAV/UCAV operations. *Proceedings of the Human Factors and Ergonomics Society, 45th Annual Meeting*, 1, 162-165.

Murray, W. (Editor). (2001). Army transformation: A view from the U.S. Army War College. Carlisle Barracks, PA: U.S. Army War College.

Nalepa, G.F. (1993). Marine Corps humanitarian and disaster relief operations. *Marine Corps Gazette*, 22-24.

National Research Council. (1997). *Tactical Displays for Soldiers: Human Factors Considerations.* Washington, DC: National Academy Press.

Naval Air Test Center. (1981). Impulse noise measurement of the SJU-4/A ejection seat operation. Patuxent River, MD: Naval Air Test Center. Report No. SY-98R-81.

Nindl, B., Leone, C., Tharion, W., Johnson, R., Castellani, J., Patton, J., and Montain, S. (2002). Physical performance responses during 72-h of military operational stress**.** *Medicine and Science in Sports and Exercise*, 34(11), 1814-1822.

North Atlantic Treaty Organization. (2001). Can soldiers be peacekeepers and warriors? *NATO Review*, Web edition, 49(2), 16-20, Retrieved from http://www.nato.int/docu/review/2001/0102-04.htm

Pozos, R.S., and Danzl, D.F. (2001)**.** Human physiological responses to cold stress and hypothermia. In Pandolf, K., and Burr, R.E. (Editors), *Medical aspects of harsh environments*. Washington, DC: Borden Institute. 375.

Rash, C.E. (Editor). (2000). *Helmet-Mounted Displays: Design Issues for Rotary Wing Aircraft.* Bellingham, WA: SPIE Press.

Rash, C.E., Kalich, M.E., van de Pol, C., and Reynolds, B.S. (2002). The issue of visual correction compatibility with helmet-mounted displays. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL report No. 2003-04.

Rash, C.E. (2004). Unpublished data developed in support of Army Technical Objective, Warfighter Visual Performance Prediction Model for Advanced Display Concepts.

Rushton, J.P., and Jensen, A.R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, 11(2), 235–294.

Shaw, G.B., and Trost, P.R., (2005). Statistical analysis of hearing loss among Navy personnel. Alexandria, VA: CAN. CRM D0011228.A2.

Shotwell, C.B. (1992). Food and the use of force: The role of humanitarian principles in the Gulf Crisis and beyond. *The Fletcher Forum of World Affairs*, 131-152.

Simpson, D. (1992). Military rule in drug interdiction. Air University Library, Maxwell Air Force Base, Montgomery, AL.

Stackpole, Henry, C., III, and Chase, E.L. (1993). Humanitarian Intervention and disaster relief: Projecting military strength abroad to save lives. *Marine Corps Gazette*, 16-20.

Stansfeld, S.A., and Matheson, M.P. (2003). Noise pollution: non-auditory effects on health. *British Medical Bulletin*, 68, 243-257.

Steele, D. (2001). The Army Magazine's "Hooah" Guide to Army Transformation. Arlington, VA: Association of the U.S. Army.

Sutton, E.L. (1992). The New Role of Humanitarian Assistance in National Military Strategy: How to Make It Work. Carlisle Barracks, PA: U.S. Army War College.

Tyson, A. (2005). U.S. Pushes Anti-Terrorism in Africa. Retrieved July 26, 2005 from Washington Post.com, A01.

U.S. Army Center for Health Promotion and Preventive Medicine. (2005). Stressing combat performance. Aberdeen Proving Ground, MD: U.S. Army Center for Health Promotion and Preventive Medicine. 22-002-0499.

U.S. Army Center for Health Promotion and Preventive Medicine. (2006). Hearing conservation program. Aberdeen Proving Ground, MD: U.S. Army Center for Health Promotion and Preventive Medicine.

U.S. Army Natick Soldier Center. (2004). Future force warrior. Retrieved from http://www.natick.army.mil/soldier/WSIT/.

U.S. Army War College. (1988). The military in the war on drugs: A selected bibliography. Carlisle Barracks, PA: U.S. Army War College.

U.S. Congress House. (1989). Committee on armed services: Military role in drug addiction. Washington, DC: Government Printing Office.

U.S. Department of the Army. (1999). Army Pamphlet 611-21 - Military Occupational Classification and Structure. Washington, DC. (Under revision).

U.S. Department of the Army. (2005). Office of Army Demographics. Retrieved from http://www.armyg1.army.mil/hr/Demographics.asp.

U.S. Department of the Army and Air Force. (2003). Technical Bulletin, Heat stress control and heat casualty management. Headquarters Department of the Army and Air Force. TB MED 507//AFPAM 48-152 (I).

U.S. Department of Defense. (1981). *Military Standard: Human Engineer*ing: *Design Criteria for Military Systems, Equipment, and Facilities.* Washington, DC: Department of Defense. MIL-STD-1472C.

U.S. Department of Defense. (1984). *Military Standard: MIL-STD-1295 Human Factors Engineering Criteria for Design Criteria for Helicopter Cockpit Electro-Optical Displays Symbology*. Washington, DC: Department of Defense. Cancelled October 1996.

U.S. Department of Defense. (1989). *Military Specification: Aviator's Night Vision Imaging System AN/AVS-6(V)l, AN/AVS-6(v)2*. Washington, DC: Department of Defense. MIL-A-49425(CR).

U.S. Department of Defense. (1999). *Military standard: MIL-STD-461E Electromagnetic Emission and Susceptibility Requirements for the Control of Electromagnetic Interference*. Washington, DC: Department of Defense. MIL-STD-461E.

U.S. Department of Defense. (2004). Selected manpower statistics FY 2004. Washington Headquarters Services. Retrieved March 14, 2006 from http://www.dior.whs.mil/ mmid/ M01/fy04/m01fy04.pdf.

U.S. Department of Defense. (2005). Military Personnel Statistics. Retrieved March 14, 2006 from http://web1.whs.osd.mil/mmid/ military/ms8.pdf.

U.S. Department of State. (2005). Defense Department Report, January 4: Humanitarian Relief Efforts. Retrieved from http://usinfo.state.gov/gi/Archive/ 2005/Jan/04-780600.html.

U.S. Environmental Protection Agency. (1973). *Public Health and Welfare Criteria for Noise, 550/4-73-002*. Washington, DC: U.S. Environmental Protection Agency.

U.S. Marine Corps. (1998). Marines Awaiting Training (MAT) Program, Physical Training (PT02).

van Dongen, H.P.A., and Dinges, D.F. (2000). Circadian rhythms in fatigue, alertness and performance. In Kryger, M.H., Roth, T., and Dement, W.C., (Editors), *Principles and Practice of Sleep Medicine*. Philadelphia, PA: WB Saunders, 391-399.

Wang, C., Wang, Z., and Peng, Q. (2003). Wavelength independent texture for multi-spectral scene simulation. *Third International Symposium on Multispectral Image Processing and Pattern Recognition*. In Lu, H., and Zhang, T. (Editors), *Proceedings of the International Society for Optical Engineering (SPIE)*, 5286, 717-722.

Westerman, R., Darby, D.G., Maruff, P., and Collie, A. (2001). ADF health. *Journal of the Australian Defense Health Services*, 2, April.

2 THE HUMAN-MACHINE INTERFACE CHALLENGE

Gregory Francis
Clarence E. Rash
Michael B. Russo

In Chapter 1, *The Military Operational Environment*, the military user of helmet-mounted displays (HMDs) and operational environment were described in a general fashion. In this chapter, we try to identify the essential issues that must be considered when attempting to use an HMD or head-up display (HUD). We begin by suggesting that the human-machine interface (HMI) challenge of HMD design is to *use robust technology to organize and present information in a way that meets the expectations and abilities of the user*. This chapter outlines some of the main concepts that are relevant to this challenge. Subsequent chapters describe important details about human sensory, perceptual, and cognitive systems and describe the characteristics, abilities, and limitations of HMD systems. Additional engineering-related information about HMDs can be found in this book's predecessor, *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft* (Rash, 2001). The following discussion steps back from some of the details of these systems and looks at a bigger picture to identify themes that will apply to many different situations.

Although engineering teams do not make an HMD awkward to use, HMDs often fail to live up to their promised performance (Keller and Colucci, 1998). Many of the issues with HMDs are those common to all information displays, which can either make information more useable or can increase workload and stress (Gilger, 2006). In spite of conscientious efforts by the human factors engineering (HFE) community, HMD designs have not been optimized for the capabilities and limitations of the human user (National Research Council, 1997). The progress that has been made in addressing HFE issues has been modest and largely limited to either anthropometry or to the physiological characteristics of the human senses, i.e., vision and audition. Perceptual and cognitive factors associated with HMD user performance have been almost totally overlooked. With the information-intensive modern battlespace, these factors are taking on an even greater importance.

While HMDs can be used for a wide variety of purposes and display many different types of information, fundamentally there is always a "region" where a human user interacts with the HMD. This is the *human-machine interface* (HMI). This interface serves as a *bridge* that connects the user and the machine (Hackos and Redish, 1998). The design of this interface is critically important because the information from quality sensors and computer analysis will not be beneficial unless the human user understands the information. It is important to note that the HMI is not a device; instead, it is a virtual concept, represented by the interaction of the human sensory, perceptual and cognitive functions with the HMD's information output(s).

This chapter is organized to examine the different aspects of the human-machine interface. We start with a basic description of human perceptual and cognitive systems, and consider their biases, abilities, and limitations. We then turn to a description of HMDs and consider their abilities and constraints. Finally, we discuss the interface between these two systems and consider general aspects of how they can be brought together. This discussion is kept at a relatively high-level abstraction of ideas and leaves the details for other chapters.

Human Sensation, Perception and Cognition

Sensation, perception, and cognition all refer to the acquisition, representation, and utilization of information in the world. These processes appear easy, automatic, and complete, but in reality, they are extremely complex, take substantial processing, and are surprisingly limited in terms of their relation to the veridical world.

Sensation

Sensation is one of the first steps in acquiring information about an environment. It refers to the detection of a property (or characteristic) of an object in the world. Typically this process involves responses from biological receptors that are sensitive to a particular form of energy. These receptors can be very complex and can respond to a wide range of energy forms. For vision, the receptors are cells in the back of the eye called rods and cones that respond to light energy of different wavelengths. For audition, the receptors are the cilia of the organ of Corti that sit on the basilar membrane in the cochlea of the ear. For cutaneous sensation (touch), there are several types of receptors that are embedded in the skin and respond to flutter, vibration, pressure, and stretching. Figure 2-1 shows schematic views of the receptors for vision, audition, and touch.

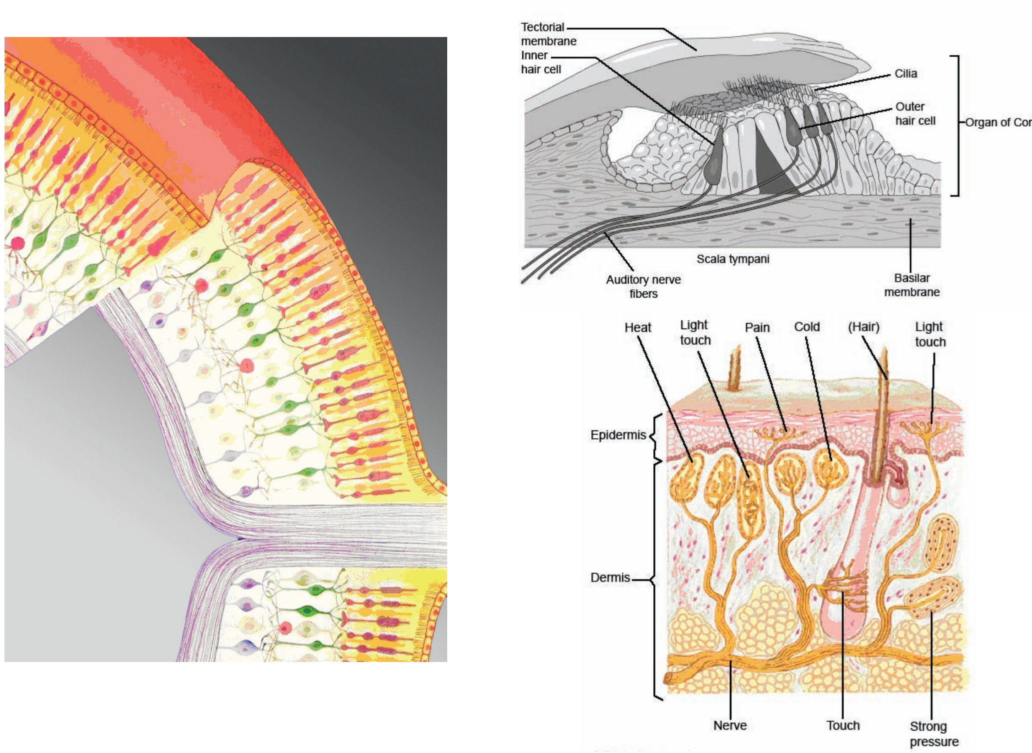


Figure 2-1**.** Different receptors are responsible for the sensation of dissimilar types of stimulus energy. Left: Cross section of the back of the eye shows photoreceptors that are sensitive to light energy. Top right: Cilia on the organ of Corti are sensitive to sound energy. Bottom right: Receptors in the skin are sensitive to forces on the skin.

Perception

Humans are not aware of sensory processes, except as they influence our perception of the world. Perception refers to the awareness of objects and their qualities. The process of perception is so accurate and convincing that the detailed mechanisms of how perception happens are mostly hidden from general awareness. We have the impression that as soon as we open our eyes, we see the world, with all of its objects, colors, patterns, and possibilities. In reality, the events that occur when the eyes open are astonishingly complex processes that depend on precise chemical changes in the eye, transmission of electrical and chemical signals through dense neural circuits, and rich interactions with both memories of previous events and planned interactions with the world.

Figure 2-2 characterizes the *perceptual loop* that is mostly hidden from awareness when looking at the world. This figure and the following discussion are adapted from Goldstein (2007). (See Chapter 15, *Cognitive Factors*, for a similar loop that describes some cognitive processes.) One could start a description of the loop at any place and could talk about any of the perceptual senses. We will focus on visual perception because it is easy to refer to the stimuli, and we will start with the *Action* node on the far right. Here, a human interacts with the environment in some way that changes the visual array. This could be as simple as opening the eyes, turning the head, or taking a step forward. It could also be a quite complex event such as jumping toward a ball, splashing paint on a surface, or changing clothes. The action itself changes the environment. Thus, the next step in the loop is the *Environmental stimulus*.

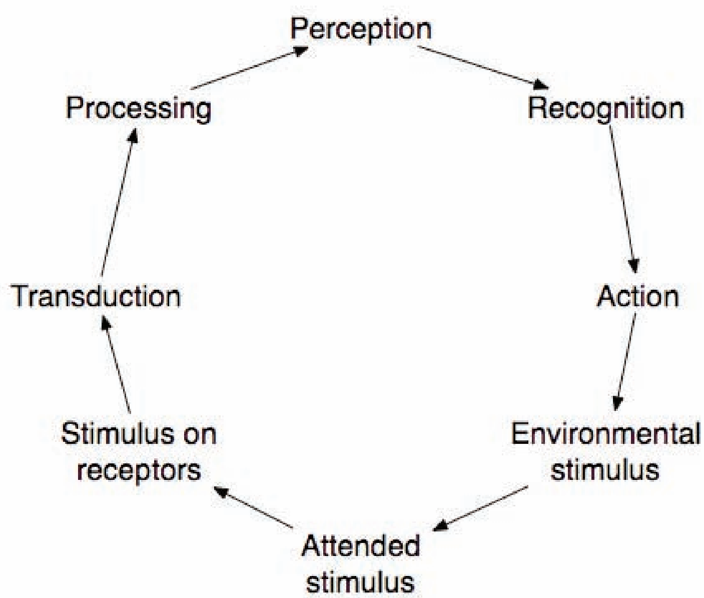


Figure 2-2. The perceptual loop demonstrates that perceptual processing involves many different complex interactions between the observer and the environment (adapted from Goldstein, 2007).

The environmental stimulus refers to properties of things in the world. This is the information that is, in principle, available to be acquired. For visual perception, this refers to the currently visible world. In practice, a person cannot acquire information about the entire environmental stimulus. Instead, perceptual processes usually focus on only a relatively small subset of the environmental stimulus, the *attended stimulus*.

The attended stimulus is the part of the environmental stimulus that is positioned in such a way that sensory systems can acquire information. The term *stimulus* may need some explanation, as it is very context specific. In some situations, the stimulus may be a particular object in the world, such as a building. In other situations the stimulus may refer to a particular feature of an object in the world, such as the color of a building's wall. In still other situations, the stimulus may be a pattern of elements in the world, such as the general velocity and direction of a group of aircraft. A person can attend a stimulus by moving the body, head, and eyes so that the relevant parts of the environmental stimulus fall on to the appropriate perceptual sensors.

The *stimulus on receptors* is the next step in the perceptual loop. Here, energy that corresponds to the attended stimulus (and some energy from parts of the environmental stimulus that are not the attended stimulus) reaches specialized cells that are sensitive to this energy. For visual perception, the specialized cells are photoreceptors in the back of the eye (Figure 2-1). These photoreceptors are sensitive to light energy (photons).

Transduction refers to the conversion of stimulus energy into a form of energy that can be used by the nervous system of the observer. The stimulus energy must be converted into an internal signal that encodes information about the properties of the attended stimulus. For visual perception, the photoreceptors of the eye initially undergo chemical transformations when they absorb photons. These chemical transformations induce an electrical change in the photoreceptor. These electrical changes are then detected and converted into a common format that the rest of the brain can use.

After transduction, the stimulus undergoes *processing* by the neural circuits in the brain. This processing includes identification of stimulus features (e.g., patterns of bright and dark light; and oriented edges of stimuli). As this information is processed, the observer experiences the phenomenological experience of *perception*. At this stage, the observer gains an awareness of properties of the attended stimulus in the world. The perceptual experience is not simply a function of the attended stimulus energy, because the experience may also depend on a memory of the world from a few previous moments. It may also depend on the action generated by the observer.

Recognition refers to additional interactions with memory systems to identify the attended stimulus, relative to the observer's experience and current needs. Here, the observer interprets the properties of the attended stimulus, perhaps to identify friend or foe and opportunity or threat. As a result of this interpretation, the observer generates some kind of action, which restarts the perceptual loop.

While we have stepped through the stages of the perceptual loop one after another, in reality all the stages are operating simultaneously and continuously. Thus, actions based on one moment of recognition may occur at the same time as transduction from a previous stimulus on receptors. Moreover, some information about the environment can only be detected after multiple passes through the perceptual loop, where the observer plans a specific sequence of actions so that they change the environmental stimulus in a way that allows them to gain particular desired information (e.g., moving the head back and forth to induce a motion parallax, which allows for discrimination of an object in depth).

One of the main messages from the description of the perceptual loop is that perception is an extremely complex experience. Each stage of the perceptual loop plays an integral role in perceptual experience and contributes to how we interpret and interact with the world. What is known about the details of each stage in the perceptual loop is far too complicated to describe in this book. Some of the other chapters in this book do discuss some of the details that are especially important for HMDs. Here, we try to take a more global view of the issues.

The human perceptual systems have evolved to process only certain types of stimulus inputs. For example, the human visual system covers only a small subset of the electromagnetic spectrum (i.e., 380 to 730 nanometers). We interpret different wavelengths of light as perceptually different colors, but the visual system is unaware of electromagnetic energy at longer wavelengths (heat) or very short wavelengths (ultraviolet and beyond).

Similarly, the human visual system has evolved to detect subtle properties of the visual world by interpreting global flows of streaming motion (Gibson, 1950). As we move through an environment, individual objects in the world produce moving patterns of light across our retinas. The patterns of movement contain significant information about the world and the properties of the observer. Figure 2-3 schematizes two flow fields generated by different movement of the observer. The line projecting out from each dot indicates the direction and velocity (length of the line) of a dot at that position in the field-of-view (FOV). Figure 2-3A shows the flow field generated when the observer moves in a straight line toward a fixed point in the middle of the field. All of the motion patterns expand from the fixed point. Sensitivity to the properties of the flow field can allow a moving observer to be sure that he or she is moving directly toward a target.

Flow fields can be much more complicated. Figure 2-3B shows the flow field generated by an observer traversing on a curved path while fixating on the same spot as in Figure 2-3A. To maintain fixation on a point, the observer must change his or her head or eyes, and these movements change the properties of the flow field.

Humans can use these kinds of flow fields to estimate heading direction to an accuracy within one visual degree (Warren, 1998), and many areas of the brain are known to be involved in detecting motion and flow fields (Britten and van Wezel, 1998). Flow fields of this type exist for many different situations, and they are especially important for detecting heading and direction of motion in aircraft (Gibson, Olum and Rosbenblatt, 1955). However, there are some kinds of flow fields that humans interpret incorrectly and so produce perceptual illusions (e.g., Fermuller, Pless and Aloimonos, 1997). Thus, the perceptual systems limit the kinds of information that people can extract from flow fields.

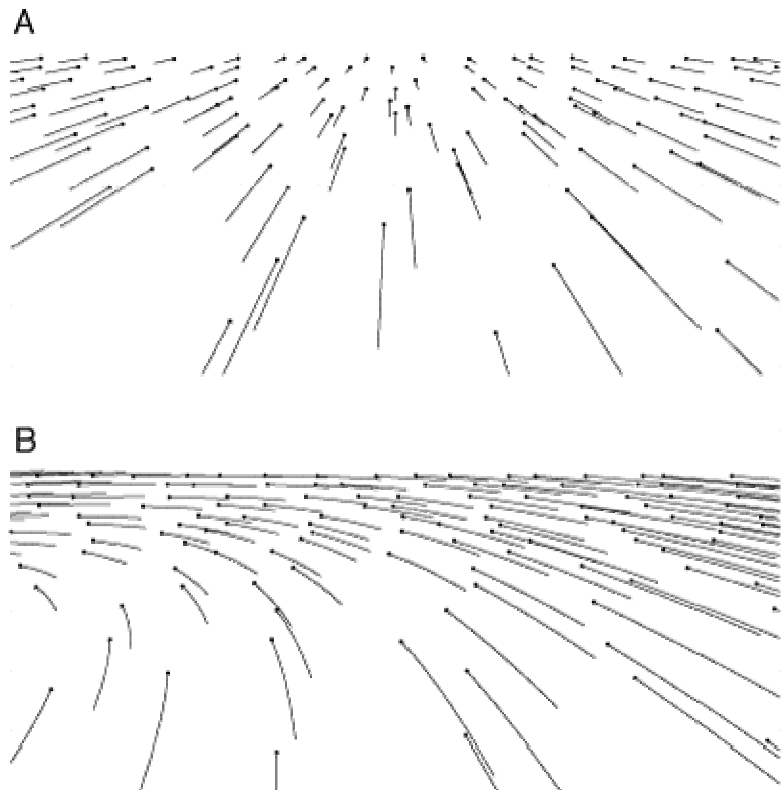


Figure 2-3**.** A) Radial dot flow generated from a straight-line path across a ground plane. The direction of motion can be determined by finding the focus of expansion, the point in the flow field where there is no horizontal or vertical motion. This may not be explicitly present, but can be extrapolated from the motion of other points in the image. B) Curvilinear dot flow generated from a curved path across a ground plane, also with a fixed gaze. From Wilkie and Wann (2003).

There are similar issues for depth perception. Objects in the world occupy three spatial dimensions, but the pattern of light on the retina of an eye is a 2-dimensional projection of light from the world. The third dimension must be computed from the differences in the projection to the two eyes, by changes in the projection over time (motion parallax), or by pictorial cues that generally correlate with differences in depth. As part of this process, the human visual system has evolved to make certain assumptions about the world. These assumptions bias the visual system to interpret properties of a scene as cues to depth. For example, the objects in the top row of Figure 2-4 generally look like shallow holes, while the objects in the bottom row look like small hills. There is a bias for the visual system to assume that light sources come from above objects. The interpretation of the objects as holes is consistent with this idea. Now rotate the page so that the figure is upside down. The same bias for light to come from above now switches the percept of the items.

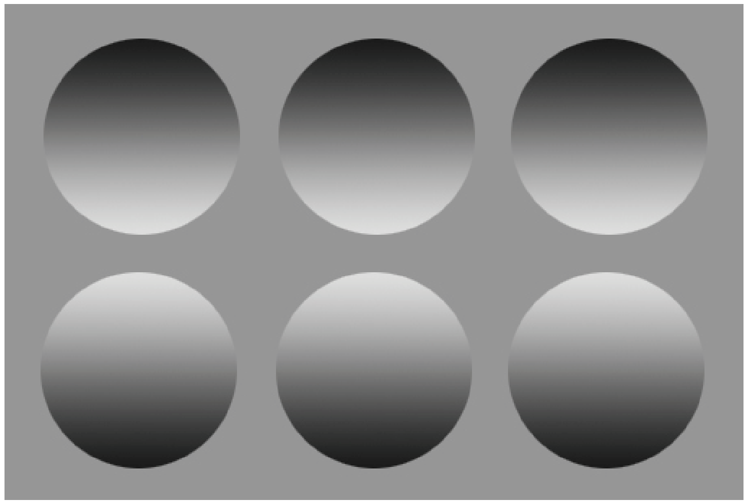


Figure 2-4**.** The visual system is biased to assume an illuminant comes from above. The perceived depths associated with the dots reveal this bias. The top row appears to be made of shallow holes because the brighter part is at the bottom and the top is darker (in shadow). The bottom row appears to be made of small hills because the top is brighter (light hitting it) while the bottom is darker (in shadow).

There are similar biases for interpreting patterns of reflected light. Figure 2-5 shows what is called the Snake Illusion (Adelson, 2000; Logvinenko et al., 2005). The diamonds are all the same shade of gray, but are set against light or dark backgrounds. They look different because the visual system interprets the dark bar on top as a transparent film in front of the gray diamonds and the white background. As seen through such a film, the gray diamonds appear brighter than the (physically identical) gray diamonds below that are not seen through a film. Here, a bias in the visual system to interpret patterns of light as indicative of transparent surfaces changes the apparent brightness of objects. Such complex interpretations of scenes are quite common (Gilchrist et al., 1999).

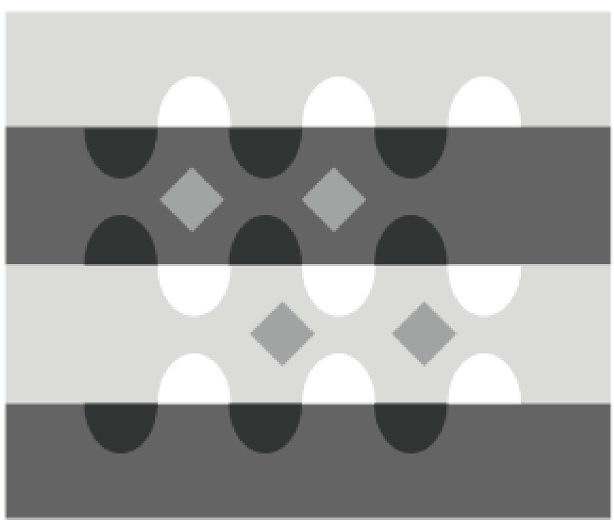


Figure 2-5**.** The Snake Illusion: The gray diamonds are physically the same shade of gray, but the diamonds in the top row appear lighter than the diamonds in the bottom row. Adapted from Adelson (2000).

Although more difficult to demonstrate in a printed format, there are similar biases and influences for other perceptual systems. Humans detect sounds only within a certain band of frequencies, and have varying sensitivities across these frequencies that are biased toward the range of frequencies that correspond to human speech (Fletcher and Munson, 1933). Likewise, segmentation of the auditory stream follows certain rules that can cause a listener to perceive multiple sound sources when only one sound source is actually present. See Chapters 11, *Auditory Perception and Cognitive Performance*, and 13, *Auditory Conflicts and Illusions,* for further discussions of human auditory perception, and Chapter 18, *Exploring the Tactile Modality for HMDs*, for a description of haptic perception.

The main lesson from these observations about human perception is that the perceptual systems have evolved to identify and extract some types of information in the environment but have not evolved to process other types of information. Evolutionary pressures have lead to perceptual systems that operate well within some environments, but these same systems will behave poorly when placed in entirely new environments.

Cognition

Similar observations can be made about human cognition (see Chapter 15, *Cognitive Factors*, for a fuller discussion of cognitive systems). Humans are very good at tasks involving face recognition (e.g., Walton, Bower, and Power, 1992) because evolutionary pressures give an advantage to being able to recognize, interpret, and remember faces. Humans also are quite good at many pattern recognition tasks that are difficult for computers, such as reading handwriting, interpreting scenes, or understanding speech in a noisy environment (Cherry, 1953). However, there are many recognition tasks where humans perform quite poorly, especially tasks that involve judgments of probability or the use of logic (Khaneman and Tversky, 1984). Moreover, biases and limitations of perceptual, attentional, memory, decision-making, and problem solving systems severely restrict the ability of individuals to perform well in many complex situations.

We complete this description of human behavior by pointing out a few common misconceptions about perception and cognition. First, evolutionary pressures rarely lead to optimal behaviors, and humans rarely act in an optimal way. Instead, evolution tends to select solutions that satisfy many different constraints well enough. Humans are good pattern recognizers, but outside of a few special situations it would be false to characterize them as optimal. Second, perception does not involve direct awareness of the world. Some researchers go so far as to claim that all of perception is an illusion, but this presupposes that one has a good definition of reality. Such philosophical discussions (Sibley, 1964) are beyond the scope of this book, so we simply note that perception actually requires significant resources and processing to acquire information about an environment. Third, contrary to centuries of philosophizing, humans are not generally rational. Studies of human cognition show that when humans appear to be rational it is not because they think logically, but because they learn the specific rules of a specific situation and act accordingly (Wason and Shapiro, 1971). Thinking rationally requires substantial training in the rules of logic, and this often does not come naturally. Finally, it is a mistake to believe that an individual can use all available information. The presence of information on a display or “known” to a person does not mean that such information or knowledge will guide human behavior in any particular situation.

Machine: Helmet-Mounted Displays

HMDs can be constructed in many different ways. Variations in sensors can make an image on a display sensitive to different aspects of the environment. Variations in the display change how information is presented to the human observer. Whatever the application, HMDs are not stand-alone devices. As integrated components of combat systems (as well as in other applications), they are used to present information that originates from optical and acoustic sensors, satellites, data feeds, and other communication sources, Even in simulation or virtual immersion applications, external signals (consisting of visual and audio data or information) must be provided. In the following discussion, we briefly place the HMD in perspective, by considering its role as just one component of the night imaging system. The function of the HMD does not come to bear until energy that is created by or reflected from objects and their environment (referred to as *stimuli*) is captured (detected) by *sensor(s)*, and then manipulated, transmitted, and presented on the HMDs *displays*. While not an exhaustive examination of the important properties of HMDs, this distinction helps to highlight some of the key features that relate to the HMI.

Stimuli

There are several ways to define a stimulus, but usually the term is used to refer to the properties of objects of interest in the world that generate sensations. This definition is important because an HMD filters and modifies the detected properties of the object. Thus, for example, a faint visual stimulus that normally would be undetected by the unaided eye can become detected with the aid of a night vision sensor; similarly, a faint sound stimulus that normally would be undetected by the unaided ear can become detected with the aid of an amplifier. A different way to describe the situation is to note that the night vision sensor converts one stimulus (the original faint stimulus) into another stimulus (visual energy in the HMD's display component). These are largely philosophical distinctions, although it is sometimes useful to switch between descriptions to explain different aspects of perception.

For human vision, input sources can be any object that emits or reflects light energy anywhere in the electromagnetic spectrum. For nighttime operations, examples include obvious naked-eye sources such as weapon flashes, explosions, fires, etc., and thermal sources such as human bodies, tanks, aircraft, and other vehicles that would serve as emissive sources during and after operation.

For human hearing, input sources are both outside and inside the personal space (e.g., cockpits for aviators and vehicle interiors for mounted Warfighters). Outside audio input sources include explosions, weapon fire, and environment surround sounds (especially for dismounted Warfighters). Inside sources include engine sounds, warning tones, and communications.

With an HMD application, properties of the external world are detected by sensors and are then converted into electronic signals. These signals are relayed to the visual or audio display component of the HMD, where an image of the external "scene" (visually or acoustically) is reproduced, sensed, and then acted upon by the user. A simplified block diagram for this visual/acoustical stimulus-sensor-display-user construct is presented in Figure 2-6. In this simplistic representation, the HMD acts as a platform for mounting the display (or, in some designs, a platform for mounting an integrated sensor/display combination).

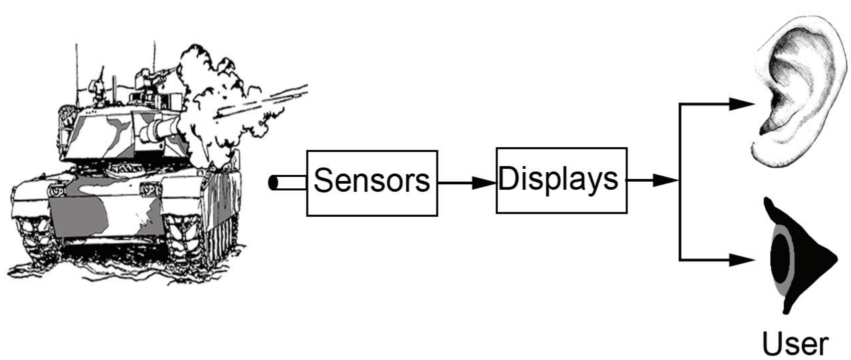


Figure 2-6**.** Simplified block diagram of the visual/acoustical stimulus-sensors-displays-user construct used in HMDs.

Sensors

Sensors are devices that acquire information about stimuli in the outside world. A sensor is defined as a device that responds to a stimulus, such as heat, light, or sound, and generates a signal that can be measured or interpreted. HMD visual imagery is based on optical sensors. Historically, the use of acoustic sensors in the battlespace has been limited, with underwater applications being the most prevalent. Generally, HMD audio information presentation has been limited to the reproduction of communications via speakers. However, acoustic sensors are rising in importance as their utility is explored. Different acoustic sensors operating in the ultrasonic and audible frequency ranges have a wide range of applications and impressive operating ranges. Optical forward-looking infrared (FLIR) imaging sensors can have an effective detection operating range as great as 20 kilometers (km) (12.4 miles) under optimal environmental conditions; acoustic sensors theoretically can operate out to approximately 17 km (10.6 miles) under ideal conditions. For both sensor technologies, identification ranges are more limited.

Many HMDs are based on optical imaging systems and are used to augment normal human vision. These systems include sensors that are sensitive to energy that is not detected by the normal human eye. The HMD displays this energy in a way that helps an observer identify objects in the environment. Optical imaging sensors can be categorized by the type of energy (i.e., range of wavelengths) they are designed to detect. Each specific category defines the imaging technology type (and therefore the physics) used to convert the scene of the external world into an image to be presented on the HMD's display. Theoretically, such sensors may operate within any region of the electromagnetic spectrum, e.g., ultraviolet, visible, IR, microwave, and radar. Currently, the two dominant imaging technologies are image intensification (I^2) and FLIR.

Image intensification (I^2) sensors

The sensor used in an I^2 system (as applied in early generation I^2 devices) uses a photosensitive material, known as a photocathode, which emits electrons proportional to the amount of light striking it from each point in the scene. The emitted electrons are accelerated from the photocathode toward a phosphor screen by an electric field. The light emerging from the phosphor screen is proportional to the number and velocity of the electrons striking it at each point. The user views the intensified image formed on the phosphor screen through an eyepiece (Figure 2-7).

I^2 sensors generally detect energy in both the visible range and the near-IR range; the actual wavelength range is dependent on the technology generation of the I^2 sensor (and sometimes the presence of optical filters).

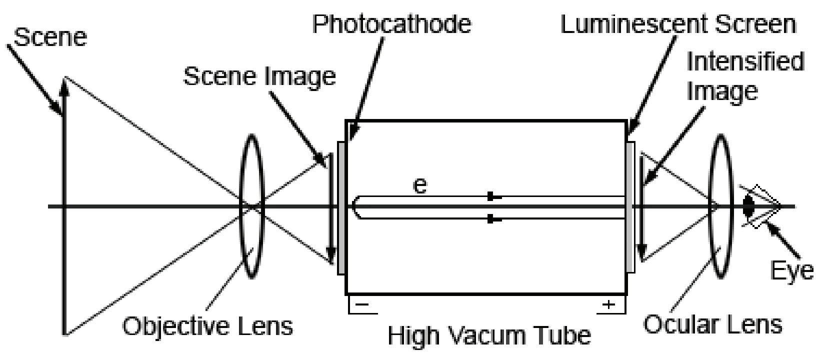


Figure 2-7**.** The basic parts of an I^2 device. The photocathode effectively amplifies the light intensity of the visual scene and projects the amplified scene on a screen that is perceived by the observer.

This process is analogous to using a microphone, amplifier and speaker to allow the user to more easily hear a faint sound. In both cases, some of the “natural fidelity” may be lost in the application process. The intensified image resembles a black-and-white television image, only usually in shades of green (based on the selected display phosphor) instead of shades of gray. However, recent advances offer the promise of pseudo-color I^2 devices based on dual- or multi-spectrum I^2 technology (Bai et al., 2001; Toet, 2003; Walkenstein, 1999).

Forward-looking infrared (FLIR) sensors

FLIR-based imaging systems operate on the principle that every object emits energy (according to the Stefan-Boltzmann Law). The emitted energy is a result of molecular vibration, and an object’s temperature is a measure of its vibration energy. Therefore, an object’s energy emission increases with its temperature. An object’s temperature depends on several factors: its recent thermal history, its reflectance and absorption characteristics, and the ambient (surrounding) temperature.

FLIR sensors detect the IR emission of objects in the scene and can “see” through haze and smoke and even in complete darkness. Although no universal definition exists for infrared (IR) energy, for imaging purposes, it is generally accepted as thermally emitted radiation in the 1 to 20 micron region of the electromagnetic spectrum. Currently, most military thermal imaging is performed in the 3 to 5 or 8 to 12 micron region. These regions are somewhat dictated by the IR transmittance windows of the atmosphere (Rash and Verona, 1992).

Thermal imaging sensors form their image of the outside world by collecting energy from multiple segments of the outside scene. The sensors convert these energy data into a corresponding map of temperatures across the scene. This may be accomplished by using one of several sensor designs. The two most common are the older *scanning arrays* and the newer *focal plane staring arrays*.

Typically a scanning array consists of a vertical row of sensor elements. This 1-D array is scanned horizontally across the focused scene, producing a 2-D profile signal of the scene. If desired, the scan can be reversed to provide an interlaced signal.

A focal plane array uses a group of sensor elements organized into a rectangular grid. The scene is focused onto the array. Each sensor element then provides an output dependent upon the incident infrared energy. Temperature resolution, the ability to measure small temperature differences, can be as fine as 0.1° C.

Acoustic sensors

The source inputs for auditory displays are often thought of as being radio-transmitted communications, automated voice commands, or artificially generated alert tones. Historically, any sensing of external sounds, such as engine sounds, weapons fire, ground vibration, etc., has been accomplished primarily by the human ear, and for lower frequencies, the skin. However, in the last decade there has been an increased interest in acquiring external sounds and using them for source identification and spatial localization, e.g., sniper fire from a specific angle orientation. This is accomplished through the use of acoustic sensors.

Acoustic sensor technology involves the use of microphones or arrays of microphones to detect, locate, track, and identify air and ground targets at tactical ranges. Target information from multiple widely-spaced acoustic sensor arrays can be digitally sent to a remote central location for real-time battlespace monitoring. In addition, acoustic sensors can be used to augment the soldier's long range hearing and to detect sniper and artillery fire (Army Materiel Command, 1997).

Acoustic sensors have been used for decades in submarines for locating other submarines. The earliest and most familiar is the “hydrophone.” The hydrophone is a device that detects acoustical energy underwater, similar to how a microphone works in air. It converts acoustical energy into electrical energy. Hydrophones in underwater detection systems are passive sensors, used only to listen. The first hydrophone used ultrasonic waves. The ultrasonic waves were produced by a mosaic of thin quartz crystals placed between two steel plates, having a resonant frequency of approximately 150 kilohertz (kHz). Contemporary hydrophones generally use a piezoelectric ceramic material, providing a higher sensitivity than quartz. Hydrophones are an important part of the Navy SONAR (SOund Navigation And Ranging) systems used to detect submarines and navigational obstacles. Directional hydrophones have spatial-localized sensitivity allowing detection along a specific direction. Modern SONAR has both passive and active modes. In active systems, sound waves are emitted in pulses; the time it takes these pulses to travel through the water, reflect off of an object, and return to the ship is measured and used to calculate the object's distance and often its surface characteristics.

No longer confined to the oceans, modern acoustic sensors are being deployed by the U.S. Army in both air and ground battlespace applications. These sensor systems have demonstrated the capability to detect, classify, and identify ground targets at ranges in excess of 1 km and helicopters beyond 5 km with meter-sized sensor arrays, while netted arrays of sensors have been used to track and locate battalion-sized armor movements over tens of square kilometers in non-line-of-sight conditions.

Regardless of the characteristics of individual sensor types, the sensors employed by an HMD system are designed to detect certain types of energy and present information about that energy to the user. The properties of the system are thus fundamentally defined by the sensitivity of the sensors to the energy they detect. For an observer to respond to this energy, the sensors must convert the detected energy into a format that can be detected by the observer's perceptual system. The converted energy is then displayed to the observer. This leads to the remaining main component of any HMD-the display.

Displays

A generic definition of a "display" might be "something used to communicate a particular piece of information." A liberal interpretation of this definition obviously should be extremely broad in scope. Examples would run the gamut from commonplace static displays (e.g., directional road signs, advertising signs, posters, and photographs) to dynamic displays (e.g., televisions, laptop computer screens, and cell phone screens).

Visual displays used in HMDs are so ubiquitous that they almost do not need any introduction. There are many different type of visual displays (e.g., cathode-ray tubes, liquid crystal, electroluminescent, etc.), but they all generate patterns of light on a 2-dimenional surface. Details about the properties, constraints, and capabilities of these display types are provided in Chapter 4, *Visual Helmet-Mounted Displays*.

Because they are less familiar to many people, we will describe auditory displays in more detail. Auditory displays use sounds to present information. These sounds can be speech-based, as with communications systems, or nonspeech-based, such as the "beep-beep" of a microwave oven emitted on completion of its heating cycle. Auditory displays are more common than might first be assumed. They are used in many work environments including kitchen appliances, computers, medical workstations, automobiles, aircraft cockpits, and nuclear power plants.

Auditory displays that use sound to present data, monitor systems, and provide enhanced user interfaces for computers and virtual reality systems are becoming more common (International Community for Auditory displays, 2006). Examples of auditory displays include a wide array of speakers and headphones.

Auditory displays are frequently used for alerting, warnings, and alarm-situations in which the information occurs randomly and requires immediate attention. The near omni-directional character of auditory displays that can be provided using an HMD is a major advantage over other types of auditory displays.

Long used primarily as simple alerts, the presentation of nonspeech-based sounds is increasing in its scope, effectiveness and importance. Sound is being explored as an alternate channel for applications where the presence of vast amounts of visual information is resulting in "tunnel vision" (Tannen, 1998). However, sound is sufficient in its own capacity to present information.

There is a vast spectrum of sounds available for use in auditory displays. Kramer (1994) describes a continuum of sounds ranging from audification to sonification. Audification refers to the use of "earcons" (Houtsma, 2004), a take-off on the concept of icons used in visual displays. An icon uses an image that "looks" like the concept being presented, e.g., a smiley face representing happiness; an earcon would use a sound that parallels that of the represented event. These are typically familiar, short-duration copies of real-world acoustical events. As an example, an earcon consisting of a sucking sound might be used in a cockpit to warn the aviator of fuel exhaustion.

Sonification refers to the use of sound as a representation of data. Common examples of sonification use include SONAR pings to convey distance information about objects in the environment and the clicks of a Geiger counter to indicate the presence of radioactivity. In both examples the sounds are the means of information presentation; but the actual sounds themselves are not meaningful. Instead, it is the relationship between the sounds that provide information to the user (in the case of SONAR, the relationship is between distance and time; for the Geiger counter, the relationship is between intensity and frequency).

It is worth reemphasizing that audification uses the structure of the sound containing the information, while sonification uses the relationship of sounds to convey the information. This implies that in the case of sonification, changing the specific sounds does not change the information, and even simple tones may be employed (Kramer, 1994).

The development of 3-D auditory displays for use in HMDs is both an example of the sophisticated level that auditory displays have achieved and an example of an application where an auditory display is superior to a visual display. The inherent sound localization produced by such displays can be used to directionally locate other Warfighters (friend and foe), threats, and targets (Glumm et al., 2005).

Auditory display technologies for HMD applications are not as diverse as visual display technologies. The dominant technology is the electro-mechanical or electro-acoustic transducer, more commonly known as a speaker, which converts electrical signals into mechanical and then acoustical (sound) signals. More precisely, it converts electrical energy (a signal from an amplifier) into mechanical energy (the motion of a speaker cone). The speaker cones, in turn, produce equivalent air vibrations in order to make audible sound via sympathetic vibrations of the eardrums.

An alternate method of getting sound to the inner ear is based on the principle of bone conduction. Headsets operating on this principle (referred to also as ears-free headsets) conduct sound through the bones of the skull (cranial bones). Such headsets have obvious applications for hearing-impaired individuals but have also been employed for normal-hearing individuals in auditory-demanding environments (e.g., while scuba diving) (MacDonald et al., 2006).

Bone conduction headsets are touted as more comfortable, providing greater stereo perception, and being compatible with hearing protection devices (Walker and Stanley, 2005). However, bone conduction acts as a low-pass filter, attenuating higher frequency sounds more than lower frequency sounds.

Auditory displays, their technologies and applications, are discussed further in Chapter 5, *Audio Helmet-Mounted Displays*.

Other issues

HMD systems face additional constraints because they are almost always a part of a larger system. In military settings, HMDs are almost always a part of the Warfighter's head protection system (i.e., helmet). As a result, the HMD must not introduce components that undermine the head protection crash worthiness of the system, e.g., impact and penetration protection (see Chapter 17, *Guidelines for HMD Designs*). One effect of this constraint is that the HMD components face restrictions on their weight and how their placement affects the center-of-mass of the combined HMD/helmet system.

Another issue that drives an HMD system design is how it interacts with the environment in which it is to be used. A key aspect is that the HMD needs to be relatively self-contained. That is, the HMD must be able to operate with a system that may change in several significant ways. While one wants the HMD to match appropriately with the larger system, it is not practical for a minor change in the larger system to necessitate a major redesign of the HMD. In addition to working well with various types of machine systems, an HMD needs to work well with various types of human users. While people's cognitive and perceptual systems are fairly similar, there can be significant differences, and the HMD needs to be functional for a variety of users. User capabilities may change over time, and an HMD needs to be usable despite these changes. Even with the best HMD development programs, system and user performance are usually evaluated under generally benign conditions and not under more realistic conditions where physical fatigue, psychological stress, extreme heat or cold, reduced oxygen levels, and disrupted circadian rhythms are present.

The HMD as a Human-Machine Interface: Statement of the Challenge

The task of designing an HMD that both meets the needs of the situation and matches the abilities of the user is a difficult one. The best efforts and intentions may still lead to a poor result. There are so many constraints on an HMD from the physical, task, and human parameters that something is almost certain to be suboptimal. Unfortunately, for the system to work well, everything must be just right.

Many of the difficulties derive from the need for an HMD to behave robustly in a complex system. From engineering and manufacturing perspectives, an HMD needs to be relatively self-contained. Unless the HMD behaves robustly, the manufacture or design of the system components can bog down development. For example, changes to one part of an HMD system (e.g., a microphone) must be relatively isolated from other parts of the HMD (e.g., the visual display).

Having described the human and machine aspects of an HMD, we are now ready to discuss how the properties of these two systems influence the design of the human machine interface. The HMI challenge is to address the following question: *How to use robust technology to organize and present information in a way that meets the expectations and abilities of the user?*

Clearly, a satisfactory solution to the challenge requires careful consideration of both the machine and human systems. Current engineering techniques tend to focus on ensuring that the machine side of the system behaves according to design specifications in a way that ensures that appropriate sensor information is present on the display. There are remaining issues to be resolved, and active development of new technologies will be needed to address these issues. For example, a continued effort in the development of miniature display technologies can improve weight, center-of-mass offset and heat generation, which in turn improves comfort. Development of more intuitive symbology (an ongoing effort) will reduce workload and error rate.

The more difficult aspect of the challenge, and the part that needs more progress, is understanding the human side of the system. Information on an HMD may be present but not be perceived, interpreted, or analyzed in a way that allows the human user to take full advantage of the HMD. Working with the human side is difficult because many aspects of human perception and cognition are not fully understood and thus there is little information available to guide the design of an HMD. Moreover, humans are exceptionally complex systems that can behave in fundamentally different ways in different contexts. These behavior changes make it very difficult to predict how they will perform in new situations. Indeed, one commonly noted aspect of fielded HMD designs is that users do not follow the "rules" for the system and instead adapt new strategies to make the HMD operate in some unexpected way. A classic example of HMD users not following the rules is AH-64 Apache pilots using the Integrated Helmet and Display Sighting System (IHADSS). This HMD has a very small exit pupil that results in great difficulty maintaining the full FOV. To compensate, pilots use a small screwdriver to minimize the image on the display, thereby allowing viewing of the full FOV (but no longer in a one-to-one relationship with the sensor FOV) (Rash, 2008).

There has been substantial progress on some aspects of the challenge. For example, studies of human vision indicate the required luminance levels that are needed for HMD symbology to be visible in a wide variety of background scenes. Likewise, the intensities and frequencies of sound stimuli that can be detected by human users are well understood and promote guidelines for HMD design (Harding et al., 2007).

Things are more challenging when the information detected by sensors does not correspond to aspects of the world that are usually processed by human perceptual systems. For example, infrared vision systems that detect sources of heat energy can provide a type of "night vision." The information from sources of heat energy is usually displayed as a visual display. Such a display requires some type of conversion from heat energy to the visible ranges of light energy. This conversion can lead to misinterpretations of the information when aspects of the sensor information are mapped onto display properties in a way that is inconsistent with the biases of the visual or cognitive systems. In the case of heat sensors, it is fairly easy to display the intensity of heat emissions as a light intensity map. This provides the human observer with an unambiguous description of what the sensor has detected. However, a light intensity map tends to be interpreted as something produced by objects that reflect illuminated light. As a result, the visual display can be misinterpreted, with columns of heat interpreted as solid objects and false identification of figure and ground.

Adding color to the display of such a system may provide additional clarity about the properties of the heat emissions, but can lead to even further confusion about the properties of the objects in the environment. In normal vision, different colors correspond to changes in the properties of surfaces (e.g., fruit versus leaves), but may correspond to something else entirely on a visual display.

Thus, the great benefit of HMDs, that they can display a wide array of sensor information, also exposes them to great risk, that they display information in a way that is inconsistent with the properties of the observer.

In optimizing the HMI for the HMD, the electrical engineer might investigate how to build better buttons and connectors (or other physical components of the HMD); the human factors engineer might investigate how to design more legible/audible and intelligible labels, alerts or instructions (e.g., perhaps, the characteristics of the symbology presented via the HMD); the ergonomicist might investigate the anatomy and anthropometry of the user population (e.g., head dimensions and interpupillary distance); but in this book we will focus on investigating HMD design from the perspective of the *in toto* human visual, auditory and neural systems (i.e., sensory, perceptual and cognitive functions). In doing so, the bidirectional flow of information will be studied via the HMD, through the sense organs (primarily the eyes and ears), through the visual and auditory pathways, through the thalamus, to and from the respective cortices. The HMI concept adopted here will incorporate the relationship between the HMD design and the user's visual and auditory anatomy and physiology, as well as the processes by which we understand sensory information (perception) and the neural activities associated with recognition, memory, and decision making with this information (cognition).

All of the issues are addressed in the following chapters. There is, as yet, no complete solution to the HMI challenge, but progress is being made in many areas. One goal of this book is to identify where solutions do exist, identify situations that require additional study, and outline possible solutions to some of those problem situations.

References

Adelson, E.H. (2000). Lightness perception and lightness illusions. In Gazzaniga, M. (Editor), *The New Cognitive Neurosciences*. Cambridge, MA: MIT Press, 339–351.

Army Materiel Command. (1997). *Command and Control Research Program. Army Science and Technology Master Plan (ASTMP).* Alexandria, VA: Army Materiel Command.

Bai, L., Gu, G., Chen, Q., and Zhang, B. (2001). Information obtaining and fusion of color night vision system. *Data Mining and Applications, Proceedings of SPIE*, 4556, 65-70.

Britten, K.H., and Wezel, R.J.A. van. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nature Neuroscience*, 1, 59-63.

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25(5), 975-979.

Fermüller, C., Pless, R., and Aloimonos, Y. (1997). Families of stationary patterns producing illusory movement: Insights into the visual system, *Proc. Royal Society London B*, 264, 795-806.

Fletcher, H., and Munson, W.A. (1933). Loudness: Its definition, measurement, and calculation. *Journal of the Acoustical Society of America*, 5, 82-108.

Gibson, J.J. (1950). *Perception of the Visual World.* Boston: Houghton Mifflin.

Gibson, J.J., Olum, P., and Rosenblatt, F. (1955). Parallax and perspective during aircraft landings, *American Journal of Psych*ology, 68, 372-385.

Gilchrist, A., Kossyfidis, C., Agostini, T., Li, X., Bonato, F., Cataliotti, J., Spehar, B., Annan, V., and Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, 106(4), 795-834.

Gilger, M. (2006). Information display-the weak link for NCW. *Defense, Security, Cockpit, and Future Displays II, Proceedings of SPIE,* 6225, 622511-11 to 622511-12.

Glumm, M.M., Kehring, K.L., and White, T.L. (2005). Effects of tactile, visual, and auditory cues about threat location on target acquisition and attention to visual and auditory communications. *Human Factors and Ergonomics Society Annual Meeting Proceedings, Cognitive Engineering and Decision Making,* 347-351.

Goldstein, E.B. (2007). *Sensation and Perception, 7th Edition*. Pacific Grove, CA : Thomson Wadsworth.

Hackos, J., and Redish, J. (1998). *User and task analysis for interface design*. New York: John Wiley and Sons.

Harding, T.H., Martin, J.S., and Rash, C.E. (2007). The legibility of HMD symbology as a function of background local contrast. *Head- and Helmet-Mounted Displays XII: Design and Applications, Proceedings of SPIE*, 6557, 622570D.

Houtsma, A. (2004). Nonspeech audio in helicopter aviation. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL 2004-03.

International Community for Auditory Displays (ICAD). (2006). What is ICAD? Retrieved May 10, 2007 from http://www.icad.org.

Kahneman, D., and Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39, 341-350.

Keller, K., and Colucci, D. (1998). Perception and HMDs: What is it in head-mounted displays that really makes them so terrible? *Helmet- and Head-Mounted Displays III, Proceedings of SPIE,* 3362, 46-53.

Kramer, G. (1994). An introduction to auditory display. In: Kramer, G. (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces.* Reading, MA: Addison Wesley.

Logvinenko, A.D., Adelson, E.H., Ross, D.A., and Somers, D.C. (2005). Straightness as a cue for luminance edge classification. *Perception and Psychophysics*, 67, 120-128.

MacDonald, J., Henry, P., and Letowski, T. (2006). Spatial audio: A bone conduction interface. *International Journal of Audiology*, 45, 595-599.

National Research Council (1997). *Tactical display for soldiers: Human factors considerations*. Washington, DC: National Academy Press.

Rash, C.E. (Editor). (2001). *Helmet-mounted displays: Design issues for rotary-wing aircraft*. Bellingham, WA: SPIE Press.

Rash, C.E. (2008). Information obtained via interviews with AH-64 Apache pilots.

Rash, C.E., and Verona, R.W. (1992). The human factor considerations of image intensification and thermal imaging systems. In: Karim, M.A. (Editor), *Electro-Optical Displays*. New York: Marcel Dekker.

Sibley, F. (1964). Book review: *Perception and the Physical World*, *The Philosophical Review*, 73, 404-408.

Tannen, R.S. (1998). Breaking the sound barrier: Designing auditory displays for global usability. *Proceedings of 4th Conference on Human Factors and the Web*. Basking Ridge, NJ.

Toet, A. (2003). Color the night: applying daytime colors to nighttime imagery. *Enhanced and Synthetic Vision, Proceedings of the SPIE,* 5081, 168-178.

Walkenstein, J.A. (1999). Color night vision: A critical information multiplier. *Proceedings of American Lighting Institute.*

Walker, B., and Stanley, R. (2005). Thresholds of audibility for bone conduction headsets. Presented at the ICAD 05, Eleventh Meeting of *the International Conference on Auditory Display*, Limerick, Ireland, July 6-9.

Walton, G.E., Bower, N.J.A., and Power, T.G.R. (1992). Recognition of familiar faces by newborns. *Infant Behavior and Development*, 15, 265-269.

Warren, W.H. (1998). Heading is a brain in the neck. *Nature Neuroscience*, 1, 647–649.

Wason, P., and Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, 63-71.

Wilkie, R.M., and Wann, J.P. (2003). Eye-movements aid the control of locomotion. *Journal of Vision*, 3, 677-684.

Part Two

Helmet-Mounted Displays

Historically, the helmet-mounted display (HMD) has been thought of as an optical/visual system. Thus, it is important to understand the optical parameters involved in the design of HMDs and the impact these parameters have on image quality. Equally important are the characteristics of the sensor(s) that produce the visual imagery. However, advanced HMD designs include significant audio information. This requires the HMD designer to also consider auditory factors such as noise attenuation and communication speech intelligibility.

3 INTRODUCTION TO HELMET-MOUNTED DISPLAYS

Michael M. Bayer
Clarence E. Rash
James H. Brindle

In order to fully understand the sensory, perceptual, and cognitive issues associated with helmet-/head-mounted displays (HMDs), it is essential to possess an understanding of exactly what constitutes an HMD, the various design types, their advantages and limitations, and their applications. It also is useful to explore the developmental history of these systems. Such an exploration can reveal the major engineering, human factors, and ergonomic issues encountered in the development cycle. These identified issues usually are indicators of where the most attention needs to be placed when evaluating the usefulness of such systems.

New HMD systems are implemented because they are intended to provide some specific capability or performance enhancement. However, these improvements always come at a cost. In reality, the introduction of technology is a tradeoff endeavor. It is necessary to identify and assess the tradeoffs that impact overall system and user sensory systems performance. HMD developers have often and incorrectly assumed that the human visual and auditory systems are fully capable of accepting the added sensory and cognitive demands of an HMD system without incurring performance degradation or introducing perceptual illusions. Situation awareness (SA), essential in preventing actions or inactions that lead to catastrophic outcomes, may be degraded if the HMD interferes with normal perceptual processes, resulting in misinterpretations or misperceptions (illusions).

As HMD applications increase, it is important to maintain an awareness of both current and future programs. Unfortunately, in these developmental programs, one factor still is often minimized. This factor is how the user accepts and eventually uses the HMD. In the demanding rigors of warfare, the user rapidly decides whether using a new HMD, intended to provide tactical and other information, outweighs the impact the HMD has on survival and immediate mission success. If the system requires an unacceptable compromise in any aspect of mission completion deemed critical to the Warfighter, the HMD will not be used. Technology in which the Warfighter does have confidence or determines to be a liability will go unused.

Defining the Helmet-Mounted Display

Melzer and Moffitt (1997) describe an HMD as minimally consisting of "an image source and collimating optics in a head mount." From the perspective of U.S. Army rotary-wing aviation, Rash (2000) extended this description to include a coupling system that uses head and/or eye position and motion to slave one or more aircraft systems, typically a head-directed sensor. Using this description, Figure 3-1 presents a basic block diagram in which there are four major elements: image source (and associated drive electronics), display optics, helmet, and head/eye tracker. The *image source* is a display device upon which sensor imagery is reproduced. Early on, these sources were miniature cathode-ray-tubes (CRTs) or image intensification (I^2) tubes. More recently, miniature flat panel display technologies have provided alternate choices. The *display optics* is used to couple the display imagery to the eye. The optics unit generally magnifies and focuses the display image. The *helmet*, while providing the protection for which it was designed originally, also now serves as a platform for mounting the image source and display optics. The *tracking system* couples the head orientation or line-of-sight with that of the pilotage sensor(s) and weapons.

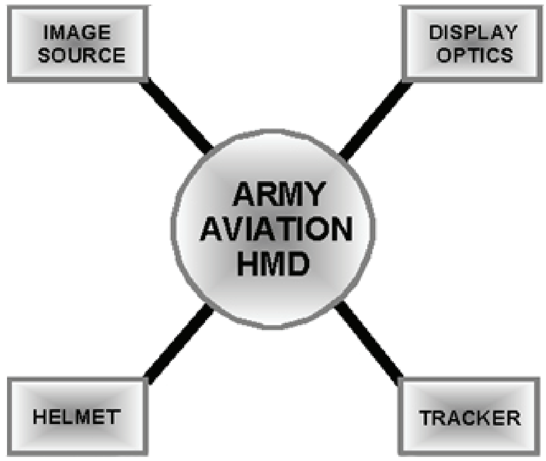


Figure 3-1. Block diagram of a basic U.S. Army rotary-wing aviation HMD.

However, this extended description of HMDs is still limited by its close association with use in military rotary-wing aircraft as well as being focused only on the visual system. Manning and Rash (2007) provide a more generalized description of visual HMDs that is applicable to both military and commercial applications, where the name "head-worn displays" (HWDs) has been gaining acceptance. The same basic four building blocks are employed but are expanded in scope:

- A *mounting platform*, which can be as simple as a headband or as sophisticated as a full flight helmet. In addition to serving as an attachment point, it must provide the stability to maintain the critical alignment between the user's eyes and the HWD viewing optics;
- An *image source* for generating the information imagery that is optically presented to the user's eyes. Advances in miniature displays have produced a wide selection of small, lightweight and low-power choices at moderate cost, while meeting the demands of perceptual intensity and resolution (See Chapter 4, *Visual Helmet-Mounted Displays*.);
- *Relay optics*, which transfer to the eye(s) the information at the image source. Relay optics typically consists of a sequence of optical elements (mostly lenses) that terminates with a beam-splitter (combiner). Initial designs for visual applications were monocular with a single beam-splitter in front of one eye, but as miniature display technologies develop, binocular designs are becoming dominant;[1] and,
- A *head-tracker*, which is optional if the HWD is used only to present status information using non-spatially-referenced symbols. However, it often is required if external (outside) imagery is supplied by a sensor or a synthetic database. If such imagery is to be presented, the user's directional line-of-sight must be recalculated continuously (updated) and used to point the sensor or to select the synthetic imagery data correlated with the user's line-of-sight. Presentation of head-referenced information (imagery and/or symbology) via a head tracker requires a preflight calibration procedure called boresighting, which aligns the sensor's and user's lines-of-sight.

Each of these fundamental HMD building blocks has engineering, sensory, perceptual, cognitive, and ergonomic considerations that will be explored in future chapters. All of these engineering and human factor

[1] For the audio realm, three-dimensional (virtual) audio technologies are being developed. Tactilely, small vibrators are being explored for 360 degree enhanced awareness.

considerations are interrelated; therefore, tradeoffs are required in order to achieve a design that will be functionally acceptable for a specific operational application. As the tradeoffs are implemented, it is essential that the developer and the user be aware of the performance implications of these tradeoffs. The following sections will use the visually based HMD as an example of these considerations.

Classifying Visual Helmet-Mounted Display Designs

Since visual HMDs are complicated systems, there are several classification schemes that can be employed. These include those based on image source, image display technology, imagery presentation mode, and optical design approach. The image formed by an optic system, e.g., an HMD, can be real or virtual. At a practical level, the image is real if the light rays to be focused by the eye or a camera are spreading farther apart, i.e., diverging. This is the case when we view a real object directly or in a flat mirror, a photograph, the screen at a movie theater, or view an image focused by a convex lens from beyond its focal plane. The image formed is outside the optical system; the light rays (or wave front) from the image points that reach the eye are diverging. An image is virtual if the light rays to be focused by the eye are moving closer together, i.e., converging. Examples of virtual images include those from telescopes or microscopes focused by the user, a real scene viewed through a concave lens, or looking into a convex lens from a point inside its focal plane.

Real-image HMD designs are rare. A direct-view image source like a miniature liquid crystal display (LCD) would have to be located no closer than reading distance, which is not practical. Putting the appropriate optics in front of the miniature display to move it closer to the eye would likely make the image virtual. All currently fielded HMDs are set to produce virtual images (although a slightly diverging system than produces some accommodation in the eye for presented symbology while viewing a real scene through the display may have some attentional advantages).

Virtual image displays offer several advantages (Seeman et al., 1992). At near optical infinity, virtual images theoretically allow the eye to relax (reducing visual fatigue) and provide easier accommodation for older users. By providing a virtual image, a greater number of individuals (but not all) can use the system without the use of corrective optics. A collimated image also reduces effects of vibration that produces retinal blur.

Shontz and Trumm (1969) categorize HMDs based on the mode by which the imagery is presented to the eyes. They define three categories: One-eye, occluded; one-eye, see-through; and two-eye, see-through. In the *one- eye, occluded* type, imagery is presented to only one eye, to which the real world is blocked, with the remaining eye viewing only the real world. The *one-eye, see-through* type, while still providing imagery to one eye, allows both eyes to view the real world. (Note: The optics in front of the imagery eye will filter the real world to a lesser or greater degree.) The Integrated Helmet and Display Sighting System (IHADSS)[2] employed on the AH-64 Apache helicopter is an example of this type. In the *two-eye, see-through* type, imagery is presented to both eyes, while the real world also is viewed by both eyes.[3] The Thales TopOwl™ is an example of this type.

Another classification scheme, which parallels the three types described above, uses the terms monocular, biocular, and binocular. These terms refer to the presentation mode of the symbology and/or sensor imagery by the HMD. For our usage, *monocular* means the HMD sensor imagery is viewed by a single eye; *biocular* means the HMD provides two visual images from a <u>single</u> sensor or multiple sensors, but each eye sees exactly the same image from the same perspective; *binocular* means the HMD provides two visual images, one for each eye, from two sensors displaced in space, thus providing perspective. (Note: A *binocular* HMD can use a single sensor, if the sensor is manipulated to provide two different perspectives of the object scene.) Both *biocular* and *binocular* HMDs will have two optical channels (one for each eye). Note that a two-eyed HMD presenting *biocular* imagery

---

[2] The IHADSS system now is owned and manufactured by Elbit EFW, Fort Worth, TX.

[3] Not included in this classification scheme is a “two-eye, occluded” category such as Night vision Goggles (NVGs)

from one sensor/database is still capable of presenting *binocular* symbology overlays as long as it has two independently controllable image sources

Typically, binocular HMDs use optical designs that fully overlap the images in each eye. In such HMDs, the field-of-view (FOV) is limited to the FOV of the display optics. However, in order to achieve larger FOVs, recent HMD designs partially overlap the images from two optical channels. This results in a partially-overlapped FOV consisting of a central binocular or binocular region (simultaneously seen by both eyes) and two monocular flanking regions (each seen by one eye only) (Figure 3-2). Such overlapping schemes can be implemented by either divergent or convergent overlap designs. In a divergent design, the right eye sees the central overlap region and the right monocular region, and the left eye sees the central overlap region and the left monocular region (Figure 3-3a). In a convergent design, the right eye sees the central overlap region and the left monocular region, and the left eye sees the central overlap region and the right monocular region (Figure 3-3b).

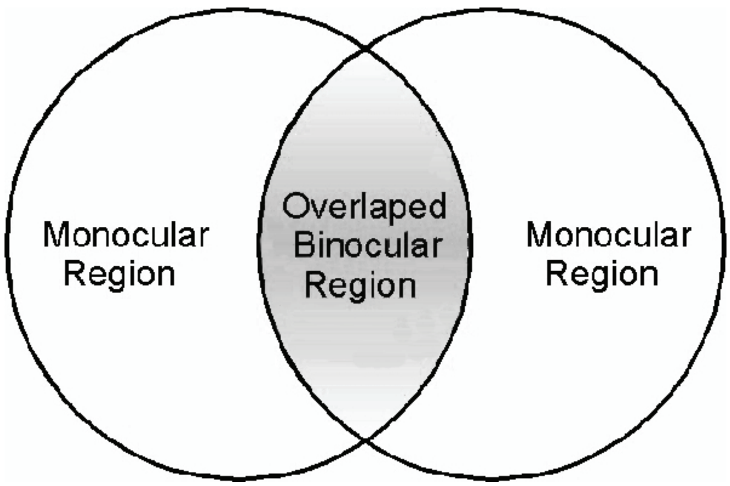


Figure 3-2**.** Partially overlapped FOV with a central binocular region and two monocular regions

The IHADSS is an example of a monocular HMD; the Aviator's Night Vision Imaging System (ANVIS) is an example of a 100% overlapped binocular HMD; and the Kaiser Electronics' CRT-based Helmet Integrated Display Sight System (HIDSS) design is divergent and has an overlap of approximately 30% (based on a 17º overlap region within the 52º horizontal FOV).

Classifying HMDs by optical design is even more complicated. The simpler and more predominant types use optical designs based on reflective and refractive lens elements that relay the HMD image source to the eye. A standard characteristic of these designs is the presence of a final partially-reflective element(s) positioned in front of the user's eye(s) called "combiners" (Wood, 1992). These elements combine the see-through image of the real world with the reflected image of the HMD image source. Reflective/refractive optical designs will be discussed in detail in Chapter 4, *Visual Helmet-Mounted Displays*.

Another HMD type is based on a visor projection design (e.g., Cameron and Steward, 1994). A simple diagram of this design approach is presented in Figure 3-4. The image source(s) is usually mounted around (top/side) the helmet, and the image is relayed optically so as to be projected onto the visor where it is reflected back into the user's eye(s). The advantages of visor projection HMDs include lower weight, improved center-of-mass (CM), increased eye relief, and maximum unobstructed visual field. A possible deficiency is image degradation that can result in a high vibration environment. An optical problem that can show up with this design is the production of ghost images. Also, this design requires that the visor be able to be placed consistently at the same position. Recently, visor projection designs have been revisited (Chapter 4, *Visual Helmet-Mounted Displays*).

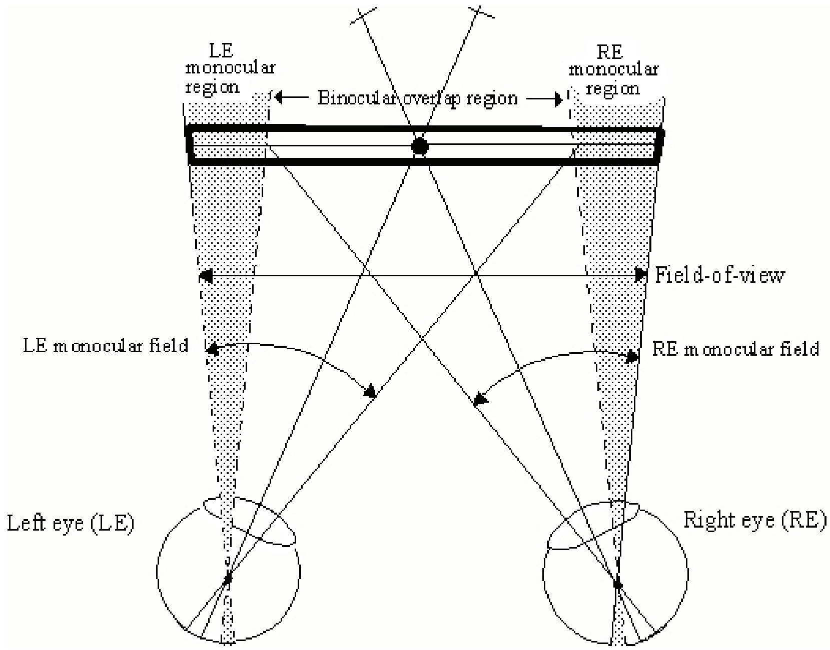


Figure 3-3a. Visual interpretation of the divergent display mode.

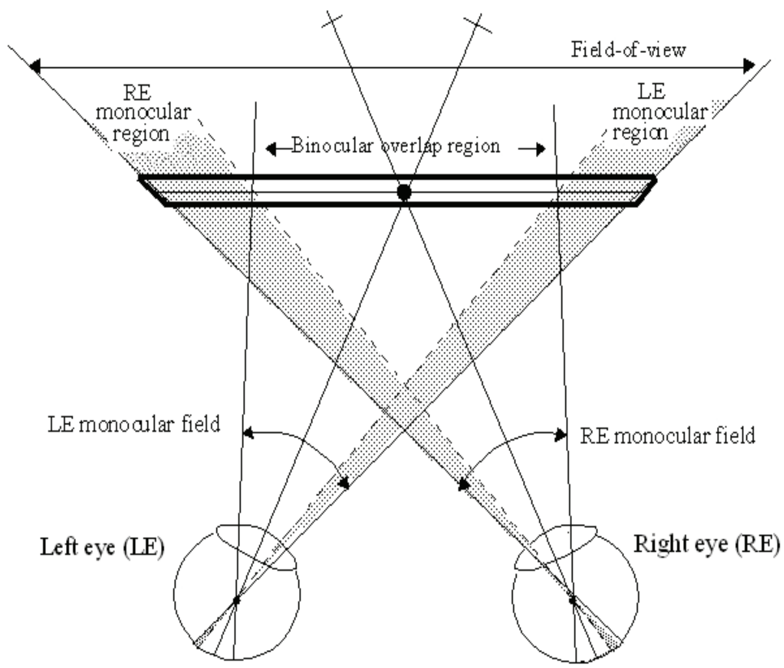


Figure 3-3b. Visual interpretation of the convergent display mode.

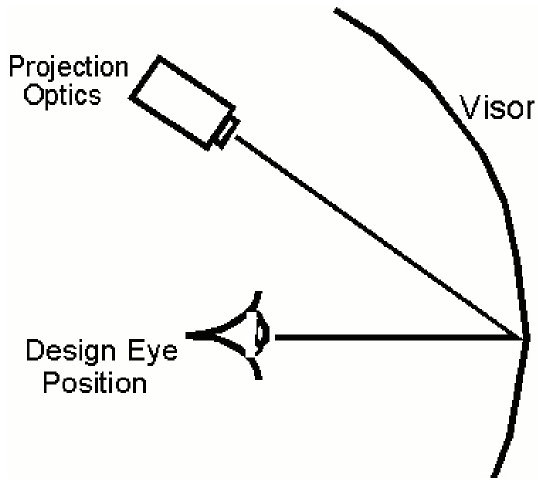


Figure 3-4. Visor projection HMD design approach.

Another approach, which again allows for low weight and provides a compact design, is one using holographic optical elements (Vos and Brandt, 1990). A holographic combiner is used to merge the standard combiner function with the collimation function usually performed by an additional refractive optical element. This merging implies that the holographic combiner acquires optical power, hence the term *power combiner* (Wood, 1992). In some designs, the visor serves as the combiner, with a holographic coating on the visor substrate. Disadvantages of this approach include the problem of preventing humidity and temperature effects from degrading the holograms. Considerable progress has been made in mitigating these problems in the last few years.

One of the most recent entries into HMD design approaches is the use of lasers that scan an image directly onto the retina of the user's eye (Johnston and Willey, 1995). Figure 3-5 provides a diagram of the basic retinal scanning approach. This approach eliminates the need for a CRT or flat panel (FP) image source, offering the potential of improving both weight and CM. Other cited advantages of this system include diffraction (and aberration) limited resolution, small volume (for monochromatic), full color capability, and high brightness potential. Disadvantages, at least potentially, include scanning complexity, susceptibility to high vibration environments (as with helmet slippage in military environments), limited exit pupil size, and safety concerns.

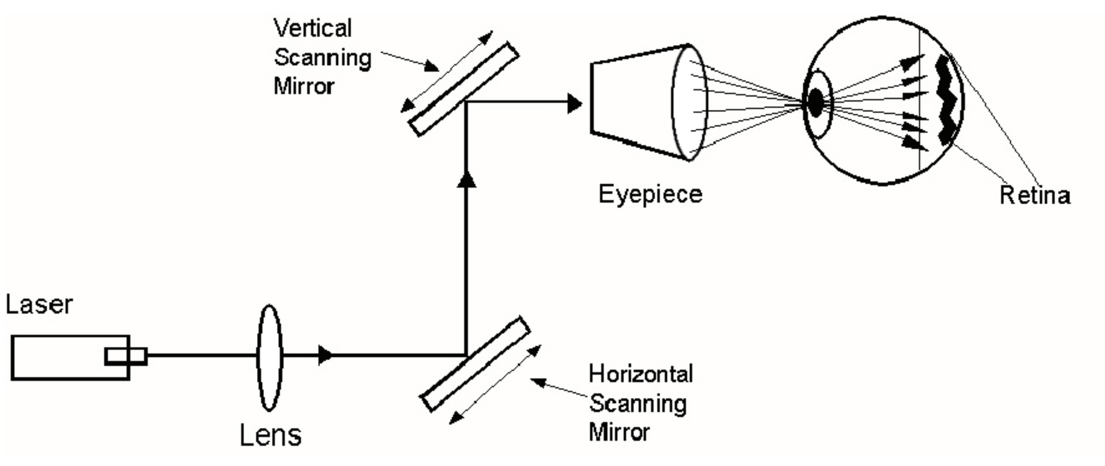


Figure 3-5. Basic diagram of retinal scanning display (adapted from Proctor, 1996).

A recent optical design for HMDs developed by BAE Systems uses wave-guide technology. This system uses holographic optics embedded between two transparent plates to direct the image to the eye. The potential advantages to this system are simplicity, large eye relief, ability to use in conjunction with existing night vision goggles (NVGs), lower cost, reduced weight and ability to adapt to existing military helmets. Although most of the disadvantages are unknown at this time, safety related to the plate placed in front of the eye and the eventual FOV have not been fully addressed. This approach is used in BAE System's Q-Sight™ HMD discussed in the *Current and Future HMD Programs* section of this chapter.

Regardless of the actual optical approach used, a visual aviation HMD also must include an image source, a head/eye tracker (if sensor is remotely located), and a helmet platform. At one time, the traditional approach was to integrate the optics and image source into a subsystem which was then mounted onto an existing helmet (Melzer and Larkin, 1987). This after the fact add-on approach was used with ANVIS. As one might expect, attaching one subsystem to another subsystem may not produce the optimal design. Instead, an integrated approach in which all elements and components of the HMD are designed in concert generally will result in the best and most functional overall design. The IHADSS was the first HMD product of the integrated approach, i.e., the helmet and the HMD optics were developed as a system, even though the optics is a removable component.

Even when using an integrated approach, the desired application of an HMD will impact design, leading to a variety of configurations. There is no one-design-fits-all scenario. In fact, the various missions, and the conditions under which they must be performed, are so different, that a single HMD design, while optimal for one set of conditions, may be significantly deficient for other mission scenarios. A solution to this problem may be a modular approach (Bull, 1990), where the HMD system consists of a base mounting unit (e.g., helmet platform), and interchangeable modules that can be attached, each for a specific set of mission requirements. This modular approach can be effective as long as an integrated approach is used that does not compromise the basic requirements of any subsystem. For example, the helmet, while now being used as a platform to attach optics, still must serve its primary functions of providing impact, visual, and acoustical protection. The HIDSS HMD design for the now cancelled U.S. Army Comanche program was an example of the modular approach.

The visually-coupled system (VCS) concept

Head-position sensing or head tracker technologies provide the pilot's/operator's "caged eyeball" line-of-sight as a control input to the aircraft/vehicle and its on-board sensors and weapons. This class of head-mounted system has sometimes been called a helmet-mounted sight. HMD technologies provide virtual image display capability integral to the user's helmet. When combined, they form a class of systems many times referred to across the military community as VCS, as illustrated in Figure 3-6. With closed loop VCS, the head tracker technology serves as the control path input to sensors, weapons, avionics, or the vehicle itself, while the HMD technology provides the display symbology/imagery feedback. It should be noted that even the most basic head tracker requires at least a simple display reference a "crosshair" or "reticle" so the user knows what line-of-sight is being sensed. It is also worth noting that the image intensification technology (commonly referred to as NVGs) that has evolved over this same timeframe represents a "self-contained" VCS, in that NVGs present spatially-referenced image intensification information to the wearer.

VCS take advantage of the psycho-motor skills of the operator to provide an intuitive visual interface to the vehicle, its on-board systems, and the surrounding environment. VCS provide a "look-and-shoot" vs. a "point-the-vehicle-and-shoot" capability for effective targeting of airborne and ground, and stationary and moving ground targets. This class of systems provides an expanded off-axis visual capability for the entire range of mission requirements. As time has gone on, there has been an increase in situations where the individual Warfighter is the "weapon platform" of choice with rapid adaptability and real-time decision-making before the enemy can react. Human systems, and in particular, visually-coupled display systems, optimize and sustain the human role in combat operations.

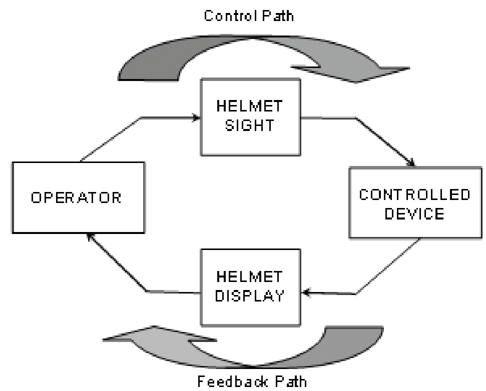


Figure 3-6. Visually-coupled system concept block diagram.

The History of Helmet-Mounted Displays

The official history of HMDs starts almost a century ago, with Albert Bacon Pratt, of Lyndon, Vermont. During the height of World War I, between 1915 and 1917, Pratt was awarded a series of U.S. and U.K. patents (Marshall, 1989), for an “Integrated Helmet Mounted Aiming and Weapon Delivery System” for a marksman (Figure 3-7).

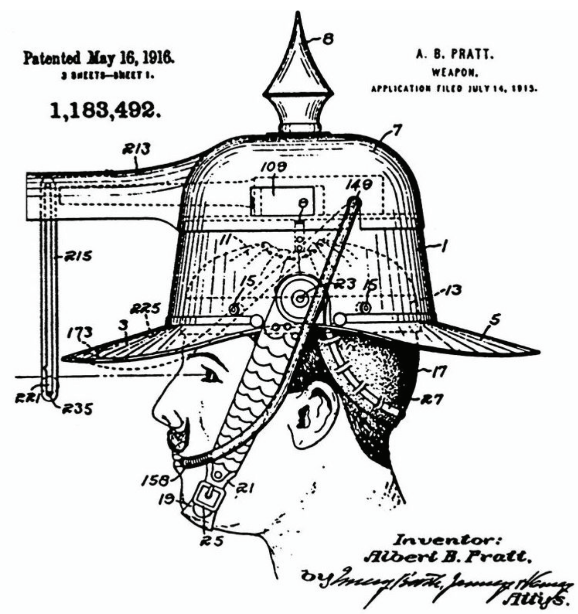


Figure 3-7. Albert Pratt’s helmet-mounted display (Marshall, 1989).

Pratt, a chemical engineer, claimed a few features in his patent that have survived through the years and are as valid today as they were 100 years ago. A couple of comparisons between Pratt’s patent claims and features of today’s HMD designs will help establish his design as the precursor of current HMDs.

- Size and Fit

“The helmet preferably will be made in two sizes, a large size and a small size. To adapt the helmets to fit different size heads the lower section is provided with flexible linen.”

Today’s flying helmets are designed in small, medium, large and extra large sizes; the liner is customized to individual pilots.

- Target Acquisition

“The gun is automatically aimed unconsciously to the turning of the head of the marksman in the direction of the target. In self-protection one instinctively turns the head in the direction of attack to see the enemy. Thus, the gun is automatically directed toward the target.”

Today’s HMDs embody the same “look-and-shoot” philosophy; sophisticated technology with Kalman filtering tracks the instantaneous pilot’s line-of-sight to guide missiles to the target.

- Dual Use

On a lighter note, the crown of Pratt’s helmet (Item #7) doubled as a cooking pan, with the gun barrel safeguard (Item #213) serving as the handle. Whereas some might think the top spike (Item #8) is intended for hand-to-hand combat, it is simply stuck into the ground to support the pan while dining in the field.

Also, despite conducting an in-depth literature research, the authors of this chapter were not able to identify a like-functionality for modern helmets.[4] Advantage, Pratt!

The concept and the potential applications of HMDs in aircraft cockpits have fascinated military aviation strategists for decades. The idea of placing a virtual image focused at infinity in the visual path of the pilot and overlaying computer-generated images so that mission critical information is always available with “eyes-out,” has mobilized incredible technical and financial resources over the last decades. It is generally acknowledged that an HMD, when part of a Visually Coupled System (VCS), is among the most valuable visual aids in the arsenal of a military pilot. Experience has shown that nothing can be added to a tactical aircraft that give more “bang for the buck” or operational payoff-per-pound-added than a VCS.

Military HMD development: historical overview

The various militaries across the world have actively pursued the research, development, application, and fleet introduction of a variety of helmet-mounted technologies for over forty years. A complete overview of the HMD technology development over the last forty years would be difficult as there have been hundreds of head tracker and HMD development efforts. Additionally, in recent years the concept of virtual reality has spurred interest in HMDs within industry and the general population. One artifact of the vast interest in HMDs has been the failure of the military (and more recently the commercial) communities to develop and accept an overall plan that would establish unambiguous guidelines for HMD development, not that such efforts have not been attempted.

Within the U.S., in 1995 (Brindle, Marano-Goyco, and Tihansky, 1995) under the auspices of a Tri-Service Working Group reporting to the Office of the Undersecretary of Defense for Research and Engineering, a technology-development taxonomy was established to help the HMD community properly categorize and

---

[4] The modern plastics-composite helmet has lost considerable functionality, as the early steel-pot was used to cook, wash, dig, etc.

articulate the diverse spectrum of research and development (R&D) programs underway at any point in time. The taxonomy's main categories included:

- *Human System Integration*, dealing mostly with efforts on safety, anthropometry, vision, situation awareness, spatial disorientation, symbology, and audio performance/hearing protection.
- *Component Development*, focusing on optics, image intensification, head trackers, image sources, three-dimensional (3-D) audio, and voice recognition; interconnect technology/ systems, and symbol generation/graphics.
- *System Development* for air and ground vehicles, the individual warrior, and simulation.
- *System Integration and Analysis*, coordinating all R&D efforts dealing with a) helmet system integration – both the integration of the various VCS components with each other and with existing personal life support equipment; and b) vehicle/laboratory system integration for properly integrating the helmet-mounted system with the vehicle and the vehicle sensors/weapons/subsystems.
- *Application Demonstration/Measurement and Evaluation*, oriented toward laboratory measurements, simulation evaluations, flight-worthiness testing, flight evaluations, concept demonstrations and field trials.

In order to highlight and summarize the wide range of HMD developments over the past decades, it may be useful to briefly describe those efforts that have progressed all the way from initial R&D, through prototyping and production, and into fielding (even if limited). Some of these programs will be summarized in greater detail in the *Current and Future HMD Programs* section of this chapter.

One of the earliest (1970s) sighting HMD systems to be fielded was the electro-mechanical linkage head-tracked sight used to direct the fire of the gimbaled gun in the U.S. Army's AH-1G Huey Cobra attack helicopter (Braybrook,1998). The pilot aimed the gun by superimposing a helmet-mounted reticle over the target.

Not too long after the Cobra head tracker system (1973-1979), the Navy introduced an electro-optical head-tracking system into its later Phantom models F-4J and F-4N fixed-wing jet aircraft, coupled with the radar and AIM-9H Sidewinder missiles (Klass, 1972). The Visual Target Acquisition System (VTAS), shown in Figure 3-8, consisted of photo diodes on either side of a "halo assembly" that mounted on the standard fixed-wing flight helmet. Sensor surveying units on either side of the cockpit scanned the helmet in the "head motion box." The pilot used a visor-projected reticle and cueing discretes to interface with the fire-control radar and missiles for daytime, off-boresight, air-to-air targeting.

As was the case with the Cobra application, these head trackers yielded a significant reduction in the time required to bring weapons to bear on target. VTAS was discontinued in the 1970's (Dornheim, 1995) due to its technological limitations.

The first complete VCS system to see operational use was the introduction in the early 1980s of the IHADSS by the U.S. Army in the AH-64 Apache attack helicopter (Figure 3-9). The head tracking technology in the IHADSS was the electro-optical technology similar to the Navy VTAS. However, the HMD technology was much more capable and provided higher resolution dynamic video imagery by using a miniature 1-inch CRT with relay optics.

The monocular IHADSS serves as the crew interface for both the pilot and copilot/gunner. The pilot's IHADSS is interfaced with a 30° x 40°-FOV thermal sensor (mounted on the nose of the aircraft) to form a head coupled, one-to-one magnification pilotage system. The copilot's IHADSS is interfaced with a switchable-FOV thermal targeting sensor to form an effective off-boresight interface with the head-slaved gun and missiles. In both cases, the appropriate flight-control or fire-control symbology is mixed electronically with the thermal imagery. The systems have been used effectively for both day and night missions for almost three decades (Rash, 2008).

Recently, in the fixed-wing community, the U.S. Air Force and U.S. Navy have introduced the Joint Helmet-Mounted Cueing System (JHMCS) into the F-15, F-16 and F-18 aircraft. The JHMCS utilizes magnetic head tracker technology and provides a monocular, visor-projected display of stroke-written dynamic symbology from a ½-inch miniature CRT and relay optics (Figure 3-10). The JHMCS provides a daytime air-to-air and air-to-ground off-boresight targeting capability, especially valuable when used with high off-boresight missile seeker technology.

Figure 3-8. Visual Target Acquisition system (VTAS) HMD.

Figure 3-9. Integrated Helmet and Display Sight System (IHADSS).

Figure 3-10. Joint Helmet-Mounted Cueing System (Vision Systems International).

The U.S. military Services began working on a class of VCS called multi-mode HMDs in the mid 1980's. A multi-mode HMD, in a single integrated system, functionally provides an image-intensified view of the wearer's environment (similar to NVGs), as well as a day/night display of spatially-referenced imagery (e.g., low-light level TV, forward-looking infrared [FLIR]) and symbology like a traditional VCS. This is illustrated in the example shown in Figure 3-11. The U.S. Navy first implemented a developmental model based on IHADSS, and numerous R&D efforts including U.S. Army Comanche HIDSS program and U.S. Navy Advanced Helmet Vision System program, which pursued both discrete optics and visor-projected versions of this class of system. These types of head-coupled systems not only functionally perform the night NVG and day/night HMD mission, but they also provide "sensor fusion" capability by simultaneously presenting correlated, spatially-referenced information to the user in the visible and near/far infrared regions of the electromagnetic spectrum. Recent developmental multi-mode HMDs, e.g., the Comanche HMD and Advanced Helmet Vision System programs, current HMD efforts for the Joint Strike Fighter (JSF) for the U.S. Navy and U.S. Air Force, and the AH-1 upgrades for the U.S. Marine Corps, are binocular/biocular, helmet-mounted vision systems.

Outside the United States, the first "modern" helmet-mounted sight (HMS) was the optically-sensed Russian design, developed to support the Vympel R-73/AA-11 Archer high off-boresight seeker, air-to-air missile, carried by the MiG-29 Fulcrum and the Su-27 Flanker, and built to attach to the ZSh-5 series Russian helmet (Beal and Sweetman, 1997). Even though this HMS (Arsenal's Zh-3YM-1) was relatively rudimentary, lacking missile-cueing symbols and using only a flip-down monocle with a light-emitting-diode (LED) reticle for aiming, the combination of the HMS and R-73 missile provided the Soviets with a greatly improved close combat capability (Merryman, 1994). The Arsenal Design Bureau (Kiev, Ukraine) subsequently improved on this first HMS with newer versions, like the Sura and Taurus. The combination MiG-29/ AA-11 were sold to the air forces in India, Iraq, North Korea, Libya, Syria, Iran, Yugoslavia and potentially Cuba (Lucas, 1994).

During the Cold War the Russians developed and deployed force-multiplier HMD and HMS systems that gave them an edge on air superiority and then sold these systems to (then) unfriendly nations. The combination of an

HMD-guided, 4th generation (GEN-4) missile and even inferior aircraft reduced to zero the technology advantage enjoyed by U.S. fighter aircraft. This caused a surge in HMD development programs in the Western countries.

The Israeli Display and Sight Helmet (DASH) 3/ Python 4 combination (1990s) had an equally important impact on HMD development. The Python-4 was a missile system that had limited "fire-and-forget" capability, as well as helmet-sight guidance. The DASH HMS system by Elbit Systems was developed for Israeli F-15s and F-16s and will be discussed in more detail in the *Current and Future HMD Programs* section later in this chapter, as it is considered to have played an important role in the development history of today's HMD.

Advantages of Helmet-Mounted Displays

There is little argument that displays and their ability to provide information are a distinct advantage in any operational setting. It would be unthinkable to offer an automobile design that failed to provide the driver with displays that provide real-time presentations of such operational parameters as speed and fuel status. While such information is not critical to the second-to-second operation of the automobile, drivers depend on being able to "look down" at the display console and obtain this information as needed.

However, there are operational settings where certain displayed information is critical on a second-to-second basis. For example, in fast-moving aircraft flying close to the ground, the operational environment changes so rapidly that even the brief time it takes a pilot to glance down at one or more displays to obtain aircraft flight status information may severely degrade his/her situation awareness. This short-coming of "head-down" displays gave rise to the development of head-up displays (HUDs) (Figure 3-12). HUDs employ fixed, transparent pieces of glass or plastic mounted inside the aircraft windscreen (e.g., combiners or beamsplitters). HUDs allow critical flight data to be accessed in a head-up, eyes-out scenario. This offers a tremendous advantage in applications where the time taken to view head-down displays can negatively impact safety and performance. The use of HUDs is not limited to aircraft. They have been employed in racecars, another application where outside operational conditions change so rapidly that a constant eyes-out requirement exists (Qt Auto News, 2006).

HUDs also are finding applications in less demanding vehicles. In an attempt to reduce accidents by preventing extended attention to head-down radio and CD-player knobs and buttons, a number of car manufacturers offer a windshield HUD. General Motors offers a HUD option on its Cadillac XLR/SRS models. The HUD presents a speedometer, turn signal indicators, audio system data, gear indication and cruise control settings (Dupont Corp, 2004).

But, as advantageous as HUDs are, they are fixed forward and are not as useful when the user is required to exercise constant head movement, e.g., constantly searching for enemy aircraft in a 360° environment. This factor played an important role in the motivation to mount the display on the head (or other head-mounted platform such as a helmet).

The potential benefits of HMDs have captivated the aircraft community for 40 years. The HMD concept can be extended and transferred to other areas where a wide field-of-regard is beneficial. While early HMD development was aviation driven, their utility beyond aviation has not been overlooked. Tank commanders can benefit by staying in touch with the "outside world" while remaining protected. Dismounted soldiers (classic infantry) can maintain constant situation awareness of the digital battlefield as well as expanded and enhanced sensory inputs via HMDs.

Nevertheless, the basic virtue of HMDs is to provide the ability to "look and shoot" at a target as fast as possible after target identification is completed. A dog fight usually lasts 30 to 60 seconds – the few seconds saved by eliminating aircraft pointing gives the pilot a vital advantage. Using the HMD, the pilot can quickly "tag" the enemy aircraft, launch a missile, and then turn to the next target and repeat the procedure. Sequential targeting enables a pilot to deal with multiple threats simultaneously, by eliminating the limitation posed by aircraft maneuverability.

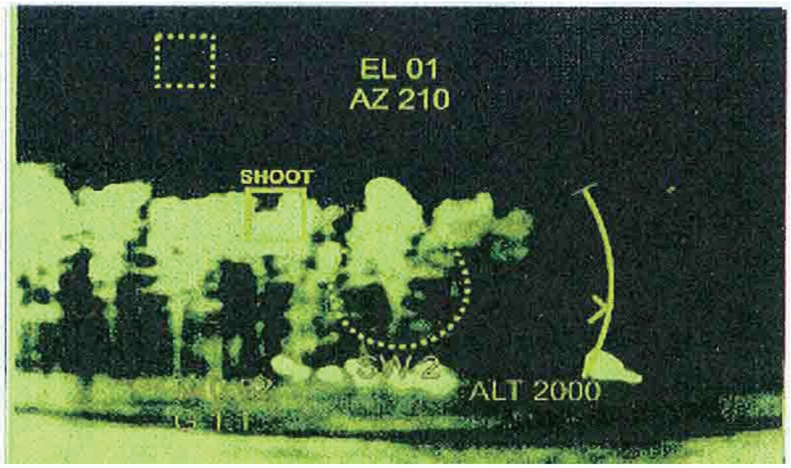


Figure 3-11. Example of multi-mode imagery with dynamic symbology.

Figure 3-12. Example of head-up display (HUD) in F/A-18C (National Aeronautics and Space Administration).

The dramatic threat coverage improvement provided by the wide field-of-regard of HMDs is shown in Figure 3-13. Comparisons are shown for a HUD, typical forward-looking radar, and off-boresight missile system.

The process of actively “tagging” targets is not limited to the individual platform: the pilot can identify a target and pass the information to an air surveillance and control platform (e.g., the Airborne Warning and Control System [AWACS] and Joint Surveillance Target Attack Radar System [JSTARS]), to other own sensors, or to another aircraft. Similarly, the opposite is useful as well - a detected threat by another platform or aircraft can be used to add cueing information to the HMD (Chapman and Clarkson, 1992).

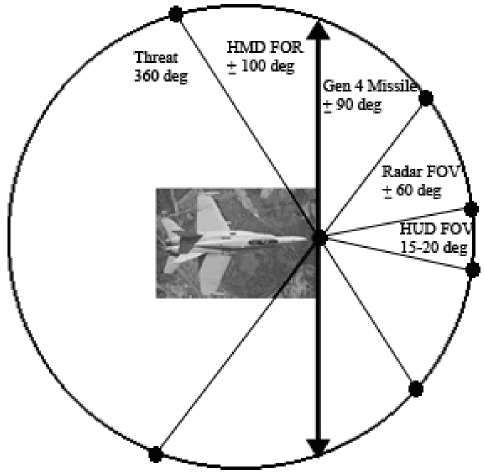


Figure 3-13. Depiction of expanded HMD/ HMS threat coverage.

For this reason, HMDs have increasingly been replacing and augmenting standard console-mounted head-down and traditional HUDs in advanced crew station designs. HMDs offer potentially greater direct access to critical visual information, while offering greater flexibility of head movement, less total system (but not user head-borne) weight, and greater flexibility in use of vehicular interior space, although at the cost of greater system complexity and possible expertise degradation in the case of system malfunctions.

More importantly, it is argued that HMDs provide users with increased situation awareness. Situation awareness encompasses the total information available, used to create an accurate picture of a battle theater, including spatial position and orientation of the aircraft, the surrounding areas, and any aircraft-relevant information. The pilot has to be aware of many different forms of information which is used to make judgments on how to respond to a given situation; any subtle level of perceptual cognizance to one's immediate environment can be vital for success in most situations (McCann and Foyle, 1995). The following operational definition of situation awareness has been proposed by a U.S. Air Force Staff Group: “A pilot’s (or aircrew’s) continuous perception of self and aircraft in relation to the dynamic of flights, threats, and mission, and the capability to forecast, then execute tasks based on the perception” (Geiselman, 1994).

In general situation awareness can be classified into Global (the "far domain") and Tactical (the "near domain"), covering close combat and navigational areas (Lucas, 1994). *Global situation awareness* refers to the range between 50 and 200 miles from the aircraft and related information is available from the main display on the instrument panel; whereas, *Tactical situation awareness* is the close range area within 50 miles, with information in the forward visual path. Each of these has associated temporal drivers as well, with faster reactions required the closer the relevant stimulus. This makes it physically impossible to see both domains simultaneously. As a result, pilots adopt a sequential acquisition scanning strategy by transitioning back and forth from the head-down instrument display to outside viewing, sampling information from first one domain, then the other. This recurrently interrupts the process of information acquisition and requires time-consuming actions, such as eye and head movements, eye accommodation, and becoming reacquainted with the alternating domains. Furthermore, as long as the pilot is looking at one domain, a sudden event (or sudden state change) in the other domain may be undetected.

By centralizing critical flight information within a user’s line-of-sight, overall performance is increased and operational safety is enhanced. HMDs offer users the advantage of monitoring critical information without having to repeatedly look down to scan instrument displays. Another proven benefit is that, with the ability to keep their eyes fixed to the outside world, users are more likely to detect important changes within the FOV (Harris and Muir, 2005: Manning and Rash, 2007). A specific example of the utility of this advantage is the greater probability in identifying runway incursions in military, civil and commercial aviation due to increased ability to maintain eyes out of the cockpit. Figure 3-14 depicts a typical HMD image. Note: This centralizing of critical flight information on front of the user's eye(s) should not be confused with the placement of the information (symbols) themselves, as early development of HMDs showed that symbology is most effective when placed around the periphery of the HMD imagery.

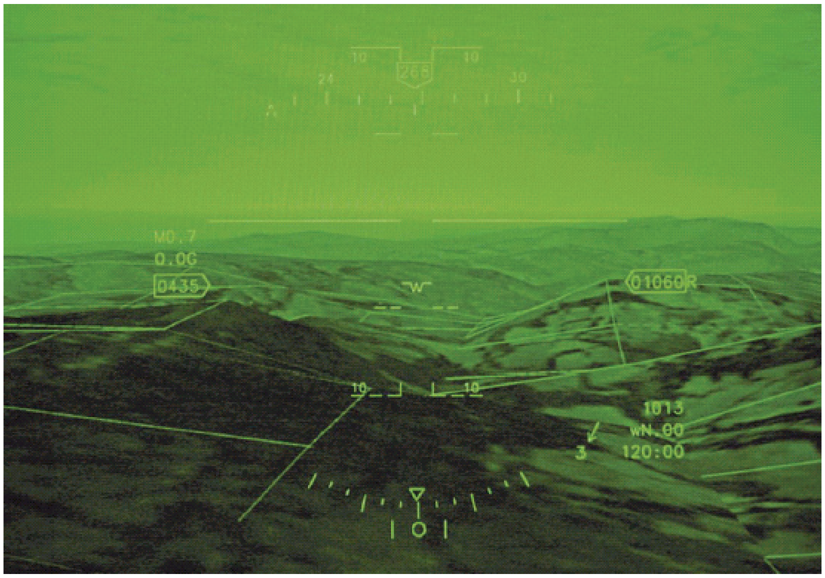


Figure 3-14. HMD Display (BAE Systems).

Limitations and Disadvantages of Helmet-Mounted Displays

Unfortunately, HMDs are not without their limitations and disadvantages. Some of the disadvantages are common to their predecessor, the HUD. First is the phenomenon of "attention capture" – or tunneling – which is the unwanted tendency for pilots to pay too much attention to the HUD and not enough attention to events in their field of vision outside the airplane (Foyle et al., 1993; McCann et al., 1993; McCann and Foyle, 1995). Attention capture with HUDs mounted just inside a windshield has been blamed for undetected runway incursions – one of the types of events that HUDs are to prevent. Numerous studies have attempted to understand attention capture and how it can be mitigated. Most disturbing is a developing consensus that HUDs (and hence HMDs) limit a pilot's ability to simultaneously process information derived from HUDs and from the real world (McCann et al., 1993).

Many HUD and HMD symbols are not "conformal" – that is, they are not overlaid in a one-to-one relationship to match shapes and features in the real world. Therefore, the symbols are perceived as different from the scene outside an aircraft's windows. This causes pilots to deliberately shift their attention to view either the symbols or the outside scene. The transition to conformal symbology may mitigate the attention capture problem (Wickens and Long, 1994). This conformity must be required for video imagery presented in HMDs. In other words, information is generated and presented based on conventions that users have to learn (train) to recognize: cognition processes as intuitive as they may be, are always slower than the instincts.

A second disadvantage is the possibility that HUD symbols or other imagery could obscure critical objects in the outside scene (Foyle et al., 1993). This problem can be reduced by keeping the number of symbols presented to a minimum and within the recommended size. Reducing the clutter caused by too many symbols also can decrease the potential for attention capture.

In addition to these general HUD-related disadvantages, other concerns are unique to HMD, as well as unique to the concept of mounting the display to the head. The first of these is user acceptability, which is important when any new technology is introduced; without user acceptance, the technology will not be used. The primary factors affecting acceptance are the head-supported weight, center-of-mass offset, required modification in head movement, display image quality/legibility, and display jitter and lag.

Most non-military pilots are not accustomed to wearing more than a headset on their heads. Current civil and commercial aviation headsets are generally lightweight, typically 12 to 18 ounces (340 to 510 grams) (Rash, 2006a). HMDs can increase head-supported weight by at least 16 ounces (454 grams). Military pilots wear helmet-based HMDs that weigh in excess of 4 pounds (lbs) (1.8 kilograms [kg]).

Because the HMD's display optics must be placed around the helmet with at least the combining element/visor in front of the eye, the HMD's additional weight is likely to be above and forward of the human head's natural center of mass - a factor that, as a flight progresses, may result in muscle fatigue.

For HMDs to present sensor and synthetic imagery that represent what a user is seeing, the HMD must incorporate head-tracking. The need for head-tracking increases the cost and the complexity of HMDs.

The head-tracking process of determining the user's head position, relaying this position to the sensor, the sensor's movement to the correct line-of-sight, the sensor's acquisition of the scene, and transmitting and presenting the final imagery on the HMD takes time (Rash, 2000). This time is called system latency. Latency times are typically hundreds of milliseconds (ms). The largest contributor is the "slew rate" of the sensor, or the time for the sensor to move to the line-of-sight defined by the new head position. Studies have shown that total system-latency times approaching one-third of a second or longer (~300 ms) are unacceptable from a performance standpoint. Many in the VCS community today are trying to achieve a total system latency time of less than one display frame time (typically 33 ms).

These latency times have been blamed for motion sickness. The onset and severity of motion sickness symptoms are difficult to predict, and such occurrences in commercial aviation would be unacceptable. Studies by the U.S. National Aeronautics and Space Administration (NASA) have documented the need for improvement in image alignment, accuracy and boresighting of HMDs to help mitigate this problem (Bailey et al., 2007).

Helmet-Mounted Display Applications

There is general agreement that HMDs have great potential applications; why, then, have only a few systems (mostly military) been fielded? Many factors contribute to this situation: cost, lagging technology, less than optimal ergonomics design (Keller and Colucci, 1998), unfinished search for that "application" that will excite users, unawareness of the potential benefits, and simply the "visceral dislike" (Hopper, 2000) of wearing a monitor on ones head. Four decades into the HMD exploration, the "killer application" that will propel the technology has not yet been identified.

Ivan Sutherland (1965) proposed the "Ultimate Display", more than 40 years ago (Figure 3-15). While at the Department of Computer Science, University of Utah, Sutherland imagined a display in which all-powerful computers would generate graphics of objects that would behave exactly (in all sensory modes) as their real-world counterparts. Implied in his concept were certain characteristics and expectations: a) the need for a complete sensory response: sight, sound, smell, feeling (haptic), and kinetic feedback to create the new reality and b) the use of HMDs will serve as a step toward an intuitive interface between human and machine, a natural way to add 3-D to an otherwise flat computer imagery. This display is still far into the future, but the anticipated technologies have come to fruition as we have moved into the 21^{st} century. Others still are found only in science fiction. Nonetheless, Sutherland's HMD concept opened the way to computer-generated 3-D stroke images coupled with head trackers – the same basic principles applied today.

Figure 3-15. Ivan Sutherland's HMD (late 1960’s) (Department of Computer Science, University of Utah).

The military has led in the applications of HMDs, and there is a growing interest in industrial and consumer applications. Some current and future potential applications are listed below. It must be noted that there are no rigid boundaries between these applications, as some applications have multiple usage across these boundaries. The use of HMDs in simulation and training has been adopted by both military and industrial users, and has served as a precursor to consumer gaming.

Military applications include:

- Navigation and situation awareness
- Targeting
- Night vision systems
- Visual enhancement
- Security monitoring
- Simulation and training
- Maintenance and inspection
- Remotely-piloted vehicle interface

Commercial applications include:

- Computer-aided design/ Computer-aided engineering (CAD/CAE)
- Surgical aid - microsurgery, endoscopic surgery
- Emergency medical telepresence
- Security monitoring
- Maintenance, Repair and Overhaul (MRO)

Consumer applications include:

- Gaming
- Mobile Internet access
- Private DVD viewing
- Fire-fighting

The following sections briefly describe and discuss some of the more important and interesting applications within the three areas: military, commercial and consumer.

Military applications

Military applications are the focus of this book – the merits of HMDs for both fixed- and rotary-wing aircraft are beyond questioning, and HMDs already have become an integral part of the next-generation cockpits. Much of this success is due to the use of head/helmet-tracking to produce visually-coupled HMD systems.

Use of visually-coupled systems (VCS) for pilotage, navigation and/or situation awareness

VCS technologies have been used for a tremendous variety of mission applications over the years. As previously noted for early applications of helmet-mounted sights, head-position sensing was used for a variety of line-of-sight designation and targeting in conjunction with onboard weapons and sensors. Some of the earliest investigations of HMD technologies were designed as a way to investigate a wider FOV display in cockpits or crew stations of various air, ground, and maritime vehicles.

Over the years, the military has interfaced helmet-mounted sights and HMDs to a wide variety of vehicle systems and weapons. They have been linked with radars, electro-optical/TV missile systems, reconnaissance sensors, long-range target identification sensors, pilotage sensors, head-slaved guns (both air-to-ground and surface-to-air), and angle-rate bombing sensors. They have been interfaced with distributed aperture sensor systems for a total coverage “windowless cockpit” synthetic vision system capability for both aircraft and ground vehicles. They have been used to present spatially-referenced “highway-in-the-sky” type flight control information for both fixed-wing ejection seat aircraft and rotary-wing operations and for shipboard landings, and to present “predictor” fire control dynamic symbology such as “hotline gun sight.” These are fairly typical VCS applications.

There have also been some “non-traditional” VCS applications attempted by the military over the years. One example is the use of a head tracker and HMD as an effective operator interface with a remotely piloted vehicle. By using VCS, the “illusion” can be created for the operator that they are “out there onboard the vehicle.” The military has successfully interfaced VCS with airborne, ground-based, and undersea unmanned vehicles for a wide variety of missions including reconnaissance, targeting, bomb disposal, undersea operations and other teleoperator applications.

Virtual cockpit

The “Virtual Cockpit” is a second application that has moved forward in the military with the main goal of providing a “software reconfigurable cockpit.” In the late 1990s the U.S. Army’s Program Manager-Aircrew Integrated Systems (PM-ACIS), Huntsville, Alabama, initiated the Virtual Cockpit Optimization Program (VCOP) to integrate advanced technologies into a single system. VCOP technologies included a Retinal Scanning Display (RSD); fully integrated 3-D cockpit audio technologies with speech recognition and synthesis; an Integrated Caution, Warning and Advisory Annunciator (ICWAA); and an Electronic Data Manager (EDM); all integrated and managed by the Rotorcraft Pilot’s Associate (RPA) Software. These technologies were intended to enhance situation and threat awareness, while at the same time providing a cost-effective technique to modernize legacy aircraft. In its simplest configuration, VCOP goals were to:

- Provide efficient access to critical information with minimized “head-down” time;
- Formulate “standardized” dashboard panel requirements; and

- Establish an environment for rapid avionics prototyping, integration, test and evaluation of multiple aircraft configurations.

A similar program was initiated in Japan, in the early 2000's, by a team coordinated by Kawasaki Heavy Industry and Yokogawa Electric Corporation (Bayer, 2007). Similar to U.S. Army's VCOP, this program's goals were to:

- Minimize cockpit cost and weight;
- Develop reconfigurable configuration between manned- and unmanned combat aircraft; and
- Increase pilot's situation awareness.

Virtual Reality (VR)

We have seen that HMDs can be designed to be see-through (transparent), in which case the sensor- or computer-generated (synthetic) imagery is overlaid on the actual physical world outside, or nonsee-through (occluded), where the user only sees sensor- or computer-generated imagery. In the former case, the HMD is said to create an Augmented Reality (AR), i.e., adding information to the world around the user. In the latter case, specifically when the HMD presents only computer-generated imagery, the situation is referred to as Virtual Reality (VR); the real world is completely obscured, with computer-generated imagery being the only visual information the user receives.

AR and VR are related, and it is valid to consider the two concepts together in terms of a continuum linking purely virtual environments (VEs) to purely real environments. The VR environment is one in which the participant/observer is totally immersed in a completely synthetic world, which may or may not obey the properties of a real-world environment. Indeed, it is possible in VR to exceed the bounds of physical reality by creating a world in which the physical laws governing gravity, time and material properties no longer hold. In contrast, the strictly real-world environment clearly is constrained by the laws of physics.

Rather than regarding the two concepts simply as antitheses, however, it is more convenient to view them as lying at opposite ends of a continuum, which is referred to as the *Reality-Virtuality (RV) continuum*. This concept is illustrated in Figure 3-16 (Milgram, 1994).

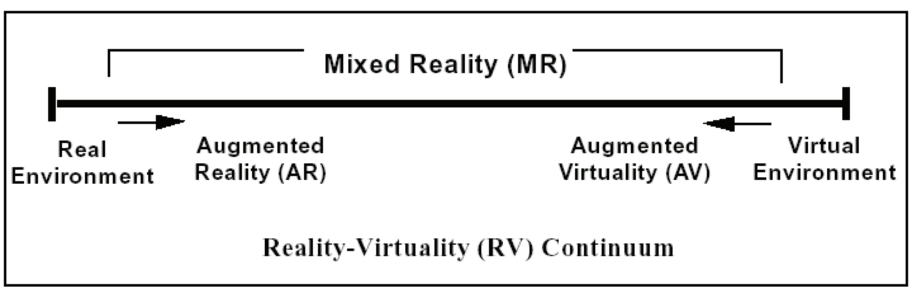


Figure 3-16. Reality-Virtuality Continuum (Milgram, 1994).

The *Real Environment* (RE) (extreme left) consists solely of real objects and is observed when viewing a real-world scene either directly, or through a 100% transparent window. The *Virtual Environment* (VE) (extreme right), defines environments consisting solely of virtual objects, e.g., computer graphic simulations; RE is completely suppressed here. The Mixed Reality (MR) environment is one where real and virtual world objects coexist and are presented together. The HMD is the mechanism that brings the MR to existence. Its level of transparency to the real world positions the "instantaneous" reality on the MR continuum line, depending on whether the HMD is a "see-through" or "opaque" configuration.

Whether the environment is Augmented Reality or Augmented Virtuality, depends of whether the presented environment is primarily real, with added computer generated graphics, or is primarily virtual, but augmented through the use of real (i.e. un-modeled) imaging data (Drascic and Milgram, 1996).

In summary, AR systems bring the computer to the user's real environment, whereas VR systems bring the world into the user's simulated computer-generated environment. This paradigm for user interaction and information visualization constitutes the core of a very promising new technology for many applications. However, real applications impose strong demands on AR technology that cannot yet be completely met at the current level of technology.

Simulation, training and mission rehearsal

Next to aviation applications, simulation, training and mission rehearsal are probably the best known HMD-based VR applications for military purposes (Haar, 2005). The military and NASA have had substantial R&D efforts aimed at using VCS as an alternative to large domed simulators. By doing this, resolution and graphics power can be concentrated into the instantaneous FOV of the subject, providing a higher performance system. Special techniques such as foveal/peripheral image generation and eye position sensing (eye tracking) have enhanced the operator interface in some of these systems. By creating a virtual world and a virtual cockpit, changes in crew station design can be investigated in this “virtual world” before real-world hardware is redesigned and modified.

Combat simulators are well established and offer an excellent fit with HMD-based applications. In conjunction with powerful computer systems, they can simulate and integrate entire environments within a single display. The fundamental difference between simulation and training is that the former often is used as a tool for development, evaluation and validation of new designs or to visualize results of complex computations that result in large 3-D graphics (Casey, 1991). Training is presenting the same sets of video scenarios with already known solutions to multiple users and interactively evaluates their response time and degree of accuracy of the solutions offered.

Simulation techniques and applications have greatly expanded with the apparently never-ending increase in computer processing power - from flight training into war simulation with a complete air fleet. Display performance requirements for such application are among the most demanding of all. For best results, simulation fidelity must match physical reality that will be encountered in the field. HMD-based simulation arguably is the best way to perform realistic simulation.

Flight training

The Aviation Combined Arms Tactical Trainer – Aviation Reconfigurable Manned Simulator (AVCATT-A) (Figure 3-17) is an aviation training simulator for both active U.S. Army and National Guard units. It is a dynamic reconfigurable system used for combined arms collective training and mission rehearsal through networked simulators in a simulated battlefield environment. AVCATT-A provides five functional cockpits: the OH-58D Kiowa Warrior, the AH-64A Apache, the AH-64D Apache Longbow, the CH-47D Chinook, and the UH-60A/L Blackhawk helicopters.

The AVCATT-A is purely a helicopter *combat* trainer and not a *flight* trainer. There is no extent of motion, and it does not give the trainees a sense of flying the helicopter. Only instruments that are specific for combat operations are usable. Its greatest asset is that it provides a unique capability to allow units to train as units and not as individual aircrews. The AVCATT-A provides the capability to conduct realistic, high intensity, task loaded collective and combined arms training exercises and mission rehearsals of current Army attack, reconnaissance, cargo, and utility aircraft.

The physical layout of an AVCATT-A suite consists of two trailers connected by a platform. One trailer includes three reconfigurable manned modules and a 20-person After-Action Review facility. The second trailer includes three reconfigurable manned modules, a Battlemaster Control room, and a maintenance room.

AVCATT-A provides a total capability of six manned module cockpits per suite, networked together to help train an aviation company or air cavalry troop. Each manned module is reconfigurable to current Army attack, reconnaissance, cargo, and utility aircraft. AVCATT-A has the capability to be linked via local area network (LAN) and/or wide area network (WAN) with other AVCATT-A suites, and other combined arms tactical trainers such as the Close Combat Tactical Trainer (CCTT). This provides the capability to conduct collective training from team through combined arms levels (Simons et al., 2002). The AVCATT-A visual system (Figure 3-18) creates the Out-the-Window (OTW) and sensor imagery view.

Figure 3-17. Pilot in the AVCATT-A System.

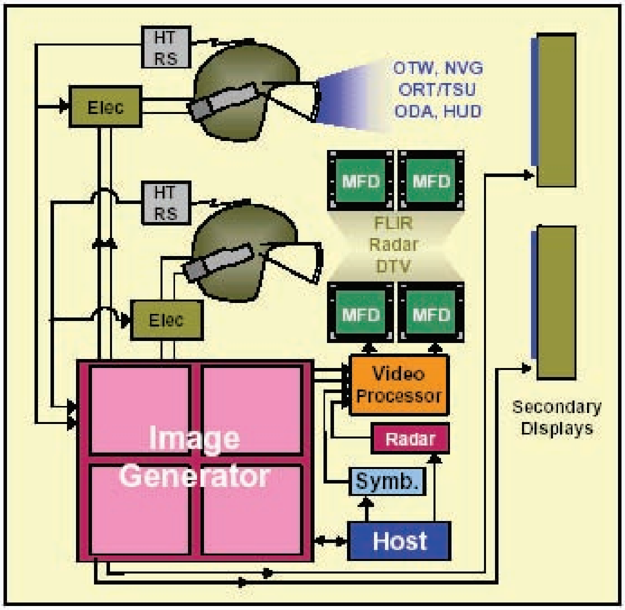


Figure 3-18. AVCATT-A visual system.

The major components of the AVCATT-A visual system are the Image Generator (IG), two HMDs, two Multifunction Displays (MFDs), and two secondary (backup) displays. The IG provides the imagery for the pilot and copilot, as well as two sensor channels. The HMD (a Rockwell Collins Model SimEye™ XL100A) is a high-resolution, full color head-mounted display that traces its origins to the Wide-Eye™ HMD designed for the U.S. Army's Light Helicopter Experimental (LHX) program (1980s), which was the predecessor to U.S. Army's Comanche program of the 1990s.

Driver trainer with mission rehearsal

Some U.S. Army vehicles now have embedded training, such as the simulator built into a Bradley Fighting Vehicle, based on the BAE Systems (U.K.) Bradley A3 Embedded Tactical Training Initiative (BETTI). This system enables soldiers to train with realistic-looking simulated terrain while they are sitting in their vehicles in the belly of a C-130 cargo plane, in route to the area of operation. When the aircraft's ramp drops to the runway, Warfighters drive directly from the virtual world into the real one. However, for Mission Rehearsal Exercises (MREs) to work, the simulations must have enough fidelity to earn troops' confidence that they will be able to draw on their simulated lessons in the heat of battle. The key challenge is to achieve a "real immersion," to faithfully replicate scenarios, and to represent the physical world in the display environment in a believable manner.

Commercial applications

The basic concept of an HMD as a head-up mode for information presentation has been of interest to various sectors of the commercial and industrial communities. However, in spite of less demanding environments, non-military applications must face a number of unique hurdles that include:

- What are the benefits an HMD-based system brings to the application (e.g., easy access to information, privacy, stereo imagery, wide field of regard)?
- What are the logistical, human factors, and ethical issues associated with the choice of an HMD over that of current direct view displays, e.g., privacy, transportability, storability?
- Is the technology mature enough to perform acceptably in the application?
- Do the cost and added inconvenience justify an HMD approach?

In general, once the cost/benefits issues have been evaluated and found acceptable, one of the remaining chief barriers to commercial applications of HMDs is user acceptance. Most commercial and industrial workers (construction workers being an exception) are not used to having to wear any type of head-gear. Head-supported weight, center-of-mass offsets, pressure points, sweating, and overall discomfort are common complaints of such devices, and such issues have certainly had a negative impact on user acceptance and, hence, the implementation of HMDs. Developers, aware of these problems, have pursued such solutions as designs no more cumbersome than simple eyeglasses. However, eye-wear HMDs come with their own set of limitations, with a narrow FOV (usually less than 20°) being probably the most critical.

Nonetheless, a number of commercial applications do exist. As such issues as head-supported weight and overall discomfort are addressed by low-weight designs, the advantages of HMDs will eventually increase this number. Potential application areas will be those where users can benefit from visualized information otherwise not available or difficult to obtain due to certain task constraints.

In the following sections, a few commercial applications are briefly described. While as in military applications, the aviation-related ones are predominate, many medical applications presenting diagnostic and surgical imagery are emerging, as HMDs offer an alternative method of presentation of this imagery.

Instrument landing

The National Research Council's (NRC's) (Canada) Cockpit Technologies Program has flight tested a stereoscopic 3-D display format to determine the feasibility of using HMD-presented pictorial and stereoscopic cues during helicopter Instrument Approach Procedures (IAP) (Jennings, 1997). Pilots were able to complete approaches to safe landings and reported that the pictorial format improved their situation awareness during the approaches. While lacking stereo cues, the pictorial display contained several strong monocular depth cues such as occlusion, linear perspective, and visual field-flow (motion). This type of system would be extremely useful during Instrument Meteorological Conditions (IMC), when the outside world is obscured, and pilots can no longer use external visual cues for maintaining control of the aircraft.

Training

The potential of HMD-based VEs for training simulation has been recognized right from the emergence of this technology. The Federal Aviation Administration (FAA) is pursuing research focused on the aircraft inspection processes. Existing training for inspectors in the aircraft maintenance environment tends to be mostly on-the-job training;, however, feedback to the trainee, may be infrequent, unmethodical, and/or delayed. One of the most viable approaches in the aircraft maintenance environment, given its many constraints and requirements, is computer-based training which is efficient, facilitates standardization and supports distance learning.

A recent example is the Automated System of Self Instruction for Specialized Training (ASSIST), featuring a personal computer (PC)-based aircraft inspection simulator. Despite the advantages, the simulator is limited by its lack of realism, as it uses 2-D sectional images of airframe structures. More importantly, the inspectors are not immersed in the environment, and, hence, they do not get the same look and feel as when conducting an actual inspection. To address these limitations, a VR-based inspection simulator using an HMD has been developed (Duchowski, 2000).

Analysis of performance data with this environment (Vora, 2002) revealed a significantly greater number of defects identified within a significantly shorter visual search time in the VE in comparison with the ASSIST environment. When these results were coupled with subjects' perception of the two systems, the VE system was preferred to the ASSIST as an aircraft inspection training tool by a ratio of almost 3:1, proving the potential effectiveness of an HMD-presented VE in improving both speed and accuracy of visual search.

Surgical planning and diagnostic tasks

A see-through HMD has been used by surgeons to view preoperatively scanned images (e.g., ultrasound, x-ray, Magnetic Resonance Imaging [MRI]), as if looking through the patient at the internal organs (Bajura, 1992). Key to the implementation, of course, is accurate color rendition and accurate registration of the 3-D graphics to the real world.

Surgery

Great advances have been made in reducing the invasiveness of surgical procedures. Many surgeries today are performed through either natural body openings or through small incisions, with the surgeon viewing the surgical field indirectly via a remotely operated camera which has been inserted into the operative field. Today, surgeons routinely remove appendixes, gallbladders, spleens and other organs and tissues by laparoscopy. The most qualified are now macerating and removing kidneys, pancreases, colons, adrenal glands and other more complicated organs, or repairing them without open surgery. In the vast majority of cases, the surgeon views the imagery on monitors located at some distance away. HMDs can allow the surgeon increased eye-hand coordination, situation awareness and flexibility as compared to viewing remotely positioned monitors, especially when coupled to teleoperated and robotically assistive instruments. In demonstrations of this application, computer generated graphics (i.e., AR) have been integrated into the HMD imagery (Ackerman, 2002).

Molecular studies

At University of North Carolina (UNC) at Chapel Hill, three major fields of research: interactive molecular studies, medical imaging and virtual building exploration are making use of the advantages of HMDs (Chung, 1989).

Macromolecules have complex 3-D structures and understanding them is often the key to explaining material chemical properties. Researchers at UNC envision a system where chemists use an HMD to view a room-sized, 3-D virtual molecule to study its external structure by “walking” around its exterior, to “enter” the molecule to examine the internal connections, and perhaps (in the future) to cause the molecule to respond to changes in ambient conditions.

Virtual Reality Dynamic Anatomy (VRDA)

A cooperative effort of Optical Diagnostics and Application Laboratory (ODALab) (Orlando, FL), 3-D Visualization (3DVIS) Laboratory (Tucson, AZ), and Media Interface and Network Design (MIND) Laboratory (Tucson, AZ) has investigated a couple of interesting WR applications. One of these is the VRDA concept, which is a visualization tool for teaching complex anatomical joint motions (Rolland, 2002). The VRDA allows a trainee to manipulate an anatomical joint and visualize the virtual model of the inner anatomy superimposed on the body using marker based techniques. Coupled with tactile phantoms, this can become a very immersive experience.

Airway management visualization and training

To open blocked airways, it is sometimes necessary to perform an endotracheal intubation (ETI) which consists of inserting a tube through the mouth into the trachea and then sealing the trachea so that all air passes through the tube. In an effort to improve training and keep them current, the U.S. Army Simulation, Training and Instrumentation Command (STRICOM) (Orlando, FL), and Medical Education Technologies, Inc. (METI) (Sarasota, FL), who provided the human patient simulator, teamed with ODALab to develop the Airway Management Visualization and Training for paramedics (Davis, 2002). This is an HMD-based AR system that allows paramedics to practice their skills and provides real-time feedback of their performance and suggests improvements/corrections.

Telepresence

Conventional telepresence usually is implemented through a pan and tilt camera system controlled by a joystick. This requires significant operator training and can be expected to lead to longer task execution. This is due to the constant requirement for the operator to adapt to the frame of reference from the camera. Nevertheless, studies have shown that telepresence, when accompanied by stereoscopic displays, brings definite benefits to the person operating remote equipment (Reinhart, 1991). Some applications include telerobotic fields, e.g., remote mining, nuclear sites inspection, space exploration, mine clearing equipment, instances where it is impractical for the human operator to be at the immediate location, whether for safety or other reasons.

A more advanced telepresence, currently in development, proof-of-concept stage is anthropometric telepresence (Primeau, 2000). Anthropometric telepresence is the next best thing to actually “being there,” with the added benefit of safe operation away from areas deemed too hazardous to have an operator on site. It is based on a camera system that is slaved in real-time to the operator’s line-of-sight. The information relayed back is presented in a natural way, which makes most training unnecessary; the operator is fully immersed in the remote site. Applications exist in space, military (piloting unmanned air/land/sea vehicles), law enforcement, industry (mining, oil exploration) and many other situations where hazardous situations exist or may develop without warning (nuclear, biological, chemical).

Consumer applications

Burdened with the same problems associated with commercial applications (e.g., discomfort, lack of acceptance of head-supported weight), it is not surprising that consumer applications have lagged even further behind. It is one thing to be paid to wear an uncomfortable device; but, doing so and having to pay for it is something else again. An exception to this argument is full-immersion computer games. Besides the VR aspects of state-of-the-art games, wearing a near-true-to-life HMD while flying an F-18 Hornet adds to the realism and the thrill. The potential for gaming is absolutely limitless. In general, the 3-D interactive games mimic military flight missions, space war games and otherwise unobtainable adventures.

Gaming

Personal gaming applications using head-worn displays are extensive. At annual gaming industry expositions, sophisticated full-immersion games, virtually all requiring some type of HMD, are the center of attention. Figure 3-19 depicts one of the latest entries in the fast-moving industry. It is the Trimersion HMD manufactured by 3001 AD,[5] touted as the "next level of realism by offering greater immersion inside the game via an HMD [acting as] a realistic and natural interface" (Gizmag, 2006).

The design uses built-in headphones and a headband system. The headband is described as a mask that surrounds the display optics. The manufacturer contrasts this design to others that employ either visors that allow external light to come into the line-of-sight of the player or eyecups that are uncomfortable. The Trimersion HMD mask curves around the player's cheekbones using soft rubber, providing a complete lightless enclosure.

Figure 3-19. 3001 AD's Trimersion gaming HMD.

[5]3001 AD, 430 South Congress Ave, Delray Beach, FL 33445

Current and Future Helmet-Mounted Display Programs

In this section, brief synopses of the more significant HMD programs will be presented. While most of these will be programs that achieved at least limited fielding, some are still in their research and development phase, and others were never fielded due to a variety of reasons but still represent significant advances in HMD design. The majority of these programs is military-related and represents worldwide efforts. However, a few commercial systems also are presented.

While most military programs were for rotary- and fixed-wing aircraft platforms, more recent programs have developed HMDs for use by both vehicular-mounted and dismounted Warfighters. Since training applications have increased, simulation HMD programs are also included. In a few cases, an HMD system may have applications on more than one platform.

The salient programs are presented first for fixed-wing platforms, then for rotary-wing platforms, and finally for the mounted-vehicle, dismounted and simulation platforms.

Military HMD programs: Fixed-wing platforms

A main HMD application that drove early development is target cueing. Since the mounting of machine guns on airplanes in World War I marked the official beginning of the evolution of pilot-centered weapons, pilots invariably had cued on the targets by pointing the nose of the aircraft in the direction of the target.[6] Introduction of the HUD marked the first step toward allowing pilots to cue their weapons with an out-of-the-cockpit aiming device. A giant leap forward in terms of pilot-to-aircraft interface, the HUD displayed not only accurate weapons-aiming symbols, but also relevant flight data such as airspeed, altitude, and heading. For the first time, pilots could view such information without looking back inside the cockpit.

The dynamics of airborne combat require pilots to outmaneuver each other. Air Forces around the world have run a technological race aimed at gaining superiority through increased propulsion and maneuverability of fighter aircraft that continued with second and third generation heat-seeking missiles. Although visually-coupled systems (VCS), the concept of linking helmet sighting systems with radars and missiles, as an operational capability dates back to the early 1970s, advances in both helmet vision systems and high off-boresight missile seeker technology of the current day brings a much more significant tactical capability to the Services today. Capable Air Intercept radars had several dogfighting modes that were designed to rapidly acquire and track a target. When the first fourth generation missiles appeared, e.g., the Soviet Vympel R-73 (AA-11 Archer) and the Israeli Rafael Python 4 (Beal and Sweetman, 1994), it was clearly apparent that with very large off-boresight angles, typically of the order of 90 degrees of arc, the old flight dynamics would no longer be adequate. Subjected to high-G forces, pilots risked loss of consciousness and extended incapacitation. Performance limitation moved beyond hardware to the human operator.

The arrival of the HMD as a cueing tool changed, and is continuing to change, this scenario. Superior aircraft speed and maneuverability agility are no longer essential factors to a successful engagement. The use of HMDs allows slaved air-to-air missiles, capable of more than 50Gs, to execute the high-G turn instead of the pilot; the HMD is a true force multiplier. Less proficient pilots flying inferior aircraft armed with a GEN-4 missile enjoy a distinct advantage because of the HMD. Essentially, HMDs are “must have” equipment on GEN-4 fighter aircraft, since high off-boresight weapons and visual cueing outweigh any aircraft-performance advantage during a dogfight. Experts believe that HMD cueing systems significantly increase the win probability for the same aircraft armed with a GEN-4 high off-boresight missile

[6] Exceptions are the use of gun-turrets in multi-engine aircraft during WWII and side gunners in modern gunships and helicopters.

Cueing HMDs make it possible to synthesize the target information by using an HMD with a cockpit computer and onboard advanced weapons' capabilities. Position sensors on the pilot's helmet track the instantaneous pilot's line-of-sight as it follows the target. The sensors relay critical information to the computer, which in turn, communicates the location of the target to the missile system. When the weapons lock onto the target, the pilot receives both audio and video signals, and then pulls the trigger located on the control stick to fire the missile. The advantage of the few extra seconds gained by getting the missile launch first, could well make the difference between life and death.

The first high-off-boresight VCS test in the U.S. military took place in early 1994, at Tyndall Air Force Base, FL (Hughes, 1994). It was a conclusive demonstration of how a Honeywell HMD, a Raytheon missile, and a Lockheed F-16 could perform seamlessly as an integrated system and achieved 72° of off-boresight deflection with a 30G acceleration.

This scenario represents a total paradigm shift in the way air-to-air fighter combat is fought and brings back the advantage of independently swiveling gun turrets of older multi-engine aircraft. The sighting reference for cueing a weapon is no longer the nose of the aircraft but rather the pilot's HMD. As long as the target is within range and the pilot can view the target via the HMD, the relative position of the aircraft to the enemy is not critical. Tactical implications are profound and serve as the major driver for many if not all of the following HMD programs directed at fixed-wing platforms.

Table 3-1 presents a partial summary of the more notable experimental, prototype, fielded and future HMD fixed-wing programs. It followed by summaries of select HMD programs. Many of these HMDs are depicted in Figure 3-20. Many of the programs involved a number of contracts with various commercial HMD developers playing differing roles. Many of the programs also were multi-national in scope. The country of development listed in Table 3-1 and ensuing program descriptions generally is based on the initial developmental phase.

Display and Sight Helmet (DASH) series (Israel)

Elbit Systems Ltd. (Israel) developed a series of HMDs known as the (DASH) in the 1ate 1970’s (beginning with DASH 1) and was installed on the Israel Air Force F-15s and F-16s. Both air-to-air and air-to-ground configurations have been deployed. DASH 2 had an improved design, but was never produced in volume.

DASH 3 (Figure 3-20) entered production during the early 1990s in conjunction with the Rafael Python GEN-4 air-to-air missile. DASH 3 is currently deployed on IDF F-15C/D, the F-16C/D, the F-15I, the F/A-18C aircraft and has been offered to export customers, as part of upgrade packages for F-5E/F and also for Russian aircraft. Dash 3 has been implemented in the Romanian Mig-21 (Lancer) platform upgrade. This HMD deserves careful examination as it has been the first of the new generation of Western HMDs to achieve operational service and it also provides part of the technology base for the Joint Helmet Mounted Cueing System (JHMCS).

The DASH 3 is an “embedded” HMD design, where the complete optical and position sensing coil package is built into a standard helmet form factor, in this instance either the U.S. Air Force standard HGU-55/P or the Israeli standard HGU-22/P. The helmet is customized to individual pilot head shapes and sizes using either poured foam or Thermal Plastic Liners (TPL™). Once the helmet is fitted to the pilot, the optics is adjusted to position the HMD’s exit pupil to the pilot’s eye. DASH 3 accommodates eye glasses and standard oxygen masks. DASH 3 weighs 1.65 kg for the larger helmet size, and the helmet center of gravity is well balanced, meeting requirements.

A visor-projection optical configuration is used for this HMD. The projection on a spherical visor eliminates the risks and cost impact of an aspheric visor. Dash 3 provides a 20-degree FOV, with a 15-mm exist pupil for the optics. All symbology is calligraphic, produced by a programmable stroke generator.

The strength of the Dash 3 lies in its maturity and compact form factor, which is advantageous in a tight canopy (Koff, 1998). The system is operational in 5 countries, on 4 continents and onboard 5 different major platforms (F-15A/B/C/D; F-15I; F-16C/D; F-5E/F; MiG-21). Over 1000 Dash systems have been delivered to customers worldwide.

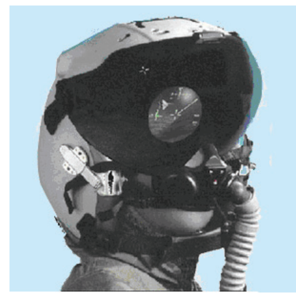
DASH 3

Agile Eye Plus (circa 1992)

Crusader

TopSight

Typhoon (Eurofighter) IDH



Joint Helmet Mounted Cueing System

HMDS

TopNight

Viper 3

Figure 3-20. Selected current and future fixed-wing HMD programs.

Agile Eye (United States)

Kaiser Electronics (now Rockwell Collins) has produced and tested a series of experimental systems since early 1980's including several Agile Eye and Agile Eye Mark I to IV systems. Agile Eye Plus (circa 1992) is shown Figure 3-20. The Agile Eye Mark V, the Visually Coupled Acquisition Targeting System (VCATS), produced in 1995, is very important to the HMD technology development.

Agile Eye uses a small CRT in the back of the helmet to project imagery (symbology and targeting data) to the pilot's eye via a set of relay optics and projection off the visor.

Table 3-1.
Summary of selected fixed-wing HMD programs

Time frame	Program	Country	Platform	Developer	Program status	Notes
1970s	Dash 1	Israel	Fixed-Wing F-15, F-16	Elbit Systems	Fielded	
Early to mid 1980s	Agile Eye	USA	Fixed-Wing Miscellaneous	Kaiser Electronics	Experimental	Mark 1 through 5 (VCATS)
Early 1990s	Dash 3	Israel	Fixed-Wing F-15, F-16	Elbit Systems	Fielded	
1990s	Viper 1-3	UK	Fixed-Wing Miscellaneous	GEC-Marconi Avionics Ltd (Delft Instruments)	Experimental	
Mid 1990s	Crusader	US/UK	Fixed- & Rotary-wing	Gentex/ BAE Systems/ Thales	Experimental	Technology demonstrator
Late 1990s	TopSight	France	Fixed-Wing Miscellaneous	Thales (Sextant Avionique)	Fielded	Day mission
Late 1990s	TopNight	France	Fixed-Wing Miscellaneous	Thales (Sextant Avionique)	Fielded	Night Mission
1999	JHMCS	USA	Fixed-Wing F-15, F-16, F-18	VSI	Fielded	
2008	Scorpion	USA	Fixed-Wing	Gentex	Operational Testing	
2008	Typhoon IHD	UK	Fixed-Wing Eurofighter	BAE Systems	Development	
2010	HMDS	USA	Fixed-Wing F-35	VSI	Development	

VCATS was extensively used as a design tool and test bed by the U.S. Air Force Research Laboratory at Wright Patterson Air Force Base, OH. The VCATS program was specifically designed to solve the technical and operational problems that historically had plagued HMDs, and it has paved the way to a successful JHMCS program. Some of the technology “building blocks” in VCATS were jointly supported by the Navy Science and Technology Base Program. Among the problems tackled on the VCATS was the introduction of a standardized helmet-vehicle interface (HVI) that uses interconnecting modules, which are easily replaced with minimal effort, down-time, or potential for error. Through the helmet and its connectors, the pilot becomes part of a closed-loop electronic system. The quick disconnect (QDC) connector also provides for emergency egress and allows “hot” disconnect without arcing.

VCATS also represents a prelude to a human-factors breakthrough. From the very beginning of air fight increased propulsion and maneuverability were the main two factors of improving the U.S. fighter pilot's advantage in the end game. The latest fighter aircraft speeds and agility levels place the pilot in the position of pulling dangerously high-force levels of up to 12Gs, maneuvers that can produce devastating results such as blackouts and extended incapacitation. With VCATS, however, the pilot continues to be limited to a safer 9Gs, while the missile may execute the high-G turn (in excess of 50Gs is now common) instead of the pilot, while in route to the target. VCATS introduced a human-centered system matching the pilot's physical and mental capabilities (the visual system, head-eye-hand coordination, decision-making abilities, and response time).

A summary of VCATS program findings and implementations is provided in Table 3-2.

Table 3-2.
VCATS findings and implementations.

Finding	Implementation
Eyeball critical sensor	No visor reflective patch
Keep system latency below the limit of being noticeable by pilot	Achieve 30-50 ms of latency; System integration
Interference suppression to smooth head bounce in high-G buffet	High update rate tracker; Accelerometers and digital filter algorithm for active noise cancellation
Keep static pointing errors < 5 mrad	Tracker algorithms; System integration
Use custom fit helmets to minimize slip under heavy G-load	Visor and mask custom trim

Viper Series (United Kingdom)

The U.K.-developed Viper HMD series included three models for fixed-wing operation. GEC-Marconi Avionics (now BAE Systems) developed the Viper 1 and 2 HMDs, which are CRT-based systems (Cameron and Steward, 1994). The Viper 1 became available in mid-1990s as a monocular, visor-projected HMD. It uses a 1-inch diameter miniature CRT display projected via an optical relay assembly, and it employs the standard aircrew spherical visor with the addition of a 70% transmission neutral density coating. This has the advantage of not coloring the ambient when viewed through. It is primarily a stroke-mode day-system, although it can also display raster images. The Viper 1 provides 20° circular FOV, with 15-mm exit pupil, and 70-mm eye relief. Excluding the oxygen mask, it weighs 3.8 lbs (1.7 kg). It was flight tested in the X-31 and also in the F-16 to demonstrate look and shoot capability.

The series continued with the Viper 2. It was BAE's first binocular visor-projected HMD and was flown in the JAST AV-8B (U.S. Version of Harrier), German Tornado, U.K. Tornado, and various F-16s. Designed in a binocular configuration, it used two of the same CRTs as Viper 1 and maintained the visor projection approach using a spherical visor with 70% transmission neutral density coating. The system was configurable to symbology only (stroke-mode), video display from an external source (raster mode) or hybrid video with symbology overlay (stroke-on-raster). It provided 40° FOV with full overlap, a 15-mm exit pupil, and a 70-mm eye relief. Excluding the oxygen mask it weighs 4.2 lbs (1.9 kg).

The Viper 3 (late 1990s) was designed to be a visor-projected NVG replacement and was first flight tested in the Dutch Air Force F-16. The Viper 3 exploits the visor projection scheme common to HMDs and employs multiple-folded optical paths to carry the imagery from a pair of 18-mm I^2 tubes to the pilot's spherical visor. This provides the pilot with an unobstructed binocular 40° FOV NVG capability on his see through visor. The I^2 tubes are mounted on the sides of the helmet, to provide the best possible balance for low fatigue and safe ejection. The helmet is considered suitable for loads of up to 5-6Gs.

An important feature of the optical design of the Viper 3 is that the addition of a dichroic beamsplitter to one of the mirrors in the optical path between the image intensification tubes and the visor allows the addition of a CRT to the Viper 3 design so that the system can become a combined projection HMD and NVG package, with the addition of a CRT and head tracking sensors. The addition of a CRT adds some weight but improves the center-of-mass of the overall system. The Viper 3 design solves the principal problems associated with conventional clip-on ANVIS.

There was also a limited development of a Viper 4 in the late 1990's, which was an extension of the Viper 2; it was extensively flown on VISTA F-16 and used for JSF development trials. Both CRT and flat panel display versions were produced.

Crusader (United States, United Kingdom)

The late-1990s Crusader HMD (Figure 3-20) was part of a technology development/ demonstrator program aimed at providing helmet solutions that can be applied into several fixed- and rotary-wing applications while at the same time maintain the protection levels and life support integration of current in-service helmets. The program was coordinated by the U.S. Navy, who very early-on expressed strong interest in the two-part helmet concept.

The Crusader HMD is a binocular, visor-projection design, has a 30 by 40 degree partial-overlap FOV, and incorporates dual, integrated camera-coupled I^2 tubes. The visor projection design is based on off-axis holographic optics, and provides unobstructed see-through vision with an eye relief of 76mm and extremely well balanced center-of- gravity. The Crusader system utilized dual, miniature solid state displays with a resolution of 1024 vertical by 1280 horizontal. The Crusader HMD is capable of presenting binocular on-helmet I^2 video, aircraft-provided FLIR video, and the merged, "sensor fusion" combination of these, all with both flight and fire-control symbology added.

TopSight (France)

Rather than designing an HMD around an existing helmet shell, Thales Avionics (Vélizy-Villacoublay, France), (at the time, Sextant Avionique) teamed with Intertechnique to design a new helmet system integrating the vision system with the oxygen positive pressure breathing and full nuclear, biological, and chemical (NBC) protection. The futuristic appearance of these helmets results from the use of a flush external face guard, contoured such as not to obstruct the pilot's FOV yet to fully cover the oxygen mask.

The TopSight (previously known as Opsis) (Figure 3-20), was evaluated originally on the Mirage 2000 fighter and subsequently has been used on both the Mirage and the next-generation multirole Rafale fighters. The TopSight is a day-only helmet, configured for air-to-air missions.

The TopSight uses a modular approach. The headgear includes two line-replaceable units: a) the basic helmet, with a custom-fitted form liner and b) a removable Day Display Module, that projects symbology on the pilot's visor for target acquisition and designation; depending on the mission, this module can be replaced by a Night Vision Module (ejection-compatible), or a Double Visor Module (for conventional helmet use).

Designed primarily for target acquisition and designation in support of the Mirage 2000 and Rafale, the air-to-air version is a monocular visor projection display with 20° FOV and 60- mm eye relief. It uses a 0.5-inch diameter CRT in stroke-only symbology, generated from target and aircraft parameters. The fully integrated system, including the oxygen mask, has a head-supported weight of 1.45 kg (3.2 lbs).

TopNight

The TopNight (Figure 3-20) is a TopSight helmet configured for air-to-ground and night mission for the Rafale fighter. It adds to the TopSight an image-intensified charge-coupled device (I^2CCD) camera and binocular display capability. It also adds FLIR image capability from an aircraft sensor or a night-vision image intensified image from the helmet-mounted CCD. The pilot can switch between the external FLIR and I^2CCD sensors. There is also the option of presenting an image received from an outside video source.

The TopNight has a binocular display with a 40- x 30-degree FOV and 60-mm eye relief. It uses two ½-inch diameter CRTs. Aircraft and targeting data are displayed both in stroke (symbology) and raster video imagery (IR, image-intensified tubes [I^2T] and television [TV]). The fully integrated assembly, including the oxygen mask and the I^2T, has a head-supported weight of 1.8 kg (4 lbs).

Joint Helmet Mounted Cueing System (JHMCS)

Following Joint Mission Element Needs Statement (JMENS) signed by the U.S. Air Force and U.S. Navy in mid-1994, the Joint Helmet Mounted Cueing System (JHMCS) (Figures 3-10 and 3-20) became the first joint office project. The JHMCS was developed over the period 1996-99 by Vision Systems International (VSI),[7] San Jose, CA, and is deployed on F-15, F-16 and F/A-18. VSI was formed in 1996 as a joint venture between Rockwell Collins (San Jose, CA) and Elbit Systems (Haifa, Israel) to address HMD opportunities for fixed-wing applications. The JHMCS is a multi-role system that enhances pilot situation awareness and provides head-out control of aircraft targeting systems and sensors.

The JHMCS uses visor projection design with a ½-inch CRT. It is monocular (right eye only), provides only daytime stroke symbology, uses an electro-magnetic tracker, and has a 20° FOV.

In May 2003, VSI was selected to develop a dual-seated version of the JHMCS so that both pilots, in a two-seater fighter, can share information. Deliveries of the modified version started in early 2007 for the Navy’s two-seat F-18F. In a dual-seat aircraft, each crewmember can wear a JHMCS helmet, perform operations independent of each other, and have continuous awareness of where the other crewmember is looking.

The JHMCS can best be described as the offspring of the Elbit Systems Dash 3, the Kaiser Electronic Agile Eye and the VCATS HMDs. Unlike the embedded Dash, the JHMCS is a clip-on package.

The system provides low-weight, optimized center-of-mass with in-flight replaceable modules to enhance operational performance – including the ability to be reconfigured in-flight to meet night vision requirements.

The JHMCS has been introduced with the main goal of slaving the AIM-9X GEN-4 air-to-air Sidewinder Missile to the pilot line-of-sight; this will provide “first look, first shot” capability when employed with high off-boresight weapons and under high-G conditions. Production representative units were delivered in mid 1998, operational tests started in 1999 (first flight test took place in January) on an Air Force F-15 Eagle and a Navy

[7] VSI is a joint venture company between EFW, Inc. of Ft. Worth, Texas and Rockwell Collins, San Jose, CA.

F/A-18 Hornet and production deliveries commenced in 2000. It is used with the HGU-55/P helmet in F-15, F-16 and F-18 fighters.

VSI was authorized to begin full scale JHMCS production in January 2004; by January 2006 VSI advertised the delivery of the 1000th JHMCS helmet. A year later VSI had delivered over 1400 units to 14 nations.

A current list of international customers by fighter aircraft deployment includes:

- F-15 – U.S. Air Force and Air National Guard, Korea
- F-16 – U.S. Air Force and Air National Guard, Belgium, Chile, Denmark, Greece, Netherlands, Norway, Oman, Poland, Turkey
- F/A-18 – U.S. Navy, Australia, Canada, Finland, Switzerland

The U.S. Navy is pursuing an approach to integrate night vision capability into the JHMCS. The goal is for a 40-degree FOV, a typical value for a binocular NVG system. The U.S. Navy would prefer for this design to employ a modular wide-FOV system, such as the panoramic NVG that could increase FOV to as much as 100 degrees by using four I^2 tubes, all of which are slightly shorter and lighter than previous ANVIS-9 version tubes, reducing head strain under increased G-forces. The idea is to inject symbology into the optical train of one of I^2 tubes worn as traditional NVGs.

Scorpion Helmet Mounted Cueing System (HMCS) (United States)

The Scorpion™ HMCS (Figure 3-21) was developed by Gentex Corporation (Simpson, PA) for targeting pod, gimbaled sensor or high off boresight missile cueing mission scenarios. It was designed to interface with existing U.S. Pilot Flight Equipment (PFE), standard oxygen mask variants, Life Support Equipment (LSE) and current fixed-wing NVGs (AN/ANVS-9).

The Scorpion uses a low profile, SVGA color display. In the case of the Gentex HGU-55/P flight helmet, the compact optical element is mounted on the standard NVG helmet attachment. In day mode operation, the ANVIS Day Visor (ADV) is mounted on the helmet NVG jet mount via a ball-detent mechanism. In night operation, the ADV is replaced by the NVGs, which are located directly in front of the display optical combiner. The NVG's night image is viewed through the combiner, providing the pilot with fused NVG scene and color symbology.

The Scorpion also utilizes a low profile, high speed magnetic tracker system to track pilot head position.

The notable discriminators for Scorpion include:

- Left or right eye monocular
- Field of View (FOV): 26° x 19.6°
- Head-supported weight: 2.8 ounces (80 grams)
- Compatible with most visor types
- Compatible with laser eye protection and corrective spectacles
- Ejection system compatible

Scorpion is scheduled to commence operational testing by the US military at the U.S. Air Force / Air National Guard Flight Test Center (Edwards Air Force Base, CA) in 2008.

Figure 3-21. Scorpion Helmet Mounted Cueing System (Gentex Corporation),

Typhoon Integrated Display Helmet (IDH) (United Kingdom)

The Typhoon Head Equipment Assembly (HEA) Integrated Display Helmet (IDH) (Figure 3-20) displays night vision and off-axis cueing information. Selected for the Eurofighter program, the IDH provides 24-hour, all-weather, and all-altitude operation over the full combat profile envelope. Capabilities include weapon/sensor slaving with real-world overlay of flight information, target cueing and night vision.

The system uses a two-part helmet design, with a single size helmet being custom fitted to individual pilots and designed to cover the 5-95th percentile anthropometric range. The helmet provides laser and NBC protection. The helmet operates in conjunction with an optical head tracker, providing low latency head position solutions and eliminating the need for cockpit mapping. It uses dual high-resolution miniature CRTs in stroke, raster and mixed modes to provide a 40° FOV with full overlap, a 15-mm exit pupil, and a 50-mm eye relief. The night vision cameras use two Omni 4 GEN-3 I^2 tubes, capable of operation down to 0.5 millilux) and are detachable.

The helmet employs a dual visor configuration, a clear blast/display visor for night operation and a glare/ laser eye protection visor for day operation.

While the exact location of the I^2 tubes on the side of the helmet is still an issue, this approach will improve helmet dynamic performance, by moving the center-of-mass backward as compared to standard in-front-of-the-eyes I^2 tube mounting. Because the distance between the I^2 tubes exceeds the normal separation distance of the two eyes, the pilot may experience hyperstereopsis. This phenomenon results in objects viewed at close distance appearing closer than in reality, which can cause false cues (Kalich et al., 2007). Flight tests have showed that these effects are perceptible when distance to ground (or objects) is less than about 1,000 feet.

Helmet Mounted Display System (HMDS) (United States)

The Helmet Mounted Display System (HMDS) (Figure 3-20) is being developed for the F-35 Joint Strike Fighter (JSF) by VSI. It has completed all required safety of flight tests, allowing in-flight seat ejections up to 450 KEAS (knots equivalent air speed). It has demonstrated structural integrity to 600 KEAS as a critical risk mitigation step towards full flight certification. The HMDS had its maiden flight on 4/10/2007 on the 10th test flight of the F-35 JSF.

The HMDS provides the pilot video with imagery in day or night conditions combined with precision symbology to give the pilot enhanced situation awareness and tactical capability. For tactical fighter jet aircraft, the F-35 JSF will be the first to fly without a dedicated HUD, with the HMDS providing this functionality.

The HMDS uses the same symbology implemented in the JHMCS. The CRT display in the JHMCS has been replaced by two 0.7-inch diagonal SXGA resolution AMLCDs. The HMDS provides a FOV of 40° (H) x 30° (V).

Military HMD programs: Rotary-wing platforms

While fixed-wing HMD applications abound, the HMD owes its increasing acceptance to rotary-wing aviation. The helicopter environment does not require the HMDs to contend with the demands of high-G maneuvers or ejection with its issue of wind blast. This does not imply that HMD designs for rotary-wing applications are easier. Indeed, the requirements for a wider FOV and increased resolution driven by the common-place nap-of-the-earth (NOE) flight profiles of military helicopters are difficult ones.

Table 3-3 presents a partial summary of the more notable experimental, prototype, fielded and future HMD rotary-wing programs. It followed by summaries of select HMD programs. Many of these HMDs are depicted in Figure 3-22. Many of the programs involved a number of contracts with various commercial HMD developers playing differing roles. Many of the programs also were multi-national in scope. The country of development listed in Table 3-3 and ensuing program descriptions generally is based on the initial developmental phase.

Integrated Helmet and Display Sighting System (IHADSS) (United States)

The first fully integrated head/helmet-mounted display, the IHADSS developed by Honeywell in late 1970’s, and was acquired by (2000) and now manufactured by EFW (Figure 3-22), was fielded by the U.S. Army in the AH-64 Apache helicopter and is still in production.

Historically, the goal of aviation helmet design has been to primarily provide impact and noise protection to the user. In 1981, the U.S. Army fielded an advanced attack helicopter that required a new helmet concept in which the role of the helmet was expanded to provide a visually-coupled interface between the aviator and the aircraft. This new combined helmet and display system, the IHADSS, uses a helmet fitted with infrared (IR) head tracker detectors and a monocular display. The IR head tracker allows a slewable FLIR imaging sensor, mounted on the nose of the aircraft, to be slaved to the aviators head movements. Imagery from this sensor is presented to the aviator through the helmet-mounted display.

The IHADSS HMD consists of a fully functional flight helmet to which the monocular display is mounted. The display can present to the pilot’s eye combinations of aircraft symbology (e.g., heading, torque, altitude, etc.), a targeting crosshair, and pilotage imagery that originates from the FLIR sensor mounted on the nose of the aircraft. The IHADSS has also been used by Boeing on OH-58D Kiowa and by Agusta, on the A-129 Mangusta.

The IHADSS’ major capabilities include:

- Slaves turreted weapons, missile seekers, and gimbaled night vision sensors to the pilot’s line-of-sight;
- Displays real-world-sized video imagery from night vision sensors directly in front of the pilot’s eye and overlays flight information and fire control symbology over the video imagery;
- Can be operated either independently from each cockpit or cooperatively from both cockpits while allowing cueing between the aircraft’s crew members; and
- Enables NOE navigation by pointing a night vision sensor with natural head movements only.

Table 3-3.
Summanry of selected rotary-wing HMD programs.

Time frame	Program	Country	Platform	Developer	Program status	Notes
1970s	IHADSS	USA	Rotary-Wing Apache	Honeywell	Fielded	First integrated HMD
Early to mid 1980s	Wide-Eye	USA	Rotary-Wing Various	Rockwell Collins	Experimental	
Mid to late 1980s	Eagle Eye	USA	Rotary-Wing	Night Vision Corporation	Prototype	
Late 1980s	AN/AVS-6 ANVIS	Multiple	Rotary-Wing Various	ITT	Fielded	
Late 1980s	MONARC	USA	Rotary-wing	Honeywell	Prototype	
1990s	HIDSS	USA	Rotary-Wing Comanche	Rockwell Collins	Prototypes	The Comanche program was cancelled in 2004
1990s	MiDASH	Israel	Rotary-Wing Various	Elbit Systems	Fielded	
Late 1990s	Knighthelm	UK	Rotary-Wing Various	BAE Systems	Fielded	

Table 3-3.(continued)
Summanry of selected rotary-wing HMD programs.

Time frame	Program	Country	Platform	Developer	Program status	Notes
Mid 1990s	Crusader	US/UK	Fixed- & Rotary-wing	Gentex/ BAE Systems/ Thales	Experimental	Technology demonstrator
Late 1990s	TopOwl	France	Rotary-Wing Euro helicopter	Thales	Fielded	Selected for the AH-1Z Cobra
Mid 1990s	ANVIS/HUD-7	Israel	Rotary-Wing Various	Elbit Systems	Fielded	
Mid 1990s	ANVIS/HUD-24	Israel	Rotary-Wing Various	Elbit Systems	Fielded	
Late 1990s/Early 2000s	VCOP	USA	Rotary-Wing Various	Microvision	Experimental	Technology demonstrator
Early 2000s	HeliDash	Israel	Rotary-Wing Miscellaneous	Elbit Systems	Fielded	
Mid 2000s	MIHDS Air Warrior Block 3	USA	Rotary-Wing Various	Microvision	Development	Spectrum SD 2500
Late 2000s	Q-Sight	UK	Rotary-Wing Various	BAE Systems	Experimental	Technology demonstrator

Figure 3-22. Current and future rotary-wing HMD programs.

Primary IHADSS performance characteristics include:

- Image brightness compatible with 2,000-foot-Lambert (fL) background luminance scene; it lacks luminance performance required for optimal gray-scale operation during most daylight missions
- Monocular, right eye only, 1-inch diameter CRT image source
- Display FOV: 40° (H) by 30° (V)
- Exit pupil: circular, 10 mm in diameter
- Video format: Raster only 525 to 875 lines (auto line lock), compatible with GEN-1 FLIR
- Optical eye relief: 10 mm

User performance of the IHADSS is well documented (Rash, 2008). Its visually demanding monocular design has been successful in its deployment in the AH-64 Apache helicopter but has been plagued since initial fielding by frequent pilot reports of visual symptoms and complaints (Hale and Piccione, 1989; Behar et al., 1990). However, during most recent challenge of combat in Iraq, these reports have decreased (Hiatt et al., 2004; Heinecke et al., 2008).

Wide-Eye™ (United States)

The Wide-Eye, ™ designed by Kaiser Electro-Optics, San Jose, CA, and first conceived in the 1980s, was a integrated binocular HMD with retractable combiners for day and night use. It had two 1-inch CRTs as well as I^2 tubes. A modular approach was employed where the optical subsystem is detachable and remains with the aircraft. The system consisted of the helmet, display electronics unit, head-tracker and boresight reticle control unit.

The Wide-Eye™ was a partial-overlap design. Each optical channel has a monocular FOV of 40°; with a 50% overlap, the binocular FOV is 40° (V) by 60° (H) (Zintsmaster, 1994). This system was the precursor to Kaiser Electro-Optics's SIM EYE™ XL 100A design (Kaiser Electro-Optics, 2007).

Tactical-Air Night Vision Display System (Eagle Eye) (United States)

The Tactical-Air Night Vision Display System, built by Night Vision Corporation, and commercially known as Eagle Eye,™ was a low-profile, helmet-mounted, image intensifying system. It was a self-contained system, consisting of two GEN-3 I^2 tubes, folded optics beamsplitters, external housing, and integrated power supply. The folded optical path was designed to allow the I^2 sensors to be located slighted below and to the side of each eye, making the total separation between centers approximately 126 mm (5 inches). The effective interpupillary distance (IPD) was approximately twice the normal 64-millimeter (mm) value. Like ANVIS, the nominal FOV was 40 degrees and fully overlapped. The objective lenses could be focused from 11 inches to infinity. While there was no eyepiece optical adjustment, eyepiece lenses could be inserted in 2-diopter increments to compensate for spherical refractive error ranging from - 6 to +2 diopters. Adjustments included fore-aft, vertical, tilt, and IPD. The Eagle Eye had a limited production in the 1980s.

Aviator's Night Vision Imaging System (ANVIS) (United States)

The ANVIS (Figure 3-22) is by far the most widely used HMD in the world. The ANVIS is a combined sensor/display optics package that mounts unto existing aviation helmets by means of a visor assembly mounting bracket. Over the last two decades, improvements in the I^2 technology used in the ANVIS have given rise to a number of generations and models, all of which loosely referred to as the ANVIS. In the U.S. Army, all ANVIS are AN/AVS-6 models, with current fielded versions identified as types 4 to 6 that define when they were procured and with corresponding performance enhancements. The ANVIS-9 designation is one used by the U.S. Navy and Air Force. It has identical performance but the helmet mount is slightly longer and at a different tilt in order to be compatible with Air Force and Navy helmets. The ANVIS-9 also has an internal filter that blocks more of the visible spectrum (related to lighting compatibility issues). The ANVIS is a binocular, 40°, 100% overlap system using GEN-3 I^2 tubes, which being head-mounted, does not require an additional head tracking system.

Typical ANVIS-6 optical characteristics include:

- Focus range: 28 cm (11 inches) to infinity
- Magnification: Unity (1X)
- 27-mm effective focal length objective (f/1.23)
- Resolution: >1.3 cycles/milliradian (cy/mr)
- Brightness gain: minimum 2000x (5,500X for newer versions)
- Diopter eyepiece focus adjustment
- Interpupillary distance (IPD) adjustment: 52-72 mm

The ANVIS housing can be flipped up or down and has an 11-15G breakaway threshold. A tilt adjustment of approximately 10º is provided. There is a minimum vertical and fore/aft adjustment range of 25 mm. They operate off of a single lithium or two “AA” batteries. A dual battery pack is Velcro™ mounted on the rear of the helmet to improve the CM. An historical summary of the ANVIS and its predecessors is provided by McLean et al. (1998).

Monolithic Afocal Relay Combiner (MONARC) (United States)

The Integrated Night Vision System (INVS), built in the late 1980s and early 1990s by Honeywell, Inc., Minneapolis, Minnesota, and commercially known as the Monolithic Afocal Relay Combiner (MONARC), consisted of a helmet subsystem, a binocular image display system, and provisions for a magnetic head tracker. The helmet included a visor, energy liner, retention system, communications, thermoplastic liner, image display, magnetic receiver mounts, and electrical interfaces. Imagery, from binocular I^2 sensors and dual (binocular) CRTs, with added symbology was designed to be displayed through the imaging system which consisted of separate modules mounted to each side of the helmet. The modules were powered by an ANVIS-style battery pack. Each module contained a GEN-3 I^2 tube, CRT, objective and relay optics and beamsplitter. (Note: The MONARC combiner used the principle of total internal reflections to relay the image from the CRT image source to the eye.) The I^2 sensors were located beside and slightly above the user's eye, making the total separation distance between sensors (and effective IPD) approximately 254 mm (10 inches) (4X normal IPD). The objective lenses could be focused from 6 meters to infinity. The vertical and lateral IPD positions of each module could be adjusted independently, but there was no fore-aft or tilt adjustments. This system provided a nominal 35°, fully overlapped FOV.

Helmet Integrated Display Sight System (HIDSS) (United States)

In the 1990s, the U.S. Army was developing the next-generation armed reconnaissance helicopter, the RAH-66 Comanche. Integral to this aircraft was an HMD designed by Kaiser Electronics, San Jose, CA. The HMD was the Helmet Integrated Display Sighting System (HIDSS) (Figure 3-22). While the Comanche program was cancelled by the Army in February, 2004, the HIDSS development program led to a number of interesting and useful concepts in HMD design.

The initial HIDSS design was based on the Wide-Eye integrated binocular design. It originally provided a 40° (V) by 40° (V) FOV with 50% partial-overlap. Ultimately, the FOV specification became at 30° (V) by 52° (H), matching the anticipated GEN-2 FLIR sensor, with at least 30% overlap. The first HIDSS design incorporated two 1-inch diameter CRTs. While image quality was found to be acceptable, the addition of a second CRT (as compared to the IHADSS single CRT) pushed the total head-supported weight beyond the Army’s acceptable safety limits (Harding et al. 1998). A follow-up HIDSS design replaced the CRT image sources with miniature LCDs.

The HIDSS also used a modular approach, partitioning the system into an Aircraft Retained Unit (ARU) and a Pilot Retained Unit (PRU). The ARU was detachable from the helmet and remained stowed in the aircraft at all times; the PRU was a custom-fitted helmet and was retained by the pilot.

The technical performance goals for the HIDSS program included:

- SXGA Resolution: 1280 x 1024 pixels
- Luminance: 1500 fL, at the eye
- Modulation transfer function (MTF): 8% (H and V with one line-on/one line-off
- Exit pupil: 15 mm
- Eye relief: 25 mm
- Head-supported mass: Not to exceed 2.4 kg (5.3 lbs)

Modular Integrated Display and Sight Helmet (MiDASH) (Israel)

The MiDASH (Figure 3-22), manufactured by Elbit Systems Limited, Haifa, Israel, helmet, was designed to provide attack and reconnaissance helicopter pilots with wide-FOV, see-through binocular night imagery, flight information and line-of-sight cueing for day and night operation (Elbit Systems, 2004).

MiDASH comprises a standard helmet shell with a personal fitting device. The left and right optical modules are referred to as Helicopter Retained Units (HRUs) and are attached to the helmet by snap-connectors.

System performance specifications:

- Binocular, night imagery FOV: 50°H x 40°V (partial-overlap)
- Symbology FOV: 30° circular
- See-through transmission: >50%
- Eye relief: >50 mm
- Night vision: "Super GEN '98" or GEN-3
- Total mass (night operation): 2.2 kg (4.9 lbs)

Knighthelm (United Kingdom)

The Knighthelm (Figure 3-22), manufactured by BAE Systems, is a first-generation HMD featuring a modular (two-part)design, with a basic form-fitted helmet designed specifically for HMD applications. The display's image sources and optical components are integrated into the helmet such that the fundamental properties of the helmet (e.g., protection, weight, CM) were not compromised (White and Cameron, 2001). The Knighthelm HMD provides a full day/night mission capable system in a binocular, 40° FOV, full-overlap configuration.

Knighthem provides night vision capability via either imagery from an aircraft-mounted FLIR sensor or a pair of GEN-3 I^2 tubes integrated into the helmet. The FLIR imagery, combined with flight and weaponry symbology, is projected onto the two combiners.

A dual-visor system is fitted to the display module: a clear visor (Class 1) that can be alternated with a laser protection visor and a neutral density visor (Class 2) for glare protection. For ease of replacement the visors are mounted on quick release pivot assemblies.

The Knighthelm's initial 1990's design has been refined and enhanced, as part of an extensive development program, for the German Army Tiger helicopter, and is optimized for the attack helicopter application (White and Cameron, 2001).

Major Knighthelm performance specifications include:

- Exit pupil: 15 mm
- Eye relief: 30 mm
- See-through transmission: 70%
- Symbology overlaid on image intensified or sensor imagery
    - Cursive (stroke) symbology visible in all ambient conditions
    - Selectable binocular/ monocular CRT symbology presentation
- Weight: 2.2 kg (4.9 lbs)

Crusader (United States, United Kingdom)

While there was never a formal developmental program for a rotary-wing Crusader HMD, the fixed-wing version was developed with the potential of rotary-wing use, with specific attention paid to the greater impact and penetration requirements for the HMD helmet platform. (See fixed-wing description in the *Military HMD programs: Rotary-wing platforms* section of this chapter.)

TopOwl (France)

The TopOwl™ (Figure 3-22) is manufactured by Thales, France. It has a fully-overlapped, visor projection system, capable of presenting FLIR, I^2 and synthetic imagery. The visor projection approach improves viewing of the outside world over standard HMD designs that require optical beamsplitters. This approach also allows for increased physical eye relief (>70 mm [>2.75 inches]), which reduces potential interference with the wearing of corrective spectacles. Dual I^2 sensors are located on the sides of the helmet with a separation distance of approximately 286 mm (11.25 inches) (an effective IPD of more than 4X normal). The I^2 imagery is optically-coupled to the visor. The FLIR imagery from a nose-mounted thermal sensor is reproduced on miniature CRTs (current production version) or LCDs (prototype) and projected onto the visor. In I^2 mode, it presents a 40° circular FOV; for FLIR imagery presentation, the FOV is 40° (H) by 30° (V).

The production CRT version is currently fielded on various models of the Eurocopter Tiger and Denel AH-2 Rooivalk helicopters and in use in 15 countries. It has been selected for use on the U.S. Marine Corps AH-1W Super Cobra attack helicopter.

The total weight of a fully configured production CRT-version of TopOwl has a mass of 1.8 kg (4 lbs) for day-only operations and 2.2 kg (4.8 lbs) for the nighttime configuration.

ANVIS/HUD-7 and -24 (Israel)

The major disadvantage of legacy I^2 systems (e.g., ANVIS series) is the lack of symbology. An approach to solve this deficiency is the ANVIS/HUD, developed by Elbit Systems. The first version is the ANVIS/HUD-7, which combines the standard ANVIS goggles image with aircraft flight instrumentation and computer graphics during night operation (Figure 3-22). The system can be installed on any type of helicopter. Figure 3-23 presents sample ANVIS/HUD-7 imagery consisting of symbology overlaid on I^2 imagery.

Major technical performance specifications of the ANVIS/HUD-7 include:

- FOV:
  - Night vision - 40°
  - Symbology - 32° overlaid on the night imagery without degradation to the ANVIS image
- Resolution: > 512 x 512 pixels
- Mass: <110 g (3.9 ounces)
- Compatible to GEN-2, GEN-3 and OMNIBUS I^2 systems
- Attachable to the right or left objective
- Compatible with NBC mask or eyeglasses
- Quick disconnect for safe egress

Elbit Systems Limited developed the Day/ Night ANVIS/HUD-24 from the ANVIS/HUD-7 system above, with the DAY HUD add-on module, the system projects imagery of flight information to enable head-out flight during the day time (Yona et al., 2004). By combining the standard ANVIS imagery with aircraft flight instrumentation symbology, the ANVIS/HUD offers 24-hour operational capability. The system supports two-pilot operation, with eight selectable display screens and can be installed on any type of helicopter; it is currently operational on more than 25 different platforms.

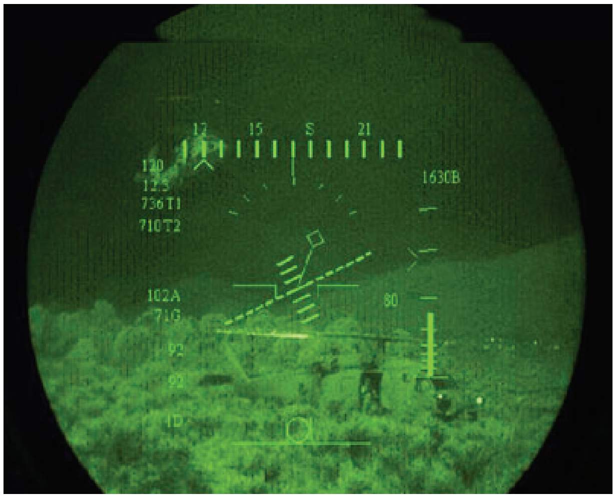


Figure 3-23. Typical ANVIS/HUD-7 imagery: Symbology overlaid on night imagery.

Performance values for night operation are identical to the ANVIS/HUD-7. The day channel performance is defined by:

- Day FOV: 25°
- See-through transmission: 36%
- Brightness: 500 fL
- Exit pupil: >12 mm
- Eye relief: >50 mm (may be used with NBC mask or eyeglasses)
- Head-supported mass: 200 grams (7.1 ounces)

EyeHUD™ (United States)

The EyeHUD™ (Figure 3-24), developed by Rockwell Collins, Cedar Rapids, IA, is a compact, light-weight monocular HMD designed as a alternative to the ANVIS/HUD. It is designed to attach to the standard ANVIS mount. Using a miniature AMLCD, its goal is to provide pilots basic HUD situation awareness capability (e.g., aircraft flight, engine performance and weapons symbology) in both day and night operations (Rockwell-Collins, 2008a) The EyeHUD™ HMD can be used with any military aviator helmet. It provides a full range of IPD and vertical adjustments while accommodating laser eye protection and aviator eyewear.

Major technical performance features include:

- Day FOV: 26° (Diagonal)
- Resolution: 800 x 600 (SVGA)
- Head-supported mass: 95 grams (2.6 ounces)
- Compatible with ANVIS Class A and B spectral response

Figure 3-24. EyeHUD.™

QuadEye™ (United States)

QuadEye™ (Figure 3-25a) was developed by Kollsman, Merrimack, NH and is an advanced Panoramic Night Vision Goggles (PNVG) providing a central 40° binocular FOV plus monocular vision of an additional 30° to either side (Figure 3-25b) (Kollsman, 2008). The impetus of this expanded FOV design is to provide a FOV similar to the normal eye's peripheral vision, thereby reducing the need to increase head movement when wearing the ANVIS. QuadEye is designed around four 16-mm I^2 tubes of which the pilot can select either only the two inner tubes or all four (panoramic) tubes. Additionally, QuadEye™ can provide HUD symbology or aircraft targeting sensor imagery using a miniature, high resolution display.

Main system performance values include:

- FOV: 100° (H) by 40° (V)
- Physical eye clearance: 32 mm
- Brightness gain: > 5,500:1
- Mass (with four I^2 tubes, display, camera): 700 grams (25 ounces)

The U.S. Army's Virtual Cockpit Optimization Program (VCOP) was a virtual cockpit simulator program. Its goal was to provide the pilot with a simulated environment where he/she could train with information such as situational awareness, sensor imagery, flight data, and battlefield information in a clear, non-confusing and intuitive manner (Moore et al., 1999; Harding et al., 2004). VCOP was comprised of six technologies:

- Full color, high resolution, high brightness HMD that incorporates Virtual Retinal Display (VRD) technology
- 3-D audio
- Speech recognition
- Situation awareness tactile vest
- Intelligent information management
- Crew-aided cognitive decision aides

Figure 3.25a. QuadEye™ (www.kollsman.com).

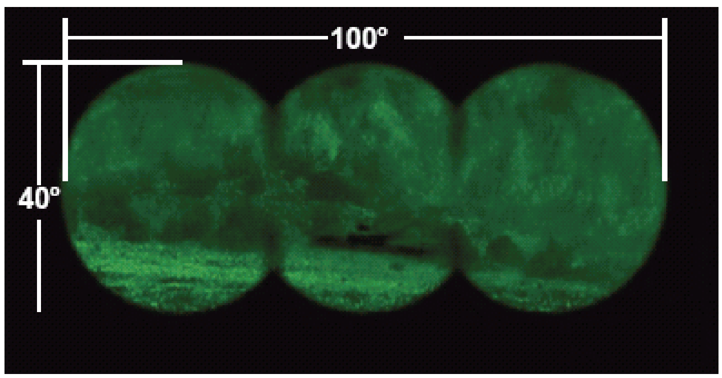


Figure 3-25b. QuadEye™ field of-view (www.kollsman.com).

Virtual Cockpit Optimization Program (VCOP) (United States)

VRD technology was invented at the University of Washington in the Human Interface Technology Lab (HIT) in 1991. Its goal was to produce a full color, wide FOV, high resolution, high brightness, low cost virtual display using miniature scanned lasers. The original VRD concept used scanning lasers to form an image directly on the retina.

Microvision, Inc., Seattle, WA, has the exclusive license to commercialize the VRD technology and was the developer of the VCOP HMD. The VRD scanning laser technology has been pursued for a number of HMD programs. HMD applications have deviated from the original VRD concept in that the scanning lasers do not scan directly on the retina but instead form an intermediate image that is viewed via an eyepiece. Figure 3-26 shows an early prototype developed under the U.S. Army's Aircrew Integrated Helmet System (AIHS) alternate image source development. Figure 3-22 shows a futuristic version of the VCOP design.

Figure 3-26. Prototype AIHS scanning laser HMD (Microvision, Inc.).

HeliDash (Israel)

HeliDash is a modular day/night display and sight helmet designed by Elbit Systems for attack, assault and utility helicopter applications. It provides the pilot with high-resolution night vision and day/night symbology. The system configuration includes electronics, clear/dark visors, a night module (ANVIS-HUD), and a day Module (DASH 20° FOV visor-projected symbology).

Modular Integrated Helmet Display System (MIHDS) (Air Warrior) (United States)

The Air Warrior program is one of a group of U.S. Army Warrior Soldier Warrior programs. General Dynamics C4 Systems (Scottsdale, AZ) is the prime contractor and system integrator for all of these systems, which additionally include Land Warrior and Mounted Warrior.

Air Warrior is intended to provide U.S. Army rotary-wing aircrew with advanced life support, ballistic protection, and NBC protection in rapidly tailorable, mission-configurable modules. Its development has in been in a 3-block format. Block 1 included the development, procurement, and fielding of a micro climate cooling system, an integrated survival gear and ballistic protection system, and a light-weight chemical and biological protection ensemble. The on-going Block 2 technology insertion phase of the program provides additional capabilities, including an Electronic Data Manager and an Aircrew Wireless Intercom System. Block 3 is focused on increasing force effectiveness by improving situation awareness and survivability. The Air Warrior systems must be compatible with multiple helicopter types, including the CH-47 Chinook, OH-58D Kiowa Warrior, AH-64 Apache and UH-60 Blackhawk. It also is required to have compatibility interoperability with the Army's Land Warrior and Future Combat Systems programs.

Integral to the Block 3 phase is the development of an HMD. General Dynamic’s program to provide the HMD is Modular Integrated Helmet Display System (MIHDS). The MIHDS will provide integration and interface of symbology, imaging sensors, and head-position tracking devices, permitting the aircrew a clear view of the external environment during both day and night operations.

Microvision’s ™ SD2500 (Figure 3-27), a descendent of VCOP, is a candidate system for the MIHDS. The SD2500 design provides a full-color, see-through, daylight and night-readable, high-resolution (800X600 pixels) display (Microvision, 2005). This HMD is fitted for attachment to the U.S. Army’s standard aviation helmet, Head Gear Unit 56P (HGU-56P), via the common Aviator's Night Vision Imaging System (ANVIS) mounting bracket.

Major performance specifications include:

- HMD type: Monocular, Color RGB
- See-through transmission: >50%
- FOV: 23º (H) x 17º (V)
- Resolution : SVGA, 800 (H) x 600 (V) pixels
- Luminance (at the eye): >1000 fL, D65 white
- Physical eye relief: > 50 mm
- Interpupillary distance (IPD) range: 29-36 mm from center

Q-Sight (United Kingdom)

The Q-Sight™ (Figure 3-22) is being developed by BAE Systems. Its design employs holographic wave-guide technology. Weighing less than 4 ounces (113 grams), it has no bulky projection optics and offers an exceptional center-of-mass.

Q-Sight's miniature display is easily adaptable to any standard helmet as either a left- or right-side configuration (at approximately 25 mm), allowing the pilot to choose his or her dominant eye. A binocular configuration also is available.

Symbology and/or video can be displayed to provide the pilot with eyes-out operation (Figure 3-28). In day (high-ambient-light) conditions, a dark visor can be deployed to improve the image contrast. Q-Sight is designed to be compatible with the current NVGs. Operation at night is achieved by attaching the NVG and deploying in the normal manner. The Q-Sight display is located in its own mount and positioned behind the NGV eyepiece (BAE Systems, 2007). Flight demonstrations of the Q-Sight system are planned for late 2007 and early 2008.

Major performance specifications include:

- FOV: 30°, monocular
- Luminance: 1800 fL
- Contrast ratio 1.2:1
- Exit pupil: > 35 mm
- Eye relief: > 25 mm
- Power consumption: <5 watts, head-mounted
- Head-supported mass: < 113 grams (4 ounces)

Military HMD programs: Mounted and dismounted

In the development and application of HMD technology, aviation has led the way. However, in the early 1990s, the potential of HMDs for mounted and dismounted Warfighters was recognized fully. This has led to a number of development programs that focused on the differing requirements that must be imposed on HMDs intended for ground applications. Not surprisingly, I^2 technology has been the sensor technology of choice in these non-aviation designs. However, the fundamental characteristics of these ground-based HMDs are the result of decades of lessons learned from aviation-based HMDs development programs.

Figure 3-27. Spectrum™ SD2500 (Microvision, Inc.).

Figure 3-28. View of symbology through Q-Sight (BAE Systems).

As HMDs move from air to ground, there will be important economic considerations. While HMDs have been fielded in both fixed- and rotary-wing aircraft for decades, their quantity have been small. This number will change drastically as HMDs are issued to every Warfighter along with his/her weapon and boots. As with any system, the larger the production demand, the smaller the unit cost.

Table 3-4 presents a summary of the more notable experimental, prototype, fielded and future HMD programs for mounted and dismounted applications and is followed by summaries of respective HMD programs.

Combat Vehicle Crew (United States)

The Combat Vehicle Crew (CVC) HMD (Figure 3-29) program, initiated in 1992, was a research and development effort to develop a high resolution, flat panel-based HMD for the Army’s M1 A2 Abrams main battle tank (Nelson, 1994).

Table 3-4.
Summary of selected mounted and dismountedHMD programs.

Time frame	Program	Country	Platform	Developer	Program status	Notes
Early 1990s	Combat Vehicle Crew	USA	M1 A2 Abrams Tank	Honeywell, Inc	Development	
Mid 1990s	Land Warrior	USA	Wearable	Rockwell Collins	Development	ProView SO-35
Late 1990s	Mounted Warrior (MWSS)	USA	Stryker Combat Vehicle	Rockwell Collins	Development	ProView SO-35
Late 1990s	DHTVS	USA	Combat and Combat Support Vehicles	Rockwell Collins	Limited fielding	
Early 2000s	NOMAD	USA	Stryker Combat Vehicle	Microvision, Inc.	Limited fielding	

Figure 3-29. Combat Vehicle Crew (Girolamo, 1997).

The CVC HMD was intended to provide a head-out HMD for tank commanders; it also would allow commanders to track near-range threats, survey the proximal terrain and avoid collision (Girolamo, 1997). The initial design was developed by Honeywell, Inc. (Minneapolis, MN), using a monochrome AMLCD panel that provided a 40° FOV at VGA (640 x 480 pixels) resolution. In 1994, the display was upgraded to SXGA (1280 x 1024 pixels) resolution for integration in the CVC HMD system used in both the Abrams tank and the Bradley fighting vehicle. The system maintained a 40° FOV and was used to project thermal imagery and tactical battlefield information. After an initial operational test, the program was discontinued in 1997.

Land Warrior (United States)

The Land Warrior program is an integrated fighting system for individual infantry soldiers which gives the soldier enhanced tactical awareness, lethality and survivability (SPG Media, 2008). The systems included in Land Warrior are the weapon system, helmet (HMD), computer, digital and voice communications, positional and navigation system, protective clothing and individual equipment. The Land Warrior system will be deployed by infantry and combat support soldiers, including rangers, airborne, air assault, and light and mechanized infantry soldiers.

The Land Warrior program is one of a group of Army Warrior Soldier Warrior programs for which General Dynamics C4 Systems (Scottsdale, AZ) serves as the prime contractor and system integrator.

The Land Warrior program was initiated in 1994. Raytheon Systems, (then Hughes Aircraft Company) was the engineering developer. Plans were drafted to build an Initial Capability (formerly Land Warrior Block 1) and then a Land Warrior Stryker Interoperable (formerly Land Warrior Block 2). In 2003, General Dynamics Decision Systems (now General Dynamics C4 Systems) was selected to enhance the Land Warrior system with integration to the U.S. Army digital communications, interoperability with the Stryker Brigade Combat Vehicle (SPG Media, 2008).

The helmet system is known as the Integrated Helmet Assembly Subsystem (IHAS). It provides required ballistic protection while serving as a platform for a helmet-mounted computer and sensor display, which serves as the Warfighter's interface to digital battlefield. Through the HMD, the Warfighter can view computer-generated graphical data, digital maps, intelligence information, troop locations and imagery from a weapon-mounted Thermal Weapon Sight (TWS) and video camera. This new capability allows the soldier to view around a corner, acquire a target, then fire the weapon without exposing himself, beyond his arms and hands, to the enemy. The thermal images are presented on the HMD.

Currently, the Land Warrior HMD is the Rockwell Collins ProView™ S0-35 (Figure 3-30). It is a monocular design and uses the eMagin's (East Fishkill, NY) full color SVGA active-matrix OLED (AMOLED) display. Major technical performances parameters include:

- Luminance: 0.1-30 fL
- Resolution/FOV:
    - SVGA resolution (800 x 600): 28° x 21° (35° diagonal)
    - VGA resolution (640 x 480): 22° x 17° (28° diagonal)
- Eye relief: >25 mm (Eyeglasses compatible)
- Exit pupil: Non-pupil forming system
- Image source type: Full-color AMOLED 800 (x3) pixels x 600 lines
- Mass: 67 grams (2.4 ounces) Display module (w/out mount), 145 grams (5.1 ounces) (with helmet mount)

Figure 3-30. The Land Warrior HMD concept and the Rockwell Collins ProView™ S0-35 (Rockwell-Collins, 2008b).

The U.S. Army merged the Land Warrior program with the Future Force Warrior (FFW) program in 2005 with General Dynamics C4 Systems as prime integrator. FFW is a Science and Technology initiative to develop and demonstrate innovative capabilities for Future Force Soldier systems. The FFW is scheduled to be fielded in 2010 and will be followed, in 2020 by the Vision Future Force Warrior. FFW is designed to provide a ten-fold increase in lethality and survivability of the infantry platoon. In May 2007, a comprehensive assessment of the Land Warrior (and Mounted Warrior) systems conducted jointly at the U.S. Army Infantry Center, Fort Lewis, WA. More than 400 soldiers of the 4th Battalion, 9th Infantry Regiment, 4th Stryker Brigade Combat Team, 2nd Infantry Division participated. The battalion was equipped with 440 Land Warrior Systems and 147 Mounted Warrior Systems. Following this test and evaluation, an initial set of Land Warrior systems was deployed with the 4-9 Infantry Stryker Battalion in late 2007.

Mounted Warrior Soldier System (United States)

The Mounted Warrior Soldier System (MWSS) (Figure 3-31) is another major component of the Army's Soldier as a System initiative (with Land Warrior and Air Warrior). It is envisioned as an integrated "system of systems" designed to improve the survivability, lethality, and combat effectiveness of Stryker-mounted crewmen. The MWSS leverages capabilities being developed in other warrior programs, such as Land Warrior, Air Warrior and Future Force Warrior.

Figure 3-31. Mounted Warrior Soldier System (MWSS) concept.

Rockwell Collins has been selected by General Dynamics C4 Systems to provide HMDs for Increment I of the Mounted Warrior Helmet Subsystem (HSS) program. The recommended HMD of choice is the ProView S0-35™ monocular. This selection illustrates design re-use opportunities across General Dynamics' warrior programs since Rockwell Collins' HMD is currently qualified for use in the Army's Land Warrior program.
The HMD provides the wearer with the capability to select and view display of information from one of three existing video sources within the Stryker:

- Driver's Vision Enhancer (DVE),
- Remote Weapon System (RWS) via the Video Display Terminal (VDT),
- Force XXI Battle Command, Brigade and Below (FBCB2) display.

In an interesting subsequent development, in September 2006 Microvision, Inc., has been awarded a contract by General Dynamics C4 Systems to supply full-color, daylight readable, see-through HMDs as part of the U.S. Army's Mounted Warrior HMD Improvement Program. Microvision, Inc., will use its scanning-laser technology. The improvement program, managed by the U.S. Army's Project Manager for Soldier Warrior under Program Executive Office Soldier, is looking for reduced size, weight, and power requirements. The contract specifies the development, design, verification, testing, and delivery of ten full-color display units for evaluation by mid-2007.

Drivers Head Tracked Vision System (DHTVS) (United States)

In the late 1990s, the U.S. Army developed a system known as the Drive Head Tracked Vision System (DHTVS) as an aid to drivers of combat and combat support vehicles (Casey, 1999). The system consisted of:

- Uncooled, gimbaled FLIR sensor
- Flat panel display
- Electronics box
- HMD

The HMD had a biocular non-see-through design that mounted onto the driver's helmet. The 30° (V) by 40° (H) FOV of the HMD matched the sensors FOV. The displays were XGA AMLCDs. An IPD adjustment was provided, and the oculars could be swung up out of the driver's field-of-vision.

NOMAD Augmented Vision System (United States)

The NOMAD Augmented Vision System (Microvision Inc.) (Figure 3-32) was developed for use in ground vehicles and has been fielded on Stryker vehicles deployed in Operation Iraqi Freedom (OIF). This HMD allows vehicle commander to stand (down) in his hatch and retain a view of the outside world, hence maintaining situation awareness. Similar NOMAD displays have been designed for use in maintenance, repair and overhaul applications. Being able to present vehicle and equipment repair checklists, parts lists, and schematics and diagrams in a head-up format right at the repair site can increase efficiency and reduce downtime (Rash, 2006b).

The NOMAD class of displays uses a scanning laser display that provides 800 by 600 pixels of resolution. Its manufacturer-cited specifications include:

- Luminance: Up to 1,000 fL
- Shades of grey (contrast metric): 32
- Mass: < 200 grams (7 ounces)
- Operating temperature range: 32-113° F (0-45°C)

Figure 3-32. NOMAD (Microvision, Inc.).

Military HMD programs: Simulation and training

Realistic training and mission rehearsal enhance crew proficiency, mission success and, most importantly, crew survivability. As a consequence of increased U.S. military involvement around the world, the military expects significant future growth in the demand for deployable virtual reality trainers. The effect of the rapid advancement in networking capability, both local-area and satellite-based wide area, computer and display technologies, has resulted in networked deployable trainers scattered around the world that allow U.S. and coalition military personnel to train collectively, in a synthetic, but realistic environment. Realism, necessary for training effectiveness, has been greatly enhanced through the use of very accurate terrain maps generated from aerial and satellite photographs. Collective training, encompassing joint aviation, naval and ground vehicle simulators based in different parts of the world, can today be performed in the same virtual battle space as the result of this networked simulation capability. Visual display capability consistently has been a critical element in successfully training military aviators.

In addition to the VCOP HMD, two major examples of U.S. aviation simulators are the Aviation Combined Arms Tactical Trainer – Aviation Reconfigurable Manned Simulator (AVCATT-A) and the Flight School XXI simulator.

Aviation Combined Arms Tactical Trainer – Aviation Reconfigurable Manned Simulator (AVCATT-A) (United States)

The AVCATT-A is a mobile, transportable, virtual simulation training system that provides Army aviation with the capability to conduct realistic, high intensity training exercises and mission rehearsals for five of the Army's current and future generations of frontline helicopters—the AH-64A Apache and AH-64D Apache Longbow, the CH-47D Chinook, the UH-60 Black Hawk, and the OH-58D Kiowa Warrior. (See earlier discussion of AVCATT in Flight training section of this chapter) Each AVCATT-A unit is housed in two 53-foot-long trailers (Figure 3-33), that have been designed to be deployable on either C-5 Galaxy aircraft or other cargo ships. The system allows pilots to train and rehearse through networked simulation in a collective and combined arms simulated battlefield environment.

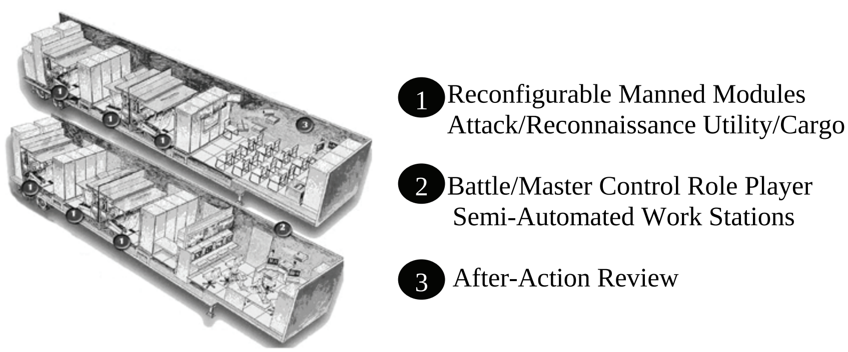


Figure 3-33. AVCATT-A Trailers (Kauchak, 2001).

Each AVCATT-A unit includes 12 HMD systems, the Rockwell Collins' model SimEye XL 100A (Figure 3-34). The SimEye features a full-color SXGA resolution (1280 x 1024) display and presents a 100° (H) x 50° (V) FOV (Rockwell Collins, 2006).

Figure 3-34. SimEye XL 100A (Rockwell Collins).

Major technical performances parameters include:

- Configuration: Binocular, see-through, color
- See-through transmission: > 20%
- Luminance: 1-20 fL (peak white)
- FOV: 100° x 50°, with 30° Overlap
- Resolution: XGA, (1024 x 768)
- Eye relief: >25 mm (Eyeglasses compatible)
- Exit pupil: 15mm
- Mass: 2.5 kg (5.5 lbs) including helmet, optics and displays

Flight School XXI (United States)

A second example of HMD application to simulation and training is in the U.S. Army's Flight School XXI (FSXXI) program. FSXXI is being implemented in the Aviation Warfighter Simulation Center situated at the U.S. Aviation War Fighting Center at Fort Rucker, AL. The primary FSXXI objective is to ensure that the aviators who leave the Fort Rucker, AL, training facility have the necessary experience in their aircraft prior to undertaking combat missions. All future army aviators will be trained under the FSXXI program.

Flight School XXI uses of three types of simulators: the Operational Flight Trainer (OFT), which is the highest fidelity training device that has a wide visual display and is motion-based; the Instrument Flight Trainer (IFT) (Figure 3-35, Top), which is essentially the same as an OFT except it is not on a motion platform and has a smaller visual presentation; and Reconfigurable Collective Training Devices (RCTDs), which enable collective training and can be reconfigured to simulate the Army's UH-60A/L, CH-47D, OH-58D, AH-64A and AH-64D aircraft (Chisholm, 2006). Integral to the IFT cockpits are HMD systems (Figure 3-35, Bottom). Currently, the HMD employed is the Advanced Helmet Mounted Display (AHMD) developed by Link Simulation and Training, Arlington, TX (an L-3 Communication company).

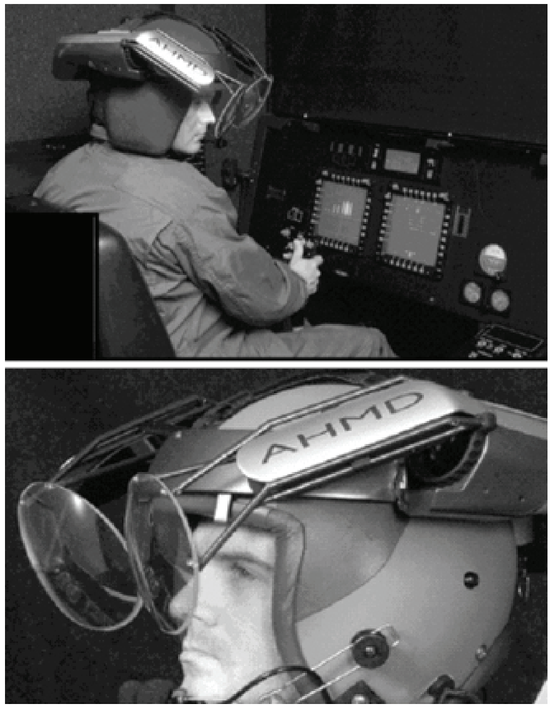


Figure 3-35. Flight School XXI simulator (Top) and Advanced Helmet-Mounted Display (Bottom) (Sisodia et al., 2007).

Major technical performances parameters of the AHMD include:

- Configuration: Binocular, see-through, color
- See-through transmission: > 60%
- Luminance: 0.02-22 fL (peak white)
- FOV: 100° x 50°, with 30° overlap
- Resolution: SXGA, (1280 x 1024)
- Eye relief: > 60 mm
- Exit pupil: 15mm

Medical platform

Advanced Flat Panel (AFP) (United States)

The medical community has developed a broad range of procedures and methodologies that require use of high resolution color video technology. The Advanced Flat Panel (AFP) program’s goal (Girolamo, 1997) was to develop color VGA and SXGA and monochrome UXGA (2560 x 1280 pixels) stereoscopic HMDs for arthroscopic and endoscopic surgical applications that meets the comfort and performance requirements for an operating room environment – including the need for sterilization. Two major applications were identified: medical surgery and diagnostic systems that use color video borescopes and portable information display systems that use high resolution computer graphics and the AFP program was initiated by DARPA in June 1994. The AFP design focused on three critical aspects of the system (Nelson and Helgeson, 1996):

- High quality color imagery comparable to that available via 21” CRT monitors used in the operation room
- Exceptional user comfort – both mechanical and visual – so as to not increase surgeon’s physical burden or stress while using the HMD
- System compatibility with the operating room, including other user-worn equipment and cleaning requirements.

U. S. Army surgeons from the Madigan Army Medical Center (Tacoma, WA) and the U. S. Army 47th Combat Support Hospital performed 15 arthroscopic knee surgeries, including the first ever arthroscopic surgeries I a field-deployed Combat Support Hospital using the system (Nelson et al., 1997). It was generally agreed that the HMD provides additional benefits for the combat medical community in warfighting environment.

One of the most difficult requirements for medical HMD systems is the color gamut and rendition quality, as surgeons relay heavily on color and color discrimination. This is further complicated by the criticality of the shades of gray accuracy to monitor subtle color changes, particularly in red and blue. The flat panel technology at the time had difficulties meeting these requirements, an Operational Requirement Document was never generated and the program terminated in 1997

User Acceptance

Every day, the “next great idea” ends up as a failure in the eyes of the consumer. Unless the need (real, induced or imagined) for a product is paramount to the task at hand or to health and safety (and that does not always win out), user acceptance usually is the more overriding factor.

From their first conception, HMDs have had to overcome their disadvantages of increased head-supported weight and center-of-mass offsets being the most difficult. These and other inherent HMD characteristics impact comfort, which is a major factor in user acceptance.

However, physical discomfort associated with HMDs may be of lesser importance when compared to potentially disastrous consequences if sensory, perceptual and cognitive issues associated with the design and use of HMDs are not as equally taken into account and carefully investigated.

This is especially true in military scenarios where the mismatch of HMD sensory inputs to the human senses may result in loss of information transfer at best and loss of situation awareness at worst, a consequence that may result in loss of life and equipment.

References

Ackerman, J. (2002). Application of augmented reality to laparoscopic surgery. Doctorate dissertation submitted to the faculty of The University of North Carolina at Chapel Hill.

BAE Systems. (2007). The Q-Sight™ family of helmet display products. Retrieved March 2, 2008 from: www.baesystems.com/BAEProd/ groups/public/documents/bae_publication/ bae_pdf_eis_q-sight.pdf

Bailey, R., Arthur J., III, Prinzel L., III, and Kramer, L. (2007). Evaluation of head-worn display concepts for commercial aircraft. *Proceedings of SPIE*, 6557, 65570Y-1 to 65570Y-16.

Bajura, M. (1992). Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. Computer Graphics, *Proceedings of SIGGRAPH.*, 26(2), 203-210.

Bayer, M. (January 4, 2007). Personal files.

Beal, C., and Sweetman, B. (1994). Helmet-mounted displays: Are we jumping the gun? *International Defense Review*, September, 69-75.

Beal, C., and Sweetman, B. (1997). A new vision: Helmet-mounted displays take multi-track approach. *Jane International Defense Review*, August, 41-47.

Behar, I., Wiley, R.W., Levine, R.R., Rash, C.E., Walsh, D.J., and Cornum, R.L. (1990). Visual Survey of Apache Aviators (VISAA). Fort Rucker, Alabama: U. S. Army Aeromedical Research Laboratory. USAARL Report 90-15.

Braybrook, R. (1998). Looks can kill. *Armed Forces Journal International*, 4, 44.

Brindle, J.H., Marano-Goyco, J., and Tihansky, P.M. (1995). Helmet-mounted sight/display system (HMS/DS) technology development status/vision. NAWCADWAR Advanced Helmet Vision System Development Program, Report No. 1.

Bull, G.C. (1990). Helmet display options-A Routemap. *Helmet- Mounted Displays II, Proceedings of SPIE*, 1290, 81-92.

Cameron, A.A., and Steward, D.G. (1994). The Viper HMD - from design concept to flight test. *Helmet- and Head-Mounted Displays and Symbology Design Requirements, Proceedings of SPIE*, 2218, 137-148.

Casey, C. (1991). Task training with a Helmet Integrated Display Simulator System. *Proceedings of SPIE*, 1456, 175-178.

Casey, C. (1999). Helmet mounted displays on the modern battlefield. *Helmet- and Head-Mounted Displays IV, Proceedings of SPIE*, 3689, 270-277.

Chapman, F.W., and Clarkson, G.N. (1992). The advent of helmet-mounted devices in the combat aircraft cockpit – An operator viewpoint. *Proceedings of SPIE*, 1695, 26-37.

Chisholm, P. (2006). 21st century helicopter training. *Military Training Technology* (Online Edition). 11(1).

Chung, J. (1989). Exploring virtual worlds with HMDs. TR89-009.

Davis, L. (2002). Application of Augmented Reality to Visualizing Anatomical Airways. *Proceedings of SPIE*, 4711, 400-405.

Dornheim, M. (1995). VTAS sight fielded, shelved in the 1970's. *Aviation Week and Space Technology*, October, 23, 51.

Drascic, D., and Milgram, P. (1996). Perceptual issues in augmented reality. *Proceedings of SPIE,* 2653, 123-134.

Duchowski, A. (2000). Binocular Eye Tracking in virtual reality for inspection training. *Proceedings of the Eye Tracking Research and Applications Symposium,* Palm Beach Gardens, FL, November 6-8, 89-96.

Dupont Corporation. (2004). Case Study: Cadillac XLR has new-generation Head-Up Display (HUD). Retrieved from http://www2.dupont.com/Automotive/en_US/applications/caseStudies/case178.html

Elbit Systems. (2004). MiDASH Modular Integrated Display and Sight System for helicopter pilots.

Foyle, D.C., McCann, R.S., Sanford, B.D., and Schwirzke, F.J. (1993). Attentional effects with superimposed symbology: Implications for head-up displays. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, Santa Monica, CA, 2, 1350–1354.

Geiselman, E.E. (1994). Utility of off-boresight helmet-mounted symbology during a high angle airborne target acquisition task. *Helmet- and Head-Mounted Displays and Symbology Design Requirements, Proceedings of SPIE*, 2218, 328-338.

Girolamo, H.J. (1997). Advanced Information Displays for the 21st Century Warrior. *Information Display,* 13(3), 10-17.

Gizmag.com. (2006). Trimersion HMD takes first person shooter games to a new level of reality. Retrieved October 31, 2007 from: http://www.gizmag.com/ go/6410/

Haar, R. (2005). Virtual reality in the military: Present and future. Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, the Netherlands.

Hale, S. and Piccione, D. (1989). Pilot assessment of the AH-64 helmet-mounted display system. *Proceedings of the 5th International Symposium on Aviation Psychology*, 1, 307-312.

Harding, T.H., Beasley, H.H., Martin, J.S., Rash, C.E., McLean, W.E., Mora, J.C., and McEntire, B.J. (1998). Optical and biodynamic evaluation of the Helmet Integrated Display Sight System (HIDSS) for the RAH-66

Comanche development and validation program phase. Fort Rucker, Alabama: U. S. Army Aeromedical Research Laboratory. USAARL Report No. 98-22.

Harding, T.H., Martin, J.S., Beasley, H.H., and Rash, C.E. (2004). Evaluation of the Microvision helmet-mounted display technology for synthetic vision application engineering prototype for the Virtual Cockpit Optimization Program. Fort Rucker, Alabama: U. S. Army Aeromedical Research Laboratory. USAARL Report No. 2004-02.

Harris, D., and Muir, H. (Editors). (2005). *Contemporary issues in human factors and aviation Safety*. Burlington, Vermont: Ashgate Publishing.

Hiatt, K.L., Rash, C.E., Harris, E.S., and McGilberry, W.H. (2004). Apache aviator visual experiences with the IHADSS helmet-mounted display in Operation Iraqi Freedom. Fort Rucker, Alabama: U. S. Army Aeromedical Research Laboratory. USAARL Report No. 2004-18.

Heinecke, J.K., Rash, C.E., Ranaudo, R.J., and Hiatt, K.L. (2008). Apache aviator evaluation of dual technology night vision systems in Operation Iraqi Freedom urban combat. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2007-05.

Hopper, D.G. (2000). Reality and Surreality of 3-D Displays: Holodeck and beyond. Invited Keynote Paper in the *Electronic Information Display Conference,* Society for Information Display ExCel, England U.K., November 21-23, 2000.

Hughes, D. (1994). Off-boresight missile test highlights U.S. Capability. *Aviation Week and Space Technology*, 6, 66-69.

Jennings, S. (1997). An in-flight evaluation of stereoscopic imagery in helicopter cockpits. *SPIE Aerosense*, *Proceedings of SPIE,* 3057, 108-117.

Johnston, R.S. and Willey, S.R. (1995). Development of a commercial retinal scanning display. *Helmet- and Head-Mounted Displays and Symbology Design Requirements II, Proceedings of SPIE,* 2465, 2-13.

Kalich, M.E., Rash, C.E., McLean, W.E., and Ramiccio, J.G. (2007). A limited flight study for investigating hyperstereo vision. *Head- and Helmet-Mounted Displays XII: Design and Applications*. In: Brown, R.W., Reese, C.E., Marasco, P.L., and Harding, T.H. (Eds.), *Proceedings of SPIE,* 6557, 65570I.

Kaiser Electro-Optics, Inc. (2007). *XL Series Head-mounted displays*.

Kauchak, M. (2001). AVCATT-A enters low-rate initial production. *T&S 2001, An Armed Forces Journal Global Publication*, Winter, 56-57.

Keller, K., and Colucci, D. (1998). Perception in HMDs: What is it in head-mounted displays (HMDs) that really make them all so terrible? *Proceedings of SPIE,* 3362, 46–53.

Klass, P.J. (1972). Navy pilots to use helmet sight. *Aviation Week and Space Technology*, January 31, 37-40.

Lucas, T. (1994). Advances in helmet-mounted displays. *Avionics*, June. pp. 20-27.

Manning, S.D. and Rash, C.E. (2007). A moveable view. *AeroSafetyWorld*. Flight Safety Foundation. 2, 41-44.

Marshall, G.F. (1989). Back from the past. helmet-mounted display, *Proceedings of SPIE*, 1116, 2-11.

McCann, R.S., Foyle, D.C., and Johnson, J.C. (1993). Attentional limitations with head-up displays. In: Jenson, R.S. (Ed.), *Proceedings of the 7th International Symposium on Aviation Psychology*. Columbus, Ohio: Ohio State University, 70-75.

McCann, R.S. and Foyle, D.C. (1995). Scene-linked symbology to improve situation awareness. *AGARD Conference Proceedings,* No. 555, Aerospace Medical Panel Conference on Situation Awareness, 16:1-16:11.

McLean, W.E., Rash, C.E., McEntire, J., Braithwaite, M.G., and Mora, J.C. (1998). A performance history of AN/PVS-5 and ANVIS image intensification systems in U.S. Army aviation. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 98-28.

Melzer, J.E., and Larkin, E.W. (1987). An integrated approach to helmet display system design. *Display System Optics*, *Proceedings of SPIE*, 778, 83-88.

Melzer, J.E., and Moffitt, K. (1997). *Head-Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Merryman, R.K. (1994). Vista Sabre II: Integration of helmet-mounted tracker/ display and high boresight missile seeker into F-15 aircraft”. *Proceedings of SPIE*, 2218, 173-184.

Microvision, Inc. (2005). Specification sheet: SPECTRUM SD2500.

Milgram, P. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. *Proceedings of SPIE*, 2351, 282-292.

Moore, J., Dennis, S., Thome, M., and Scalf, W.D. (1999). The use of Simulation and Modeling for Acquisition, Requirements, and Training (Smart) in the Virtual Cockpit Optimization Program (VCOP). The Interservice/Industry Training, Simulation and Education Conference (I/ITSEC).

Nelson, S. (1994). Combat vehicle crew helmet-mounted display: next generation high-resolution head-mounted display. In: Lewandowski, R., Stephens, W., and Haworth, L. (Ed.), *Helmet- and Head-Mounted Displays and Symbology Design Requirements, Proceedings SPIE,* 2218, 7-16.

Nelson, A.S., and Helgeson, M. (1996). High resolution color flat panel head-mounted display design for the advanced flat panel program. *Proceedings of SPIE*, 2735, 124-133.

Nelson, S., Jones, C., Pierre, P., and Sampson, J. (1997). Arthrooscopic knee surgery using the advanced flat panel high resolution color head-mounted display. *Proceedings of SPIE*, 3058, 170-177.

Primeau, G. (2000). Wide field-of-view SVGA sequential color HMD for use in anthropometric telepresence application. *Proceedings of SPIE*, 4711, 11-20.

Proctor, P. (1996). Retinal displays portend synthetic vision, HUD advances. *Aviation Week and Space Technology*, 15, 58.

Qt Auto News. (2006). HUD technology keeps your eye on the road. *Auto News*, Retrieved from http://www.qtautonews.com/2006/04/25/hud-technology-keeps-your-eye-on-the-road/

Rash, C.E. (Editor) (2000). *Helmet-mounted displays: Design issues for the rotary-wing environment*. Bellingham, WA: SPIE Press.

Rash, C.E. (2006a). Shutting out the noise. *AviationSafetyWorld,* 1 (November), 42-45.

Rash, C.E. (2006b). Helmet-mounted displays. *Special Operations Technology*, 4(5), 6-8.

Rash, C. E. (2008). A 25-year retrospective review of visual complaints and illusions associated with a monocular helmet-mounted display. *Health and Safety Aspects of Visual Displays*, *Displays*, 29(2), 70-80.

Reinhart, W. (1991). Depth cueing for visual search and positioning. *Proceedings of SPIE*, 1457, 221-232.

Rockwell Collins. (2006). Rockwell Collins Sim-Eye 100. Retrieved February 28, 2008 from ProductsAndServices/VisionSystems/HeadMountedDisplays/SIM-EYE-SR100/SimeEye_SR100_071706.pdf

Rockwell Collins. (2008a). EyeHUD Day/Night helmet-mounted display system. Retrieved February 20, 2008 from http://www.rockwellcollins.com/ecat/gs/EyeHUD.html

Rockwell Collins. (2008b). Retrieved February 10, 2008 from http://www.rockwellcollins.com/products/ gov/defense_electronics/ displays1/index.html

Rolland, J. (2002). Head-mounted projective displays for creating distributed collaborative environments. *Proceedings of SPIE*, 4711, 395-399.

Seeman, J., De Maio, J., Justice, S., Wasson, J., Derenski, P., Hunter, M., and Walrath, L. (1992). Advanced helicopter pilotage visual requirements. *Proceedings of the American Helicopter Society*, 48, 233-252.

Shontz, W.D., and Trumm, G.A. (1969). Perceptual processes and current helmet-mounted display concepts. Minneapolis, Minnesota: Honeywell Inc. Technical Note TN-1.

Simons, R., Schaefer, J., and Melzer, J. (2002). Visualization solutions for AVCATT-A reconfigurable cockpits. The Interservice/Industry Training, Simulation and Education Conference (I/ITSEC), Orlando, FL.

Sisodia, A, Riser, A., Bayer, M., and Mcquire, J.P. . (2007). Advanced helmet mounted display (AHMD) for simulator applications. Retrieved from http://www.link.com/pdfs/AHMD_2006_Paper.pdf

SPG Media. (2008). Land Warrior Integrated Modular Fighting System, USA. Retrieved February 28, 2008 from http://www.army-technology.com/projects/land_warrior/

Sutherland, I.V. (1965). The ultimate display. *Proceeding of the IFIP Congress*, 2, 506-508.

Vora, J. (2002). Using virtual reality technology for aircraft visual inspection training: presence and comparison studies. *Applied Ergonomics*, 33, 559–570.

Vos, G. D., and Brandt, G. (1990). Use of holographic optical elements in HMDs. *Helmet Mounted Displays II, Proceedings of SPIE*, 1290, 70-80.

White, J., and Cameron, A. (2001). Knighthelm 24-hour HMD. *Helmet Mounted Displays VI, Proceedings of SPIE*, 4361, 164-175.

Wickens, C.D., and Long, J.L. (1994). Conformal symbology, attention shifts, and the head-up display. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting*, Human Factors and Ergonomics Society. Santa Monica, California: Human Factors Society, 2, 1350–1354.

Wood, R. B. (1992). Holographic head-up displays. *Electro-optical displays*, In: Karim, M. (Ed.). New York: Marcel Dekker. 337-415.

Yona, Z., Weiser, B., and Hamburger, O. (2004). Day/night ANVIS/HUD-24 (Day HUD) flight test and pilot evaluations. *Helmet- and Head-Mounted Displays IX: Technologies and Applications*. In: Rash, C.E., and Reese, C.E., (Editors), *Proceedings of the SPIE*, 5442, 225-235.

Zintsmaster, L.R. (1994). Wide-Eye^{TM} helmet mounted display system for rotorcraft applications. *IEEE Aerospace and Electronic Systems Magazine*. 9(3), 6-11.

4 VISUAL HELMET-MOUNTED DISPLAYS

Clarence E. Rash
Michael M. Bayer
Thomas H. Harding
William E. McLean

The first helmet-mounted displays (HMDs) were purely visual systems. This includes the original (but not fielded) Pratt gun sight (Pratt, 1916) (Figure 4-1, top, left) and the first image intensification (I^2) devices, Night Vision Goggles (NVGs). All NVG systems, even the most current design, the Aviator's Night Vision Imaging System (ANVIS) (Figure 4-1, top, right), are add-on devices, in the sense that they are not integrated into their helmet platform but are attached to the helmet.

All currently fielded HMDs provide visual input. Integrated HMDs, such as the 1970's Honeywell, Inc., Integrated Helmet and Display Sighting System (IHADSS) (Figure 4-1, bottom), used on the U.S. Army's AH-64 Apache helicopter incorporate both visual and audio inputs within the helmet platform. Integrated HMD designs attempt to optimize optical and acoustical performance while maintaining the protective function of the helmet. In addition, the helmet must serve as the mounting platform for the optical and acoustical elements.

This chapter introduces the fundamental concepts of HMDs from the perspective of optical design and image quality as they affect the Warfighter's visual performance. The auditory concepts of HMD designs are presented and discussed in Chapter 5, *Audio Helmet-Mounted Displays*.

A discussion of visual HMDs begins with an overview of the different approaches used in the optical design of HMDs. The most important components are the image sources and the optics that deliver the image generated by the source to the user's eye(s). Probably the most important element within the relay optics is the final reflecting surface. For see-through HMDs, this element serves as a beamsplitter.

User acceptance of and performance with are the critical measures of the success of any HMD. Acceptance depends on many factors from the fields of ergonomics and human factors. HMD parameters that impact acceptance include head-supported weight, center-of-mass (CM) offsets, fitting method, exit pupil size and physical eye relief.

User performance also is strongly correlated with the quality of the display imagery presented to the eye. Image quality is determined by a number of factors, which include luminance, contrast, resolution, ambient illuminance, and uniformity. Such factors, referred to as figures of merit (FOMs), used to indicate image quality, depend on the type of image source, e.g., cathode-ray-tube (CRT), plasma, and liquid crystal display (LCD). The level of image quality in an HMD will determine the user's ability to recognize and interpret the information content in the presented image.

Optical Designs

The optical design for any HMD has as its primary purpose the generation of a final image(s) that is then viewed by the eye(s). In all HMD designs, the image source is located some distance away from the eye(s). If this initial image at the image source is sufficiently far away it must be relayed up to the eyepiece optics, which form the final image(s) for the eye(s). In performing this task, the optical system must provide a specific field-of-view (FOV) to the viewer with sufficient eye clearance to accommodate spectacles, protective masks, and other

Figure 4-1. The Pratt gun sight (top, left); the Aviator's Night Vision Imaging System (ANVIS) (top, right); and Integrated Helmet and Display Sighting System (IHADSS) (bottom).

possible required add-on devices. The optical design must create a sufficiently sized eye box (a volume in space where the viewer's eye must be placed) to compensate for pupil displacements due to eye movement, vibration, and head/helmet slippage. For optical systems that use relay optics this eye box is called the exit pupil. Optical systems that do not use relay optics also have a designed eye position location, or eye box, that is often erroneously called an exit pupil. For optical systems that produce a real exit pupil eye movement outside of the exit pupil will result in an inability to see any part of the FOV, whereas for non-real exit pupil systems (those without relay optics) movement outside of the eye box may result in losing part of the FOV and/or in reduced image quality (blur).

Optical design parameters

There are a number of important descriptive parameters in an HMD optical design. These include:

- Field-of-view (FOV)
- Exit pupil (eye box) size and shape
- Optical eye relief
- Physical eye relief
- Transmission (optical throughput)
- Beamsplitter transmission/reflection coefficients (for see-through HMDs)
- Modulation transfer function (MTF)
- Chromatic aberration
- Distortion

- Field curvature
- Magnification
- Ghosting
- Weight (Mass)
- Center-of-mass (CM)
- Volume (Space required)

While it is tempting to identify a select few of these parameters as being universally most important, the intended use of the HMD is, in fact, the deciding factor in which parameters should drive the optical design. As an example, an HMD that has targeting as its sole purpose would require a very small FOV (e.g., 1 to 3 degrees), making FOV less of a design driver. This is in contrast to an HMD that has pilotage imagery as it primary use. In this case, a large FOV is desired, making it an important design parameter for its purpose.

Nonetheless, there are a few optical system parameters that are fundamentally important to the vast majority of designs and deserve brief discussions. These include weight (mass), FOV, MTF, exit pupil size, and eye relief.

The weight (mass) of the optics includes contributions from the optical elements themselves (e.g., lenses, beamsplitter, mirrors, prisms), the housing for these optical elements, and in most cases, the image source. The choice of the material used for the optical elements can impact the optics weight significantly. Although considerable advancement has been made in optical materials, the best image quality currently available is still obtained with optical elements composed of glass. Unfortunately, glass is the heaviest optical medium. Nonetheless, compromises via the use of plastic optical elements, which are both lighter in weight and lower in cost, have been made. Holographic elements offer even more weight savings. The use of holographic beamsplitters (combiners) in refractive optics HMD optical designs makes use of their wavelength-selective characteristics and has the added advantage of not introducing additional optical power (Wood, 1992).

The weight (mass) associated with the optics is important from both ergonomic and safety perspectives. The additional head-supported weight (mass) of the HMD can produce neck muscle fatigue, which can degrade performance, and increase the potential of injury due to dynamic loading during crashes. It is desirable to minimize head-supported weight (mass) in HMD designs. The optics and image source make up a significant portion of this weight (mass).

By the very design of current HMDs, some of the optical components (and hence the additional weight) are located in front of the face. This results in the CM of the system being forward and often above the CM of the human head/neck combination (i.e., the tragion notch). In monocular HMDs, the system CM also will be offset further, laterally. This resulting torque increases neck muscle fatigue. The issues associated with head-supported weight (mass) and CM are fully discussed in Chapter 17, *Guidelines for HMD Designs*.

Another fundamental optical parameter is FOV, defined as the maximum angle of view that can be seen through an optical device. An alternative definition is the horizontal and vertical angles the display image subtends with respect to the eye. This definition is the result of most HMD FOVs being rectangular and described as a combination of the vertical angle and the horizontal angle (e.g., the IHADSS FOV is cited as 30º vertical X 40º horizontal).

FOV is affected by magnification and the image source size, with greater magnification and/or image source size resulting in a larger field of view. Typically, HMDs present a FOV to the viewer that matches one-to-one (conformally) with the FOV of the sensor that is used to capture the original image of the outside world. In principle, the larger the FOV, the greater the amount of information made available (assuming the image source and sensor have the resolution to properly support the increased FOV). Consequently, HMDs designed for pilotage attempt to maximize FOV, ideally matching that of the human visual system. The human eye has an instantaneous FOV that is roughly oval and typically measures 120º vertically by 150º horizontally. Considering both eyes together, the overall binocular FOV measures approximately 120º (V) by 200º (H) (Zuckerman, 1954) (Figure 4-2).

Designs fielded so far all provide restricted FOV sizes compared to human vision. The size of the FOV that an HMD is capable of providing is constrained by several sensor and display parameters, which include size, weight, placement, and resolution.

In ANVIS, the FOV of a single image tube is nominally a circular 40º. The two tubes have a 100 percent overlap; hence, the total FOV is also 40º. This FOV size seems small in comparison to that of the unobstructed eye. But, the reduction must be judged in the context of all of the obstructions associated with a cockpit, e.g., armor, glare shield, and support structures. The monocular IHADSS used on the AH-64 Apache helicopter has a rectangular FOV, 30º vertical X 40º horizontal. Biocular HMD designs, such as the U.S. Army's Comanche program that is no longer in development, had a 35º vertical X 52º horizontal FOV.

The design parameter most affected by the choice of material for the optical elements is the MTF. The MTF is a metric that defines how well an optical system transfers modulation contrast from its input to its output as a function of spatial frequency.[1] A plot of such a transfer is called an MTF curve (Figure 4-3). Since any scene theoretically can be resolved into a set of sinusoidal spatial frequencies, it is possible to use a system's MTF to determine image degradation through the system.

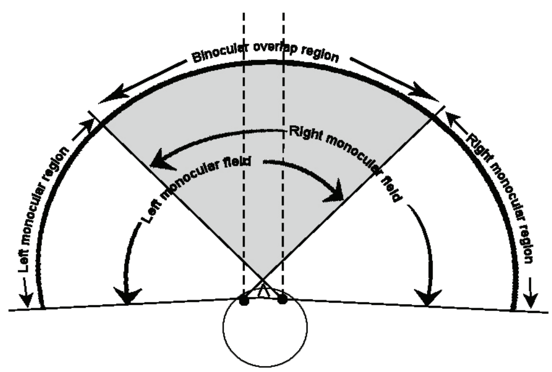


Figure 4-2. Human visual system's binocular field-of-view (FOV).

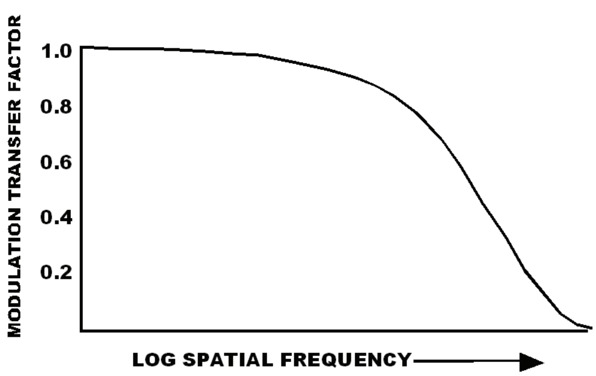


Figure 4-3. Typical modulation transfer (MTF) curve.

[1] Spatial frequency is a measure of detail in a scene, usually defined by how rapidly luminance changes within a region. A single spatial frequency is commonly represented by a series of vertical bars where the luminance varies according to a sinusoidal function. In this simple case the spatial frequency of the stimulus is just the frequency of the sinusoid used to generate the pattern. In general, the part of a scene with fine detail including sharp edges has high spatial frequencies and the part where the luminance over a region changes more slowly has low spatial frequencies.

Within an HMD system, every major component (e.g. sensor, image source, optics) has its own MTF. If the system is linear, its total MTF can be obtained by multiplying the MTFs of the system's individual components. The illustrative MTF curve provided in Figure 4-3 presents a relatively good contrast transfer for low and medium spatial frequencies (the curve is high on the vertical axis) but falls rather abruptly at higher frequencies. A system's inability to faithfully reproduce contrast at the higher spatial frequencies would indicate a loss of a user's ability to see detailed features in the environment.

To accurately predict the image quality of an HMD system, it is necessary to determine how the overall system will affect resolution and contrast. The MTF performs this function. The MTF of an optical system is perhaps the most widely-accepted metric for the quality of the imagery seen through the optical system (Velger, 1998). It defines the fidelity to which an outside scene is reproduced in the final viewed image. A perfect system would have an MTF of unity across all spatial frequencies (Shott, 1997). The degradation that is present in a practical HMD optical system's MTF is a result of the residual (uncorrected) aberrations in the system and is ultimately limited by diffraction effects, which is beyond the scope of this section.

The remaining two design parameters needing some explanation, exit pupil and eye relief, are closely related. The exit pupil is the volume in space where the eye must be placed in order to be able to see the full image. An exit pupil has three characteristics: size, shape, and location. Within the limitation of other design constraints, e.g., size, weight, complexity, and cost, the exit pupil should be as large as possible.

The 1970s IHADSS has a circular 10-mm diameter exit pupil. The planned HIDSS exit pupil was specified also to be circular but with a larger, 15-mm, diameter. While systems with exit pupils having diameters as large as 20 mm have been built, 10 to 15 mm has been the typical value (Task, Kocian, and Brindle, 1980). Tsou (1993) suggests that the minimum exit pupil size should include the eye pupil (~ 3 mm), an allowance for eye movements that scan across the FOV (~ 5 mm), and an allowance for helmet slippage (± 3 mm). This would set a minimum exit pupil diameter of 14 mm. Since the real exit pupil is the image of an aperture stop[2] in the optical system, the shape of the exit pupil is generally circular (assuming the aperture stop is circular) and, therefore, its size is expressed as a diameter.

The exit pupil is located at a distance called the optical eye relief, which is defined as the distance from the last optical element to the exit pupil (Figure 4-4). Over the years, this term has caused some confusion within the HMD community (Rash et al., 2002). What is of critical importance in HMDs is the actual physical distance from the plane of the last physical element to the exit pupil, a distance called the <u>physical eye relief</u> or <u>eye clearance distance</u> (Figure 4-4). This distance should be sufficient to allow use of corrective spectacles, nuclear, biological and chemical (NBC) protective masks, and oxygen mask, as well as, to accommodate the wide variations in head and facial anthropometry. This ability to accommodate intervening visual devices has been a continuous problem with the IHADSS, where the optical eye relief value (10 mm) is greater than the actual eye clearance distance. This is due to the required diameter of the relay optics' objective lens and the bulk of the barrel housing.

To overcome the incompatibility of spectacles with the small physical eye relief of the IHADSS, the U.S. Army investigated the use of contact lenses as an approach to provide refractive correction (Bachman, 1988; Lattimore, 1990; Lattimore and Cornum, 1992). While citing a number of physiological, biochemical and clinical issues associated with contact wear and the lack of reliable bifocal capability, the studies did conclude that contact lenses may provide a partial solution to HMD eye relief problems. Contacts have indeed provided and continue to provide the capability of vision correction for AH-64 Apache pilots. More recently, following the lead of the U.S. Air Force, the U.S. Army conducted a study that investigated refractive surgery techniques as an alternative solution (van de Pol et al., 2007). As a result of this study, a policy has been issued allowing the surgical procedure of Laser-Assisted in Situ Keratomileusis (LASIK).

---

[2] In optics, an aperture in an optical system is a structure or opening that limits the light rays that pass through the system. An optical system usually has several such apertures. In general, these structures are called *stops*, and the *aperture stop* is the stop that determines the ray cone angle, and equivalently the brightness, at an image point.

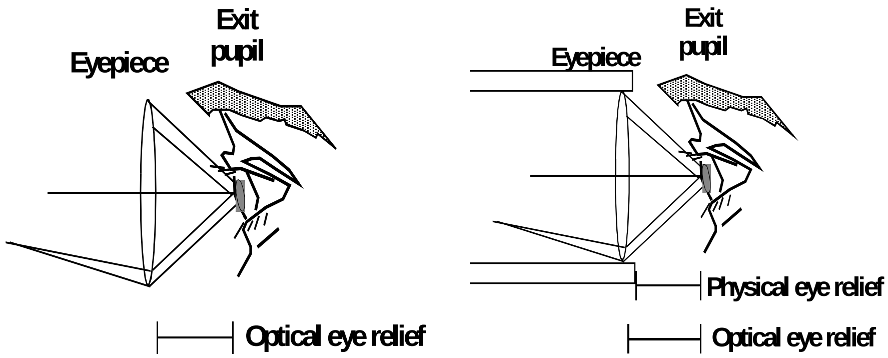


Figure 4-4. Optical eye relief (left) is defined as the distance from the last optical element to the exit pupil, where the eye would be placed. Physical eye relief (right) can be less than optical eye relief if additional structures are present (Rash et al., 2002).

HMD designs

The range of HMD design types run from a simple projection of symbolic and/or alpha/numeric information overlaying a direct-view real image for day time use, to a head slaved virtual imaging device that could be linked to remote sensors and/or computer generated imagery for day or night use, and of course, anything in between. As the design type increases in complexity, so does the optical design.

Simplest HMD design

In the Vietnam Era, a Bell Cobra helicopter (AH-1) was developed with a simple monocular helmet sight (known as the Cobra sight) that could translate an external mounted machinegun using a mechanical head tracker that attached to the top of the helmet (Braybrook, 1998). In front of the right eye was a small semi-transparent window that projected a red dot that was similar to simple commercial red dot reflex sights on some pistols and rifles. The 17-millimeter (mm) diameter combiner was located outside the helmet visor about 50 mm from the eye and could be adjusted in vertical and horizontal positions to properly align with the right eye. The size of the projected red dot was only a few milliradians (mr) in diameter, and was focused at infinity. The see-through visible transmission of the combining glass (beamsplitter) was very high, and the brightness of the aiming reticule was sufficient to be visible at the sky horizon.

Complex modern HMD designs

In contrast to the simple design of the Cobra sight is a limited-fielded visor-projection HMD currently being offered by a leading aerospace company having the following characteristics:

- Visor projection optical design
- Focused and aligned at infinity
- Binocular/biocular viewing
- Magnetic or electro-optical head tracked
- See-through vision
- FOV – >40°
- Can accommodate a wide range of eye separations

- Sufficient brightness and contrast for both day and night operations
- Incorporates direct-view image intensifiers on the side of the helmet and video links to external sensors reflected from the visor
- Flight and/or systems symbology can be projected, with unaided daytime see-through vision or at night, overlaying the image intensified or thermal image

The optical design of such a visor-projection HMD has always been challenging in obtaining a wide FOV with low-distortion, high-quality imagery and with an acceptable head supported weight. The reflective component of the visor design may be a hologram or a dichroic filter imbedded in the visor that focuses and aligns the incident rays from the relay optics. This example HMD is binocular in its optical design for thermal imagery, but since there is a single infrared sensor, the thermal image is repeated to both eyes, which results in a biocular HMD system. For the I^2 imagery, with the I^2 tubes located on the side of the helmet and further apart than the normal separation between the eyes, intensified image is truly binocular. However, this binocularity produces the visual state of hyperstereopsis, which is fully discussed in Chapter 12, *Visual Perceptual Conflicts and Illusions*. Because the near-infrared image intensifiers are located on the head, and the infrared thermal sensors are located on the outside of the aircraft, the operator can only use one or the other, since the two images can not be fused. This requires the operator to mentally move their visual location for proper perspective cues. Another challenging characteristic will be switching from the biocular thermal sensor, with no stereopsis, to a hyperstereo I^2 scene.

See-through vision of the visor would intuitively seem to be desirable. For day flight with the HMD providing flight and aircraft/weapon information (symbology), undistorted, high transmission see-through vision greatly increases the pilot's situation awareness and effectiveness. Symbology for helicopter and near-earth day-viewing must be monocular to prevent double images. Binocular symbology can only be seen single, and the outside images appear single as well, when objects are located beyond 60 meters (197 feet) (McLean and Smith, 1987). In addition, the right and left images have to be aligned both vertically and horizontally at infinity to within 1 mr. When viewing closer than 60 meters (197 feet), the difference in the eye convergence between the symbol and the outside image will exceed 1 mr and induce diplopia (double vision). An exception to using only monocular symbology could be an aiming reticle or test pattern to check the HMD for proper alignment between the right and left eye images before flight, and then switching to monocular symbology for the day mode.

At what distance should the symbology and projected image be focused? For head-up displays (HUDs), commonly used in fixed-wing fighter aircraft and are viewed binocularly, the focus and alignment would be expected to be set within 1 mr of infinity to correspond to distant outside objects. However, with thick curved canopies, such as the F-16, the alignment of the actual object and viewed symbology or image through the canopy can be slightly different, and the HUD focus and convergence are adjusted to coincide to the image shift caused by the canopy. When the HUD was initially set at infinity alignment, the symbology appeared double when viewed by an observer focusing on a distant object. In other words, the viewed image through the canopy may not be optically aligned but may appear to be at a distance other than its actual physical distance (Martin et al., 1983). However, when viewing a sensor image, whether with image intensifiers or a binocular/biocular HMD, the eyepiece infinity focus and alignment may induce slightly blurred images for many of the pilots that are very slightly myopic and not therefore required to wear corrective lenses.

If an HMD (such as the visor projection type) has a final, beamsplitter reflective element, it may induce ghost images or optical artifacts that are not desirable, compared to a standard helmet visor. One would think that having simultaneous, overlaid, unaided vision and sensor images would provide the best of both perceptions, but in almost all cases, the users are aware the two separate images (unaided and aided) never exactly align within the 1 mr tolerance, and the two images create a conflict. It's similar to the Sunday paper where the three colors do not align in a picture. The see-through vision for night imagery is easily blocked with an added opaque visor that only covers the FOV equal to the size of the sensor image. When pilots were given the option of blocking the outside see through image at night with the opaque visor, almost all preferred the non-see-through format.

Design types

There are a number of HMD optical design types that have been deployed over the decades of HMD development. Most HMD optical design types require an eyepiece to allow the user to see the HMD imagery. Figures 4-5 and 4-6 show the ray trace differences between various simplified eyepiece designs. For comparison purposes, the drawings of each eyepiece type design presented are equally scaled. The full-scaled drawings used 30-mm eye clearances and 5-mm exit pupils to obtain a vertical FOV of 40º.

The following descriptions encompass the more fundamental optical HMD optical design approaches and are only representative of the many varied designs that have been implemented. A number of extensive reviews of HMD optical designs are suggested for the more interested reader (e.g., Cakmakci, O. and Rolland, J., 2006; Melzer, J., and Moffitt, K., 1997; Velger, M., 1998).

Refractive

The simplest NVG, HUD, and HMD systems use refractive, on-axis eyepiece optics. Examples include the ANVIS (Figure 4-5, top) with no see-through vision and a reflex HUD (Figure 4-5, bottom) with a 45º angle combiner and see-through vision. The see-through vision is provided with a partial reflective beam splitter or plano combiner. IHADSS helmet display unit (HDU) (Figure 4-6, top), which is an HMD with see-through vision in the AH-64 aircraft for night pilotage, tilts the combiner to 38º from the last optical lens to improve eye relief. Refractive optical designs use lenses for imaging. The IHADSS HDU provides imagery and symbology from remote sensors, where the two night imaging sensors (I^2 tubes) are contained in the ANVIS. The primary advantage of the refractive design with a plano combiner is the high percent luminance transfer from the display to the eye. The primary disadvantages for refractive HMDs with see-through vision are excessive weight with limited fields of view and eye clearance.

The ANVIS eyepiece is a simple, well-corrected, magnifier with no see-through vision. Other NVG designs such as the Eagle Eye™ or the Cat's Eyes™ use prism combiners for see-through vision with I^2, but the see-through combiners with intensifier tubes have been used primarily by fixed-wing fighter type aircraft with HUDs. These see-through plano combiners are enclosed or sandwiched between two prisms which, when combined, form a plano refractive media with minimal prismatic deviation. The purpose of the prism combiners is to increase the combiner stability and increase the eye clearances for a given FOV and eyepiece diameter. Figure 4-6 (bottom) shows a prism combiner using the IHADSS design. The prism combiners can also be used with power reflective combiners. Figure 4-7 (top) shows a catadioptric eyepiece design without the prism combiner and Figure 4-7 (bottom) with a prism combiner.

Catadioptric optical designs use curved reflective mirrors with or without lenses for imaging (Figure 4-7). The primary advantage of catadioptric designs is larger diameter optics with less weight and without induced chromatic aberrations. By coating transmissive curved surfaces with partial reflective materials to provide see-through vision, the beam splitter is referred to as a power combiner. Figure 4-7 (top) shows the catadioptric design with a prism combiner to increase the eye clearance for a given FOV. The primary disadvantages are reduced luminance transfer with prism combiner from the display for a given percent see-through vision compared to refractive systems. Extraneous reflections have also been a problem area. The catadioptric designs can obtain slightly larger fields of view for a given eye clearance compared to refractive systems. Catadioptric designs have not been used in significant numbers for production HMDs at present, but have been used in a few HUDs (example OH-58D pilot display unit (PDU) for Stinger missiles).

Figure 4-8 shows comparison plots of the eyepiece diameters versus FOV for the refractive nonsee-through versus the various see-through HMD designs without prism combiners. The differences between the refractive and IHADSS HMDs are only in the angle of the combiner to the eyepiece and central ray to the eye. The refractive see-through HMD (Figure 4-5, bottom) uses a constant 45º combiner angle for all FOVs, where the IHADSS HMD (Figure 4-6, top) adjusts the lower FOV limit ray to run parallel with the eyepiece to minimize its

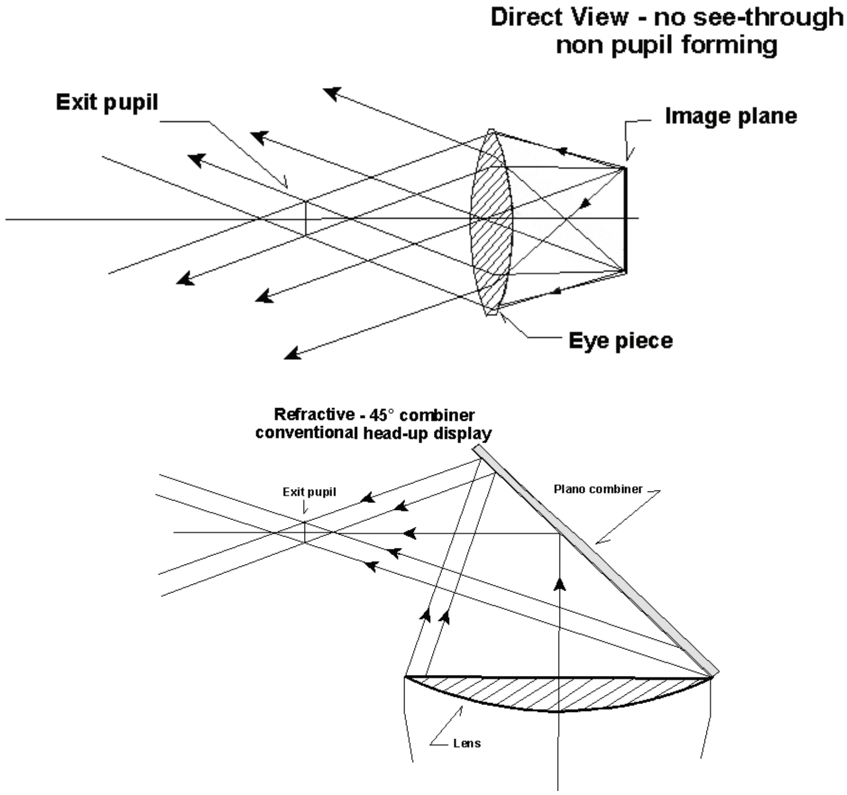


Figure 4-5. HMD eyepieces: Direct-view, no see-through, NVG type eyepiece (top) and HUD refractive see-through combiner at 45º (bottom) (Rash, 2001).

diameter. The estimated 60-mm diameter eyepiece limit is based on mechanical considerations for the smaller IPD ranges and overlapped HMD FOVs.

Catadioptric

Figure 4-9 graphs and compares the effects on the eyepiece diameter with and without prism combiners for the IHADSS and catadioptric designs. A high index of refraction (n = 1.58) plastic material (polycarbonate) was selected for the prism combiners for calculation purposes to obtain the maximum effect. Other materials could be selected for the prism combiners for the particular properties of the material such as lower weight and manufacturing qualities. Note that the surfaces closest and farthest from the eye of the prism combiners are parallel surfaces for the see-through vision. Without parallel surfaces, unwanted prismatic deviations or refractive powers would be induced. The prism combiner is actually more like a cube beam splitter, except the alignment of the beamsplitter does not have to be 45º to the central ray.

On- and off-axis designs

On-axis optical designs align the optical centers of each optical element, or slightly displace one of the elements, which can be rotated to achieve vertical and horizontal alignment for binocular designs such as

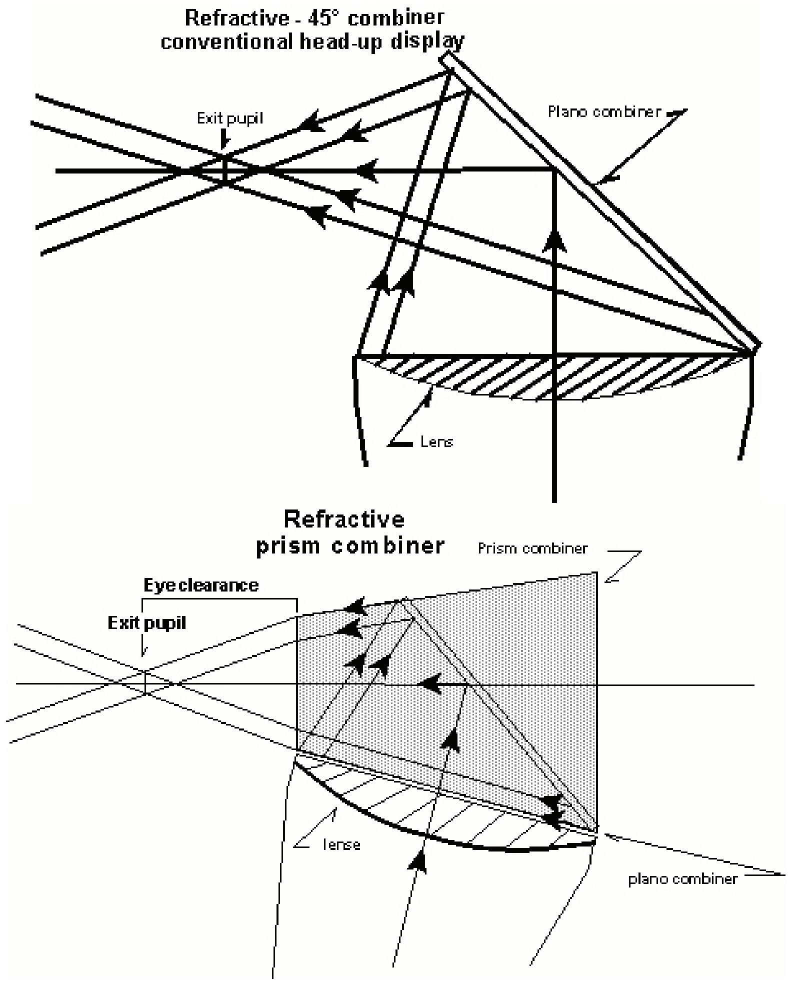


Figure 4-6. HMD eyepieces: Refractive (IHADSS) (top) and refractive prism combiner (bottom) (Rash, 2001).

binoculars. The IHADSS and the ANVIS refractive designs use on-axis alignment. The on-axis, see-through catadioptric designs include power and plano combiners. Off-axis catadioptric systems are usually referred to as reflective off-axis systems and may or may not require plano combiners. As the off-axis angle to the power combiner increases, the induced distortions and aberrations increase rapidly (Buchroeder, 1987). An example of a modest off-axis catadioptric design with a plano combiner is shown in Figure 4-10 (Droessler and Rotier, 1989; Rotier, 1989). This catadioptric design achieves a 50º x 60º FOV with a 10-mm exit pupil and 30-mm eye relief (measured from plano combiner intercept to apex of eye along primary line of sight). However, note the optical complexity with 11 refractive elements and 3 reflective surfaces with very complex coatings for both eyepiece reflective surfaces to maximize see-through and display transmissions. The modest trapezoidal distortion of 7.5% (Figure 4-11) will be aligned with the power combiner. Another promising HMD is the Monolithic Afocal Relay Combiner (MONARC), which is an off-axis, rotationally symmetrical lens system with modest FOV potential, but excellent see-through approach (Figure 4-12). However, for any of the off-axis binocular systems, the distortions will have to be corrected to achieve point for point image alignment throughout the FOV.

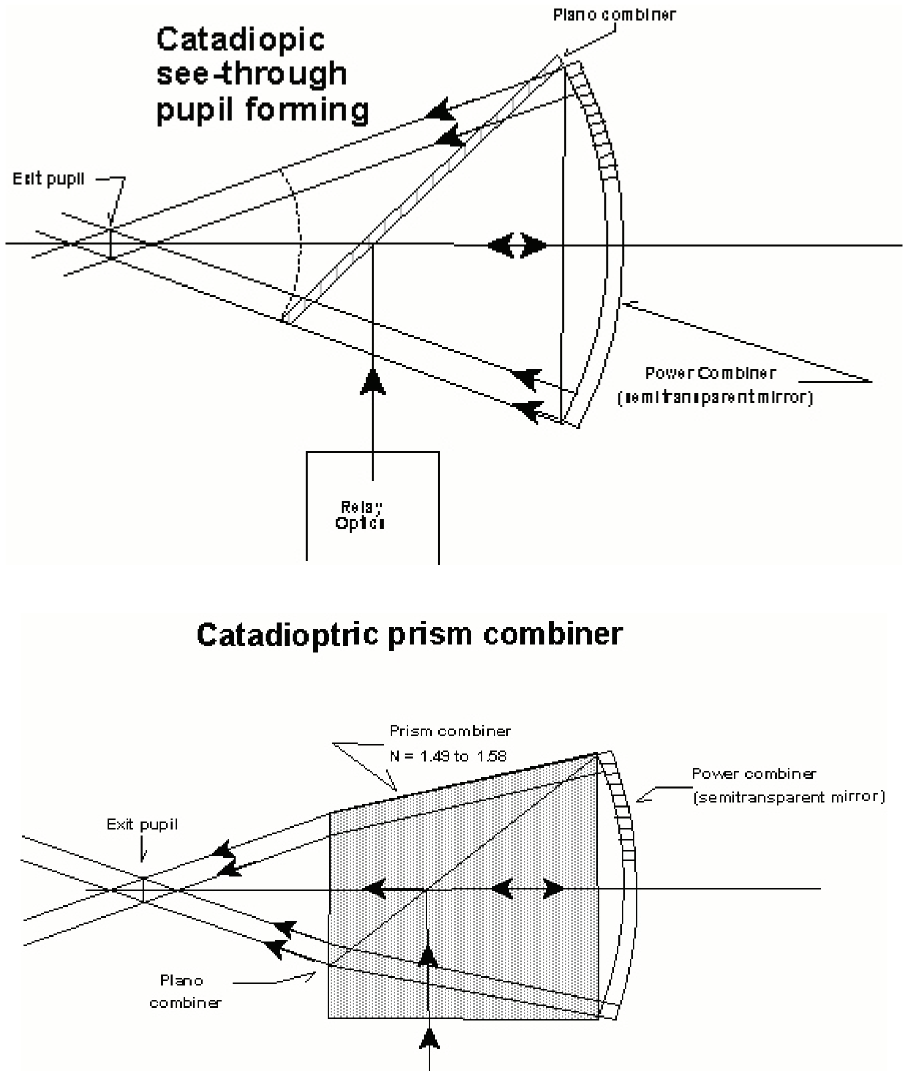


Figure 4-7. HMD eyepieces: Catadioptric (top) and catadioptric with prism combiner (bottom) (Rash, 2001).

The primary advantage of the off-axis reflective HMD design is that it provides the highest potential percent luminance transfer from the display with the most see-through vision and increased eye clearances for a given FOV. The primary disadvantages are very complex optical designs, shape distortions, and low structural integrity and stability of the reflective surface. Figure 4-13 shows the conceptual drawings (top and side view) of an off-axis HMD using the visor as the eyepiece. Note the locations of the aerial images, which are shown for the left eye. The location of the relay optics will be either on top of the helmet, or below and to the sides, where both locations have undesirable characteristics such as a high center of mass, or produce lower obstructions to unaided vision. Also, note that the head seems to get in the way of the optics or relay image. Where there are no provisions for electronic distortion correction, such as found with NVGs, the off-axis designs become unacceptable from the keystone or trapezoidal type distortions.

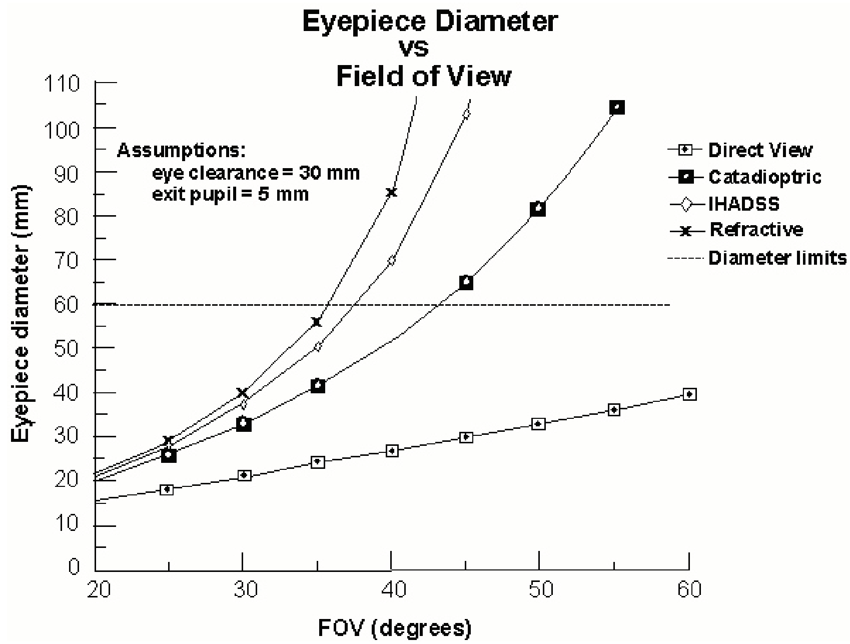


Figure 4-8. FOV versus eyepiece diameter for different designs.

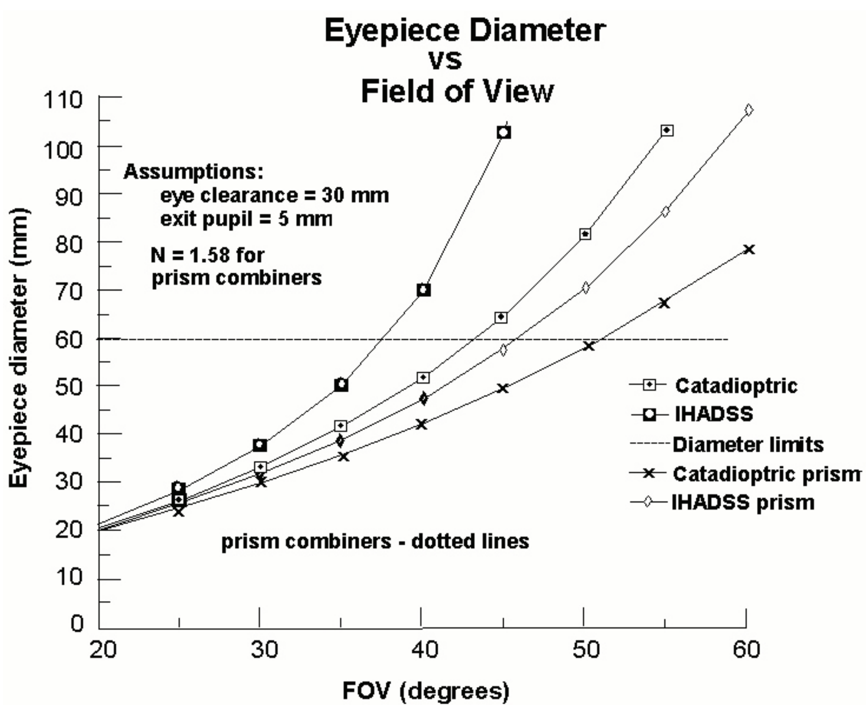


Figure 4-9. Comparisons between refractive and catadioptric HMDs with and without prism combiners.

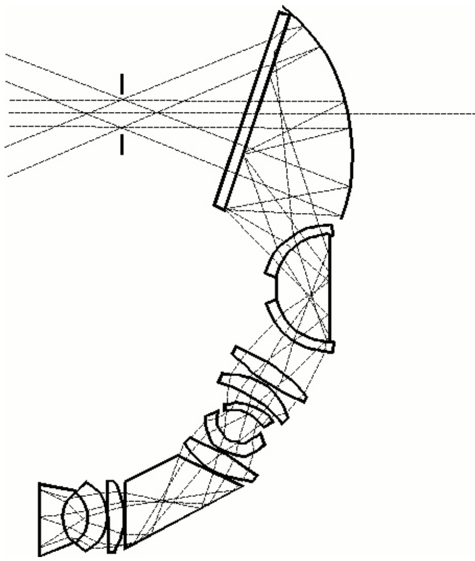


Figure 4-10. Ray trace of 50º x 60º tilted cat ocular (Droessler and Rotier,1989).

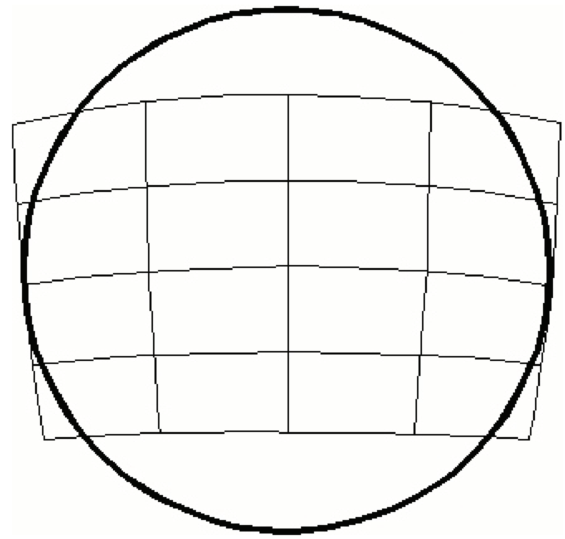


Figure 4-11. Optically induced distortion from tilted catadioptric, off-axis HMD design.

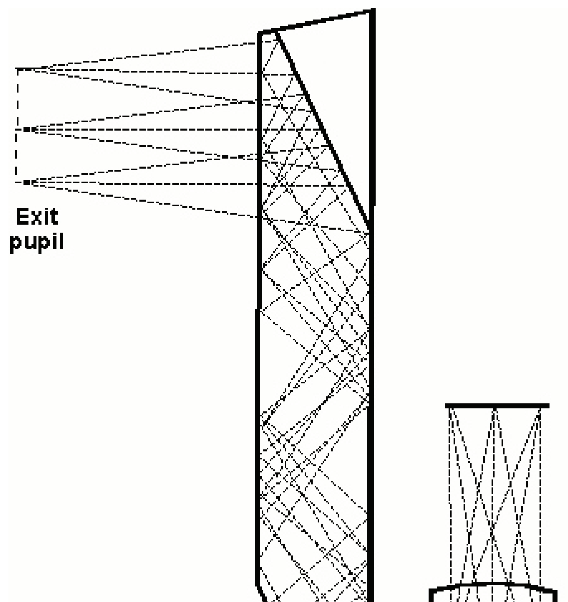


Figure 4-12. MONARC with rotationally symmetrical lens system (folded catadioptric).

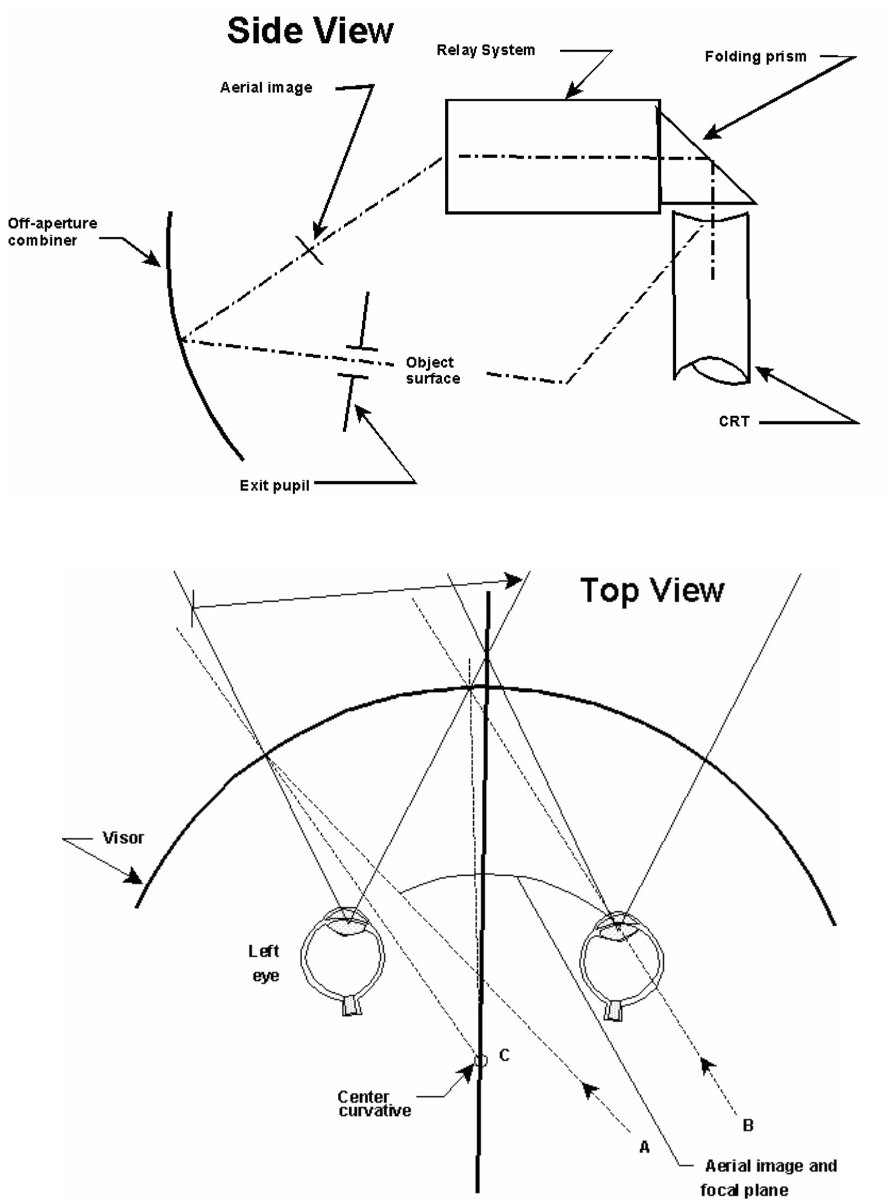


Figure 4-13. Reflective visor HMD: a) side view (top) and b) top view (bottom) (Shenker, 1987).

Pupil and nonpupil forming

A nonpupil forming virtual display uses a simple eyepiece to collimate or create a virtual image of a physical image source. An example is the ANVIS NVG where eyepieces produce virtual images of the 18-mm phosphor screens resulting in a 40º FOV. The display size, eyepiece focal length, eye clearance, exit pupil diameter, and *f/#* define the FOV relationships similar to viewing through a knot hole (Figure 4-5, top). A method to increase the apparent size of a display up to approximately 2X is with a coherent fiber-optic taper placed on the display. This approach based on a 1.5X taper was used with the Advanced I^2 program to obtain a 60º NVG FOV from the 18-

mm diameter intensifier tubes. The disadvantages of the expanding taper are a slightly increased weight compared to the 40º FOV ANVIS and reduced light transmission. However, without the taper, the increased tube diameter (from 18 mm to 27 mm) needed to obtain the same 60º FOV would weigh much more than the 18-mm tube with the 1.5X taper, but would not have a reduction in light transmission.

A pupil forming system has the same basic optical design as a compound microscope or telescope. Other common examples are rifle scopes, periscopes, and binoculars. For the pupil forming system, the eyepieces collimate real (aerial) images that are formed using relay optics. One purpose of the relay optics is to magnify the physical image source with the eyepiece providing additional magnification. Relay optics can also transport and invert the image as in the case of a periscope. The pupil forming system forms a real exit pupil that can be imaged with a translucent screen. Unlike the knothole analogy for the nonpupil-forming device, the pupil forming system requires the pupil of the eye to be positioned within a specific area to see the full FOV unvignetted. If the eye is moved closer than the exit pupil, the FOV will actually decrease. Also, if the eye is moved laterally outside the exit pupil, the complete display disappears where the nonpupil forming system merely vignettes the FOV in the opposite direction of lateral movement outside the exit pupil. The exit pupil for a pupil forming system is defined by the optical ray trace and is shown in Figure 4-14 (top) for the center of the FOV and Figure 4-14 (bottom) for the edge of the FOV. Note also the field lens, which is used to channel the aerial image to the eyepiece and adjust the eye clearance.

The relay optics of pupil forming devices usually are determined after the type eyepiece design, FOV, optical length, exit pupil diameter, and eye clearance values have been defined. To minimize the size and weight of the relay optics, the designer will attempt to use the shortest optical path possible within mechanical constraints.

Image Quality

For all of the sensor and display technology that goes into the final imagery presented to the Warfighter by an HMD, it is the quality of the imagery that determines its success. HMDs are used to present various types of information. These types include text, symbols, graphics, and video. Many factors affect the Warfighter’s ability to perceive and use this displayed information. If the information is a simple reproduction of computer generated text, symbols, or graphics, then the major factor affecting the fidelity of the information is the capacity of the HMD to faithfully reproduce the original image information. However, if the information is a representation of some external view of the world, as from an imaging system, then, in addition to the HMD’s capacity to faithfully reproduce the image, a number of additional factors will affect the user’s perception of the information. These include sensor parameters associated with the imaging system, transform functions associated with conversions of the scene from one domain to another (e.g., spatial, luminance, temporal), attenuation and filtering due to processing and signal transmission, noise, etc. However, ultimately, visual performance is limited by the quality of the final image.

What defines "acceptable" image quality varies from application to application and depends on the amount of information needed for the task(s) at hand; adequate image quality for one task may be insufficient in another. As previously stated, image quality is typically defined by a set of FOMs. Task (1979) described an extensive set of FOMs for defining image quality with CRTs. These FOMs are categorized as geometric, electronic and photometric in nature. Geometric FOMs include display source size, viewing distance, and aspect ratio. Electronic FOMs include bandwidth, dynamic range, and signal-to-noise ratio. For our discussion herein of visual HMDs, photometric FOMs are more important and include luminance, gray shades, contrast ratio, resolution, luminance uniformity, and MTF.

As flat panel displays replaced CRTs as the display technology of choice in the last two decades, the classification of image quality FOMs changed (Klymenko et al., 1997). For flat panel displays, FOMs have been categorized into four domains: spatial, spectral, luminance, and temporal (Table 4-1). These image domains parallel analogous human visual performance domains. The spatial domain includes those display parameters that are associated with angular view (subtense) of the observer and coincide with observer visual acuity and spatial sensitivity. The spectral domain consists of those parameters that are associated with the observer's visual sensitivity to color (wavelength). The luminance domain encompasses those display parameters identified with the overall sensitivity of the observer to levels of light intensity. The temporal domain addresses display parameters associated with the observer's sensitivity to changing levels of light intensity.

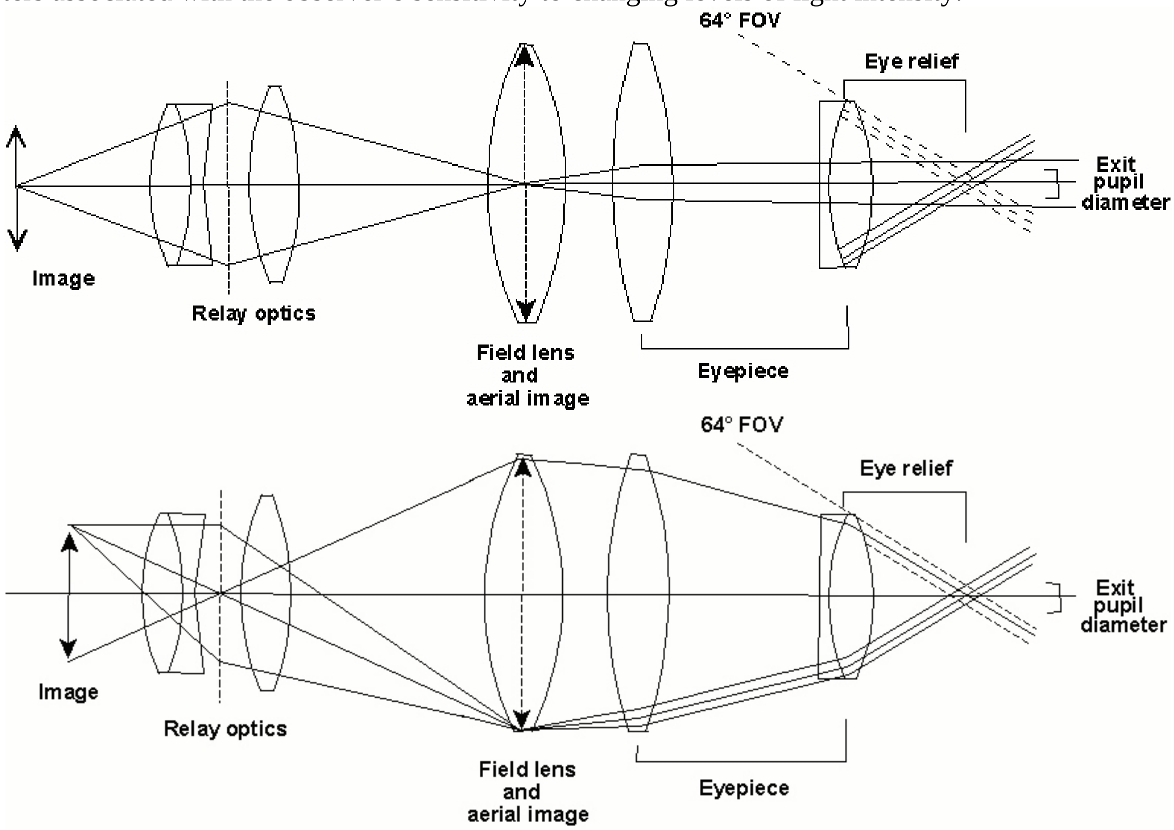


Figure 4-14. Ray trace of exit pupil formed by the center rays (top) and the marginal rays for a pupil forming optical device (bottom).

Table 4-1.
Flat panel display parameters (FOMs) (Klymenko et al., 1997).

Spatial	**Spectral**	**Luminance**	**Temporal**
Pixel resolution (H x V) Pixel size Pixel shape Pixel pitch Subpixel configuration Number of defective (sub)pixels	Spectral distribution Color gamut Chromaticity	Peak luminance Luminance range Gray levels Contrast (ratio) Uniformity Viewing angle Reflectance ratio Halation	Refresh rate Update rate Pixel on/off response rates

While all of these parametric FOMs are important, the key metrics of image quality generally are accepted to be *resolution*, *contrast*, and *distortion*. It may be argued that the most frequently asked HMD design question is "How much resolution must the system have?"

Resolution

Resolution is a measure of an imaging system's ability to reproduce scene detail (the amount of information). This will define the fidelity of the image. An HMD's resolution delineates the smallest size object (target) that can be displayed. A low-resolution image will appear blurry, lacking detail; a high-resolution image will appear sharp, presenting crisp edges and much detail.

In HMDs using CRTs as the image source, the CRT's resolution is the limiting resolution of the system. The CRT's horizontal resolution is defined primarily by the bandwidth of the electronics and the spot size. Vertical resolution is usually of greater interest and is defined primarily by the electron-beam current diameter and the spreading of light when the beam strikes the phosphor, which defines the spot size (and line width). CRT vertical resolution is usually expressed as the number of raster lines per display height. However, a more meaningful number is the raster line width, the smaller the line width, the better the resolution. Twenty microns (μm) is the current limit on line width in miniature CRTs (Rash et al., 1999).

In discrete displays (e.g., LCD, EL [electroluminescence], LED [Light Emitting Diode]), resolution is given as the number of horizontal by vertical pixels. These numbers depend on the size of the display, pixel size, spacing between pixels, and pixel shape (Snyder, 1985). Expressing resolution only in terms of the number of scan lines or addressable pixels is not a meaningful approach. It is more effective to quantify how modulation is transferred through the HMD as a function of spatial frequency. As in the discussion of optics earlier, a plot of such a transfer is called a modulation transfer function or MTF curve. Since any scene theoretically can be resolved into a set of spatial frequencies, it is possible to use a system's MTF to determine image degradation through the entire system. If the system is linear, the system MTF can be obtained by multiplying the MTFs of the system's individual major components.

Luminance contrast

Contrast is defined as the difference in luminance between two adjacent areas. An image with low contrast will appear washed out. There is often confusion associated with this term due to the multiple FOMs used to express contrast (Klymenko et al., 1997). *Contrast*, *contrast ratio*, and *modulation contrast* are three of the more common formulations of luminance contrast.

Confusion may result from the terminology, because different names are used for the two luminances involved in the definitions. Sometimes, the luminances are identified according to their relative values and, therefore, labeled as the *maximum* luminance (L_{max}) and *minimum* luminance (L_{min}). However, if the area at one luminance value is much smaller than the area at the second luminance, the luminance of the smaller area sometimes is referred to as the *target* luminance (L_t), and the luminance of the larger area is referred to as the *background* luminance (L_b). The more common mathematical expressions for luminance contrast include:

$$C = (L_t - L_b) / L_b \text{ for } L_t > L_b \text{ (Contrast)} \qquad \text{Equation 4-1a}$$

$$= (L_b - L_t) / L_b \text{ for } L_t < L_b \qquad \text{Equation 4-1b}$$

$$= (L_{max} - L_{min}) / L_{min} = (L_{max} / L_{min}) - 1 \qquad \text{Equation 4-1c}$$

$C_r = L_t / L_b$ for $L_t > L_b$ (Contrast ratio) Equation 4-2a

$= L_b / L_t$ for $L_t < L_b$ Equation 4-2b

$= L_{max} / L_{min}$ Equation 4-2c

and

$C_m = (L_{max} - L_{min}) / (L_{max} + L_{min})$ (Modulation contrast) Equation 4-3a

$= | (L_t - L_b)| / (L_t + L_b) |$ Equation 4-3b

In the preceding equations, modern conventions are adopted which preclude negative contrast values. [Classical work with the concept of contrast did not concern itself with whether the target or the background had the larger luminance value and, therefore, allowed negative contrast values (Blackwell, 1946; Blackwell and Blackwell, 1971).] The values for contrast as calculated by Equations 4-1a and 4-1c can range from 0 to infinity for bright targets and from 0 to 1 for dark targets (Equation 4-1b). The values for contrast ratio (Equations 4-2, a-c) can range from 1 to infinity. Modulation contrast (Equations 4-3, a-b), also known as Michelson contrast, is the preferred metric for cyclical targets such as sine waves and square waves. It can range in value from 0 to 1, and is sometimes given as the corresponding percentage from 0 to 100. Conversions between the various mathematical expressions for contrast can be performed through algebraic manipulation of the equations or through the use of nomographs (Farrell and Booth, 1984). Some of the conversion equations are:

$C_r = (1 + C_m)/(1 - C_m)$, Equation 4-4

$C_m = (C_r - 1)/(C_r + 1)$, Equation 4-5

$C = (2\,C_m)/(1 - C_m)$ for bright targets, Equation 4-6

and

$C = (2\,C_m)/(1 + C_m)$ for dark targets Equation 4-7

Available contrast depends on the luminance range of the display. The range from minimum to maximum luminance values that the display can produce is referred to as its *dynamic range*. A descriptor for the luminance dynamic range within a scene reproduced on a CRT display is the number of shades of grey (SOGs). SOGs are luminance steps that differ by a defined amount. They are, by convention, typically defined as differing by the square-root-of-two (approximately 1 .414).

These square-root-of-two SOGs have been used historically for CRTs, which had enjoyed a position of preeminence as the choice for given display applications. However, within the past two decades, discrete-element FPD technologies have gained a significant share of the display application market. Displays based on these various flat panel technologies differ greatly in the mechanism (physics) by which the luminance patterns are produced, and all of the mechanisms differ from that of CRTs. In addition, FPDs differ from conventional CRT displays in that most flat panel displays are digital with respect to the signals which control the resulting images. As a result, luminance values for flat panel displays usually are not continuously variable but can take on only certain discrete values. (Note: There are FPD designs which are capable of continuous luminance values, as well as CRTs which accept digital images.)

Confusion can occur when the concept of SOGs is applied to digital FPDs. Since these displays, in most cases, can produce only certain discrete luminance values, it is reasonable to count the total number of possible luminance steps and use this number as a FOM. However, this number should be referred to as “grey steps” or “grey levels,” not “grey shades.” For example, a given LCD may be specified by its manufacturer as having 64 grey levels. The uninitiated may misinterpret this as 64 shades of grey, which is incorrect. Its true meaning is that the display is capable of producing 64 different electronic signal levels between, and including, the minimum and maximum values, which generally implies 64 luminance levels. If one insisted on using a SOG FOM for discrete displays, it would appropriately depend on the value of the 1st and 64th levels.

To avoid confusion, designers should limit some FOMs to *either* discrete *or* analog displays. Contrast ratio, computed from maximum and minimum luminance, is applicable to both. The concept of SOG is most appropriate for analog displays and can be computed from contrast ratio. The number of grey levels is most appropriate for displays with discrete luminance steps, but additional information on how these grey levels sample the luminance range needs to be specified.

Other contrast FOMs may still be applicable to FPDs. However, in some cases they have to be adapted to conform to the unique characteristics of these displays. For example, because of the discrete nature of FPDs, where the image is formed by the collective turning on or off of an array of pixels, the concept of contrast ratio is redefined to indicate the difference in luminance between a pixel that is fully "on" and one that is "off" (Castellano, 1992). The equation for pixel contrast ratio is:

C_r = (Luminance of *ON* pixel)/(Luminance of *OFF* pixel) Equation 4-8

It can be argued that this pixel contrast ratio is a more important FOM for discrete displays. Unfortunately, the value of this FOM as cited by manufacturers is intrinsic in nature, that is, it is the contrast value in the absence of ambient lighting effects. The value of this FOM that is of real importance is the value that the user will actually encounter. This value depends not only on the ambient lighting level, but also on the reflective and diffusive properties of the display surface (Karim, 1992). Additional factors may need to be taken into consideration. An example is the dependence of luminance on the viewing angle where a liquid crystal display’s luminance output given by a manufacturer may only be reliable for a very limited viewing cone. Here the luminance and contrast need to be further specified as a function of viewing angle. On the other hand, the propensity of manufacturers sometimes to define "additional" FOMs that put their products in the best light must always be kept in mind.

The term grey scale is used to refer to the luminance values available on a display. (The term as used usually includes available color as well as luminance per se.) Grey scales can be analog or digital. The display may produce a continuous range of luminances, described by the shades of grey concept; or, it may only produce discrete luminance values referred to as grey steps or grey levels. The analog case is well specified by the SOG FOM and more compactly by the maximum contrast ratio of the dynamic range. Also the gamma function succinctly describes the transformation from luminance data (signal voltage) to displayed image luminance. (The MTF additionally describes the display’s operating performance in transferring contrast data to transient voltage beam differences over different spatial scales.) In an analog image, easily applicable image processing techniques, such as contrast enhancement algorithms, are available to reassign the grey levels to improve the visibility of the image information when the displayed image is poorly suited to human vision. (The techniques are easily applicable because they often simply transform one continuous function into another, where computer control over 256 levels is considered as approximating a continuous function for all practical purposes.) Poor images in need of image processing often occur in unnatural images, such as thermal images, and artificial images, such as computer generated magnetic resonance medical images. Since only certain discrete luminance levels are available in the digital case, the description of the grey scale and its effect on perception is not as simple and straightforward as in the analog case. One would like to know if there is a simple function that can describe the luminance scale; but one would also like to know how the function is sampled. A problem is that image enhancement techniques may not be as effective if the discrete sampling of the dynamic range is poor. For example, consider an infrared sensor generated image presented on an LCD with a small number of discrete grey levels. A contrast enhancement algorithm in reassigning pixel luminances must pick the nearest available discrete grey level and so could inadvertently camouflage targets by making them indistinguishable from adjacent background. Also, the original image might contain spurious edges because neighboring pixel luminance values, which would normally be close and appear as a smooth spatial luminance gradient become widely separated in luminance due to the available discrete levels, thus producing quantization noise (Rash, 2001).

Color contrast

Luminance differences are important in the ability to discriminate between two luminance values. However, even where the background and target have the same luminances, images can still be discerned by color differences (chromatic contrast). These equal luminance chromatic contrasts are less distinct in terms of visual acuity than luminance contrasts, but can be very visible under certain conditions (Kaiser, Herzberg, and Boynton, 1971).

The sensation of color is dependent not only on the spectral characteristics of the target being viewed, but also on the target's context and the ambient illumination (Godfrey, 1982). The sensation of color can be decomposed into three dimensions: hue, saturation, and brightness. Hue refers to what is normally meant by color, the subjective "blue, green, or red" appearance. Saturation refers to color purity and is related to the amount of neutral white light that is mixed with the color. Brightness refers to the perceived intensity of the light.

The appearance of color can be affected greatly by the color of adjacent areas, especially if one area is surrounded by the other. A color area will appear brighter, or less grey, if surrounded by a sufficiently large and relatively darker area, but will appear dimmer, or "more"grey, if surrounded by a relatively lighter area (Illuminating Engineering Society [IES], 1984). To further complicate matters, hues, saturations, and brightnesses all may undergo shifts in their values.

The use of color in displays increases the information capacity of displays and the natural appearance of the images. CRTs can be monochrome (usually black and white) or color. Color CRTs use three electron beams to individually excite red, blue, and green phosphors on the face of the CRT. By using the three primary colors and the continuous control of the intensity of each beam, a CRT display can provide "full-color" images. Likewise, FPDs can be monochrome or color. Many flat panel displays that produce color images are still classified as monochrome because these displays provide one color for the characters or symbols and the second color is reserved for the background, (i.e., all of the information is limited to a single color). An example is the classic orange-on-black plasma discharge display, where the images are orange plasma characters against a background colored by a green electroluminescent backlight (Castellano, 1992).

Full-color capability has been achieved within the last several years in most all of the flat panel technologies, including LC, EL, LED, field emission, and plasma displays. Even some of the lesser technologies, such as vacuum fluorescence, can provide multicolor capability. Research and development on improving color quality in flat panels is ongoing. FOMs describing the contrast and color generating capacities of displays are an ongoing area of development.

FOMs defining color contrast are more complicated than those presented previously where the contrast refers only to differences in luminance. Color contrast metrics must include differences in chromaticities as well as luminance. And, it is not as straightforward to transform chromatic differences into *just-noticeable-differences* (jnds) in a perceived color space. This is due to a number of reasons. One, color is perceptually a multidimensional variable. The chromatic aspect, or hue, is qualitative and two dimensional, consisting of a blue-yellow axis and a red-green axis. Additionally, the dimensions of saturation and brightness, as well as other factors such as the size and shape of a stimulus, affect the perceived color and perceived color differences. The nature of the stimulus, whether it is a surface color, reflected off a surface, or a self-luminous color, as present in a display, will affect the perceived color space in complex ways. Delineating the nature of perceived color space has been an active area of research with a vast literature (Widdel and Post, 1992).

As a consequence, there is no universally accepted formulation for color contrast. One FOM combining contrast due to both luminance and color, known as the discrimination index (ID), was developed by Calves and Brun (1978). The ID is defined as the linear distance between two points (representing the two stimuli) in a photo-colorimetric space. In such a space, each stimulus is represented by three coordinates (U, V, log L). The U and V coordinates are color coordinates defined by the CIE 1960 chromaticity diagram. The third coordinate, log L, is the base ten logarithm of the stimulus luminance. [A concise discussion of the discrimination index is presented in Rash, Monroe and Verona (1981).] The distance between two points (stimuli) is the ID and is expressed as:

$$ID = \sqrt{\left(\frac{\log(L_1/L_2)}{0.15}\right)^2 + \left(\frac{\sqrt{(\Delta U)^2 + (\Delta V)^2}}{0.027}\right)^2} \qquad \text{Equation 4-9}$$

where L_1 and L_2 refer to the luminances of the two stimuli, and (ΔU) and (ΔV) refer to the distances between the colors of the two stimuli in the 1960 CIE two dimensional color coordinate space.

A more recent FOM, ΔE (Lippert, 1986; Post, 1983), combining luminance and color differences into a single overall metric for contrast, has been provisionally recommended for colors which present only an impression of light, unrelated to context, only recently by the International Organization for Standardization (ISO, 1987) for colored symbols on a colored background. It is defined as follows:

$$\Delta E = \sqrt{(155\ \Delta L/L_{Max})^2 + (367\ \Delta u')^2 + (167\ \Delta v')^2} \qquad \text{Equation 4-10}$$

where the differential values (Δ) refer to the luminance (L) and chromaticity (u', v') differences between symbol and background and L_{max} refers to the maximum luminance of either symbol or background. Developing the appropriate FOM to describe the color contrast capacities of displays is an ongoing area of development (Widdel and Post, 1992).

Contrast and HMDs

This discussion has been general in nature. It is applicable to panel-mounted as well as HMDs. However, HMDs introduce additional contrast issues. For example, in IHADSS, the sensor imagery is superimposed over the see-through view of the real world. Although see-through HMD designs are effective and have proven successful, they are subject to contrast attenuation from the ambient illumination. The image contrast as seen through the display optics is degraded by the superimposed outside image from the see-through component, which transmits the ambient background luminance. This effect is very significant during daytime flight when ambient illumination is highest.

A typical HMD optical design in a simulated cockpit scenario is shown in Figure 4-15. The eyepiece optics consists of two combiners, one plano and one spherical. Light from the ambient scene passes through the aircraft canopy, helmet visor, both combiners, and then enters the eye. Simultaneously, light from an image source such as a CRT partially reflects first off of the plano combiner and then off of the spherical combiner, and then is transmitted back through the plano combiner into the eyes. If the characteristics of the various optical media are: 70% canopy transmittance; 85% and 18% transmittance for a clear and shaded visor, respectively; 70% transmittance (ambient towards the eye); 70% reflectance (CRT luminance back towards the eye) for the spherical combiner, 60% transmittance (ambient towards the eye) and 40% reflectance (CRT luminance) for the plano combiner, then one can analyze the light levels getting to the eye. An analysis of this design shows that approximately 17% of the luminance from the CRT image (and CRT optics) and approximately 25% of the ambient scene luminance reaches the eye for the clear visor (5% for the tinted visor).

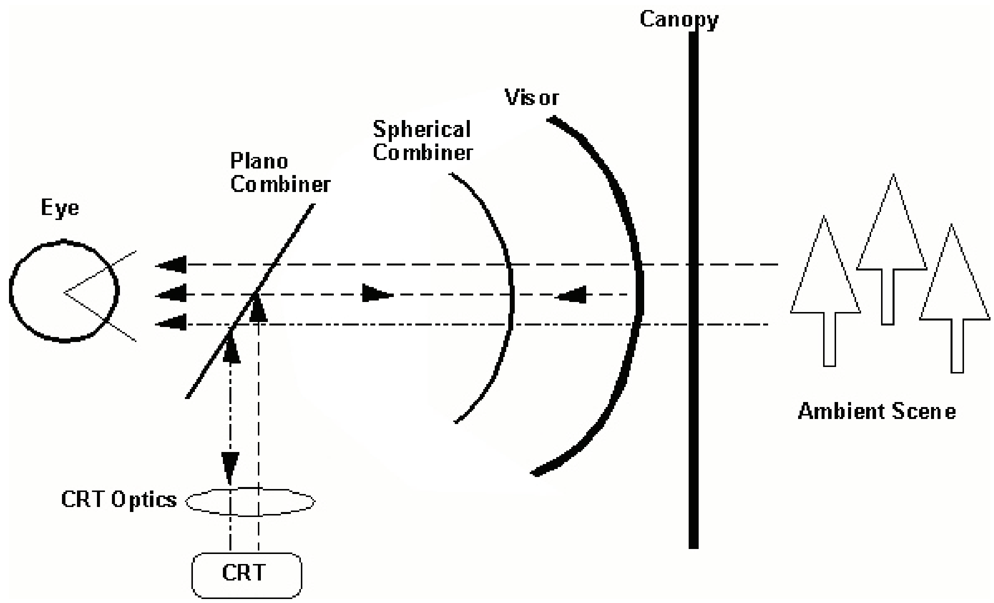


Figure 4-15. Typical catadioptric HMD optical design (Rash, 2001).

Distortion

Distortion usually is defined as a difference in the apparent geometry of the outside scene as viewed on or through the display. Sources of distortion in the display image include the image source and display optics (with combiner). For see-through designs, the combiner introduces distortion into the image of the outside scene. Distortion can exist outside the display itself, such as that caused by the aircraft windscreen. In current I^2 designs, e.g., ANVIS, the fiberoptic inverter is the primary source of distortion. Wells and Haas (1992) suggest that additional distortion can be induced in HMDs using CRTs as image sources. This distortion is perceptual and relates to a change in the shape of a raster-scanned picture on the retina during rapid eye movements (Crookes, 1957), such as those inherent in head-coupled systems.

Distortion in CRTs is rather easily minimized through the use of external correction circuitry. The CRT image also can be predistorted to allow for distortion induced in the display optics. FP image sources generally are considered to be distortion free, with the display optics being the source of any distortion present in HMDs using these sources. FP images also can be predistorted to correct for the display optics. However, this will require at least one additional frame of latency (Nelson, 1994).

In ANVIS, the optical system can produce barrel or pincushion distortion and the fiber-optic inverter can cause shear and gross (or "S") distortion. Shear distortion in fiber optic bundles causes discrete lateral displacements and is known also as incoherency. "S" distortion is due to the residual effect of the twist used to invert the image, which causes a straight line input to produce an "S" shape (Task, Hartman, and Zobel, 1993). Distortion requirements for ANVIS are cited in MIL-A-49425 (CR) and limit total distortion to 4%. Distortion for ANVIS typically is given as a function of angular position across the tube. Sample data from a single tube are presented in Figure 4-16 (Harding et al., 1996).

In Crowley's (1991) investigation of visual illusions with night vision devices, he cites examples of where aviators reported having the illusion of landing in a hole or depression when approaching a flat landing sight. Aviators also reported that normal scanning head movement with some pairs of ANVIS caused the illusion of trees bending.

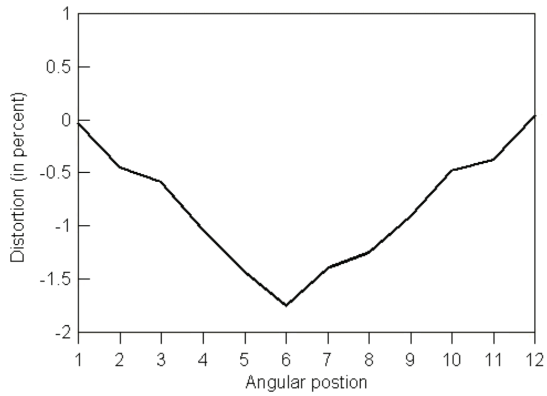


Figure 4-16. Percent ANVIS distortion as a function of angular position.

In general, for monocular, as well as for biocular/binocular, optical systems with fully overlapped fields of view, an overall 4% distortion value has usually been considered acceptable. That is, a deviation in image mapping towards the periphery of the display could be off by 4%, providing the deviation is gradual with no noticeable irregular waviness of vertical or horizontal lines. For a projected display with a 40-degree circular FOV and 4% distortion, this would mean an object at the edge of the visible FOV could appear at 40 x 1.04 (41.6º pincushion distortion) or 40/1.04 (38.5º barrel distortion). For binocular displays, differences in distortion between the images presented to the two eyes are more serious than the amount of distortion (Farrell and Booth, 1984.) Distortion is better tolerated in static images than in moving images, and therefore is of increased concern in HMDs.

Biocular/binocular HMDs having overlapping symbology will have to meet head-up display specifications of 1 mr or less difference between the right and left image channels for symbology within the binocular overlapped area if the symbology is seen by both eyes. Otherwise, diplopia and/or eye strain will be induced. However, with see-through vision, this criterion cannot be met when viewing at less than 60 meters due to eye convergence (McLean and Smith, 1987).

When imagery is used with a minimum see-through requirement, the maximum displacement between the right and left image points within the biocular/binocular region should not exceed 3 mr (0.3 prism diopter) for vertical (dipvergence), 1 mr (0.1 prism diopter) for divergence, and 5 mr (0.5 prism diopter) for convergence.

Distortion can be particularly important in aviation. For example, the apparent velocity of a target having a relative motion will change in proportion to the magnitude of the distortion (Fischer, 1997).

As an historical note, in 1988, when AN/PVS-5's were still the most common I^2 system, a number of reports from National Guard units surfaced regarding "depression" and "hump" illusions during approaches and landings (Markey, 1988). Suspect goggles were obtained and tested.

The final conclusion was that the distortion criteria were not sufficiently stringent. Based on testing, a recommendation was made to tighten both shear and "S" distortion specifications. Distortion requirements generally apply to single tubes. However, distortion differences between tubes in a pair of NVGs are more important. In fact, care should be taken to match tubes in pairs based on other characteristics; e.g., luminance, as well as distortion.

Display Technologies

While each component in an HMD design is important and plays a vital role in the design’s success or failure, it is easily argued that the image source component deserves special consideration. The selection of the image source has the largest impact on the quality of the image presented to the user.

The past several years have witnessed rapid emergence of a number of new candidate display technologies, each vying to replace the venerable CRT. Each of these new technologies has unique advantages and limitations (Table 4-2). In 1991, in order to address the need for miniature displays based on these new technologies, the Defense Advanced Research Projects Agency (DARPA) established a head-mounted display initiative as part of their High Definition Systems Program (Girolamo, 2001). The goals were to investigate and develop new display technologies that would overcome the limitations of CRTs and satisfy Department of Defense (DoD) needs for improved HMDs. At that time, the technologies selected were Active-Matrix Electro-Luminescent (AMEL) and Active-Matrix Liquid Crystal Display (AMLCD) as the most promising candidates. AMEL and AMLCD are two examples of a larger group of display technologies often referred to as Flat Panel Display (FPD) technologies. This label is somewhat inaccurately used to refer to the relatively thin profile, flat-face characteristics of displays employing these technologies. With the additional attributes of low-heat output, low-weight, and low-power consumption, this class of displays is especially attractive to HMD designers, as well as to users, such as the military, who operate in highly constrained physical environments.

Critical parameters

The role of the image source is usually two-fold. In most HMD applications, it is called upon to reproduce the picture of the outside scene for viewing by the user. In addition, the image source is used to display a range of symbology sets that represents such information as vehicle status, targeting reticules, fire-control (weapons) status, and map overlays. To perform these functions in a helmet-mounted configuration, the image source must meet a number of essential requirements that include:

- Sufficiently small physical dimensions
- Minimum weight
- Adequate image resolution
- Sufficient luminance
- Low power consumption

Size and weight

The physical dimensions of the image source need to be of appropriate size for head mounting; the optimal image plane diameter (or larger linear dimension) is 1 inch. This small size is required because in most HMD designs, the image source is collocated on the helmet and contributes to the head-supported weight (mass).

In the earliest HMD systems, the only production-available image source was the CRT. CRTs were notorious for their size, weight and power consumption, directly in opposition to virtually all of the requirements cited above for use in an HMD. This factor was a major driver in the development of miniature CRTs with diameters in the ¼- to 1-inch range.

Resolution

In any system, there is a weakest link (limiting factor). In imaging optical systems that are intended to reproduce details (resolution) of an outside scene and where this reproduced image is to be viewed by humans, it is desirable that the limiting factor be the human eye. Such a system design is said to be eye-limited. The reason for this viewpoint is that the human eye is the only component that cannot be improved. While this may no longer be rigorously true due to the development of wave front-guided laser surgery techniques, it remains an acceptable rule-of-thumb.

Table 4-2.
Summary of display technologies with advantages and disadvantages.

Category		Technology	Advantages	Disadvantages
Emissive	Scanning	CRT	Excellent resolution High conversion efficiency Infinite addressability Mature, well-known technology	Bulky, heavy, high power requirements Magnetic field sensitivity - shielding required Limited availability/ suppliers High voltage (8-12 kV)
Emissive	Matrix	EL & AMEL	Rugged Wide viewing angle Fast response time	Full-color problematic High voltage (80V) drive Limited availability/suppliers/ developers
Emissive	Matrix	FED	High luminance High conversion efficiency Uses CRT phosphors	Technology maturity High voltage (similar to CRT) Complex fabrication process Long-term reliability questionable
Emissive	Matrix	LED	Low cost Full-color available Lambertian emission	High power requirement Applications centered around illumination Miniaturization/array fabrication challenges
Emissive	Matrix	VFD	High luminance Wide viewing angle High efficiency Rugged, automotive use	Limited resolution Full-color problematic Miniaturization challenges
Emissive	Matrix	PDP	High efficiency Full-color	Miniaturization challenges High voltage drive
Emissive	Matrix	OLED	Low power/voltage operation Video speed available Full-color	Differential aging Limited availability/suppliers/ developers

The ability of a display to reproduce fine details is expressed by its resolution (the number of picture elements [pixels] producible along the vertical and horizontal dimensions of the image source). The definition of resolution depends on the class of image source technology. Virtually all image sources can be classified as matrix (discrete) or scanning. Most CRTs and some laser sources are classified as scanning sources, where the image is produced in a raster mode. A raster image is formed by moving a beam (of electrons or light [photons]) in a vertical series of horizontal lines. As a result, the image has a vertical resolution defined by the number of raster lines and a horizontal resolution defined by the bandwidth of the electronics and spot size of the electron or laser beam. CRT technology is very mature and historically has provided excellent resolution. Until the last decade, a CRT display

Table 4-2 (continued).
Summary of display technologies with advantages and disadvantages.

Non Emissive	Transmissive	AMLCD	Full-color Good image quality Video speed available Well established display technology	Limited temperature range – heater required Contrast drop at high temperature Low transmission efficiency
		Passive LCD	Low cost Simple design	Low resolution Slow response – causes smear Low multiplex capability
	Reflective	LCOS	High illumination efficiency	Response time in single panel configuration may cause smear
		FLC	Fast switching, no smear High illumination efficiency Potential for lower system cost	Limited availability/suppliers/ developers Limited temperature range
		DLP/DMD	Volume production High luminance for projection Good image quality (High contrast ratio) All-digital interface	High altitude (low air pressure) operation is problematic
	Scanning	RSD	High luminance Wide color gamut Infinite addressability	Costly Challenging packaging and ruggedization

had a preset fixed resolution. Most modern CRT displays are capable of adjusting the electron beam so as to provide multiple resolutions. Miniature CRTs are very specialized, have limited applications and limited availability. Military applications were a primary driver for miniature CRTs that were developed in ½-, ¾-, and 1-inch diameter sizes. A comparison of the characteristics of the various size tubes showed that the 1-inch tube offers the best raster imagery resolution and luminance (Levinsohn and Mason, 1997). A representative resolution of 1-inch tubes is of the order of 800 x 600. The IHADSS used on the AH-64 Apache uses a 1-inch CRT.

The development of the miniature CRT was an engineering achievement. However, even in its reduced format, the miniature CRT still has a weight, volume and power consumption footprint that challenges its choice as an image source for HMDs.

Fortunately, the 1980s brought a new class of image sources: discrete image sources. There are a number of matrix display technologies, collectively referred to as FPDs. These technologies include liquid crystal (LC), electroluminescent (EL), and light-emitting diodes (LEDs). Regardless of technology, a unique property of this class of displays is that they have individual pixels arranged in a matrix. Resolution for matrix-type or pixelated displays usually is given as the number of columns (horizontal pixels) by the number of rows (vertical pixels). As an example, a display with a stated resolution of 480 x 234 has 112,320 pixels arranged in 480 columns and 234 rows. The electronic industry has established specifications for specific standard resolutions. These include Super

Extended Graphics Array (SXGA) and Ultra Extended Graphics Array (UXGA). The SXGA specification has a 1280 x 1024 resolution; UXGA refers to a resolution of 1600 by 1200. Older, and lower, specifications of Video Graphics Array (VGA) and Super Video Graphics Array (SVGA) are most often used as a reference resolution. However, QVGA, having the lowest resolution of 320 by 240, is a popular display most often seen in mobile phones, Personal Digital Assistants (PDAs), and some handheld game consoles. Table 4-3 presents the resolution (in pixels horizontally by pixels vertically) for the more conventional specifications.

Ideally, for an optical system such as an HMD not to be display-limited, the image source should be capable of a resolution that meets or exceeds that of the human eye. For the normal human eye with a visual acuity of between 1-1.5 arc minutes and for an optimistic FOV as large as 120° (comparable with the horizontal extent of human vision), the resolution required is of the order of 4,800 horizontal pixels per display width; this exceeds by far the capability of current technologies. A more realistic FOV is 40°, requiring a resolution of 1600 pixels along the axis of the image source; this is equivalent to the UXGA specification.

Table 4-3.
Standard resolution specifications for matrix displays.

Specification	Resolution (H x V)
QVGA	320 x 240
VGA	640 x 480
SVGA	800 x 600
XGA	1024 x 728
SXGA	1280 x 1024
UXGA	1600 x 1200
HDTV	1920 x 1080

Figure 4-17 shows the required FOV of a display for a given number of pixels and at a pre-determined angular subtense of an individual pixel. For example, the very common SXGA resolution display at 1.5 arc minutes per pixel will only cover a FOV of the order of 30°, much lower than the unaided FOV of human vision.

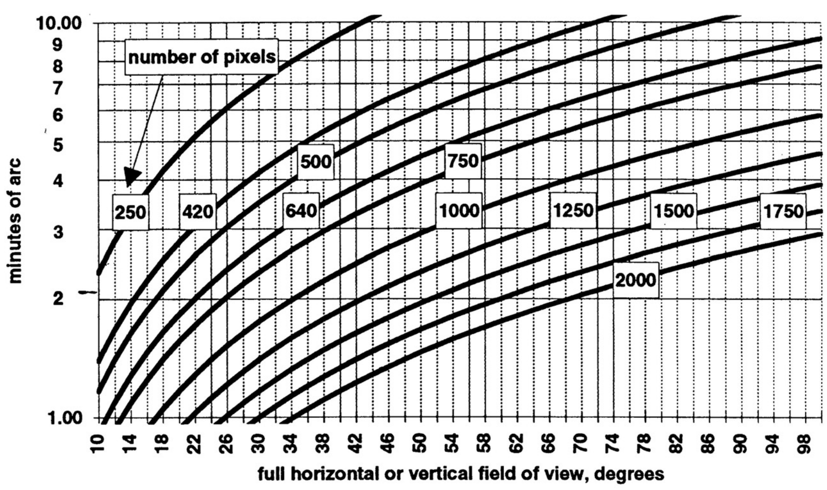


Figure 4-17. Resolution as a function of number of pixels and FOV (Melzer, 1997). When using this graph for imagery and it is assumed that the sensor has as many or more lines/pixels than the display, the resolution will be affected by the Kell factor of approximately 0.7. This means the effect number of lines of resolution is reduced by a factor of 0.7, e.g., a 1000-line or pixel display has an effective resolution of 700 lines or pixels.

Luminance

Until lasers could be packaged in a form making them useable in HMDs, image source luminance rivals resolution as the most important parameter. The image produced on the face of the image source has to successfully overcome the transmission losses incurred as the image's light rays traveled through the optics into the eye. More challenging are see-thorough HMD applications where the image luminance is required to be effectively viewed against the ambient light level of the outside world; the luminance needed for see-through HMD configurations is a strong function of the background luminance (up to 10,000 foot-Lamberts (fL) for white clouds). See-through HMDs intended for day use, require the addition of a tinted visor to reduce the level of the background luminance at the eye.

The concept of attaching a luminance value to a display image source is misleading. Melzer and Moffitt (1997) describe two luminance values that may be used to specify needed image source luminance: peak luminance and average luminance. Peak luminance is the maximum luminance that can be achieved (given maximum input). This can be defined as on-axis or off-axis for a given display source. A specification for peak luminance is recommended when symbology only is displayed (i.e., no imagery). In applications that do present scene imagery, an average luminance across the image source is recommended. Average luminance will be less than any peak luminance present in the scene and its value will depend on the content of the scene. To allow comparison between several image sources, the average luminance should be based on a universal test pattern, preferably one with both high and low spatial frequencies.

Power consumption

In vehicular HMD applications or other applications where on-site power is available, power requirements are less of an issue than for ground applications where the Warfighter must carry his power requirements with him in the form of batteries. However, even when on-site power is available, the HMD designer cannot be given carte blanche not to optimize power consumption for the image source or other HMD components.

Fortunately, the FPD technologies have greatly reduced the image source power requirements. Nonetheless, with regard to image source power consumption, two main factors still place constraints on the amount of power that can be made available in an HMD design:

- The more power consumed by an image source, the greater the heat generation. Because of the great need to reduce head-supported weight, standard mechanisms for effective heat removal – addition of a heat sink and/or a fan – are not viable options.
- In self-contained ground applications, battery power availability for man-wearable systems is limited.

Display technology classification

All display technologies are generally classified as *emissive (light generators)* or *non-emissive (light modulator)* based on their capability to either create their own light or the need to operate by modulating the transmission and/or reflection of an independent external light source. This classification and the subcategories of displays are presented in Figure 4-18. Both emissive and non-emissive displays can be further categorized as discrete (matrix) or scanning displays (Table 4-2).

Emissive displays

The underlying mechanism of emissive displays is that they emit visible light in response to some excitation action. Most emissive display technologies employ a phosphor material as the source of the visible light. These include CRTs, vacuum fluorescent displays (VFDs), electroluminescent (EL) displays, and white light-emitting diodes (LEDs) that use a phosphor coating to achieve white light output (Hur and Pham, 2001). Various LED and plasma technologies also are classified as emissive displays but use other mechanisms for light production.

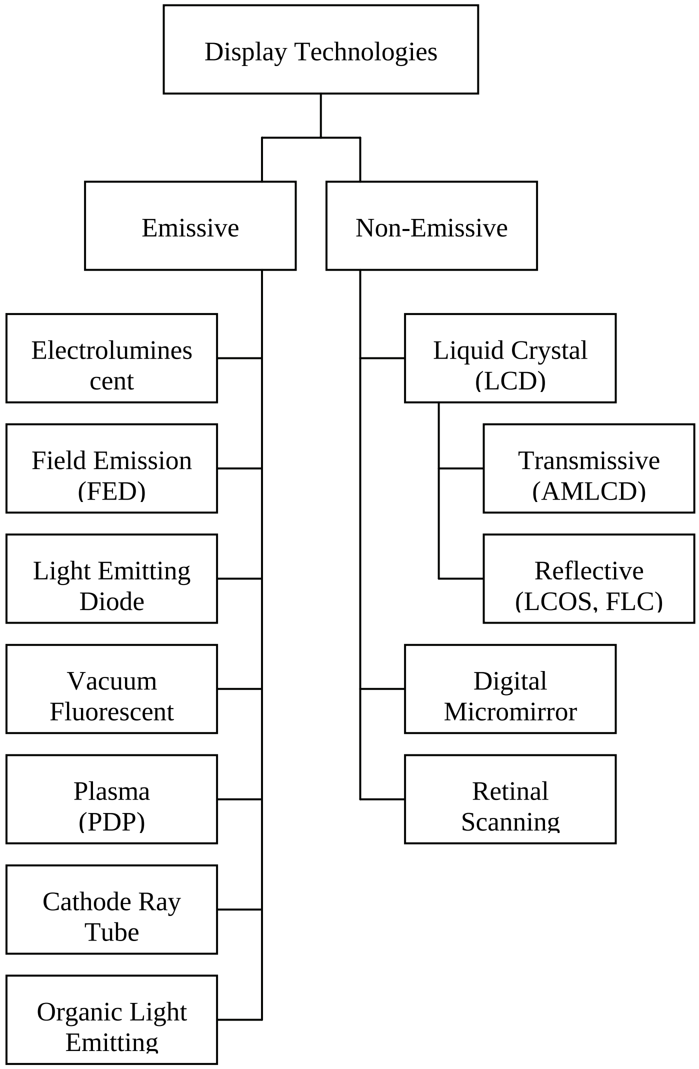


Figure 4-18. Classification of display technologies.

A phosphor is an inorganic chemical compound designed to emit visible light (fluorescence) when excited by ultraviolet radiation, x-rays or an electron beam. The amount of visible light produced is proportional to the amount of excitation energy. If the fluorescence does not terminate when the excitation energy stops, but instead decays slowly after the excitation energy is removed, the material is said to be *phosphorescent*. Succinctly, fluorescence occurs only during the period that the phosphor material is being excited and ends within approximately 0.01 microseconds after the termination of the bombardment (Farrell and Booth, 1984). Phosphorescence may persist over periods extending from a fraction of a microsecond to hours. By consensus, phosphors are designated by the letter "P" and a number, e.g., P1, P45, and P104. Each designation defines a specific chemical composition and a set of performance characteristics.

The first phosphor was created by an Italian alchemist, Vincenzo Cascariolo, in 1603, as a result of his research into transmutation of materials (Keller, 1997). This is considered by some historians to be the single most important discovery in inorganic luminescence and has become the primary basis for image production.

Phosphors have three performance characteristics that impact their selection for a specific display application: spectral distribution, luminous efficiency, and persistence (Rash and Becher, 1983). The spectral distribution of a phosphor is important in transferring display luminance to the eye. The eye's photopic (daytime, >1 fL) response peaks at approximately 555 nanometers (nm), which is in the green region of the visible spectrum. [The eye's nighttime (scotopic response) peaks at approximately 507 nm.] It is not coincidental that many phosphors employed in displays have a green or greenish yellow color (Rash, 2001) (Figure 4-19). For example, fielded ANVIS uses the P20 (older) or P22-Green phosphors; IHADSS uses the P43 (which is being fielded for ANVIS use also) and the now cancelled HIDSS planned to use P53 (Green). It is important to know that many phosphors have more than one peak wavelength. For example, P43 has three peaks (blue, red, and green). As for the phosphor employed in the IHADSS' miniature CRT, filters are used to suppress the unwanted red and blue side-lobe wavelengths.

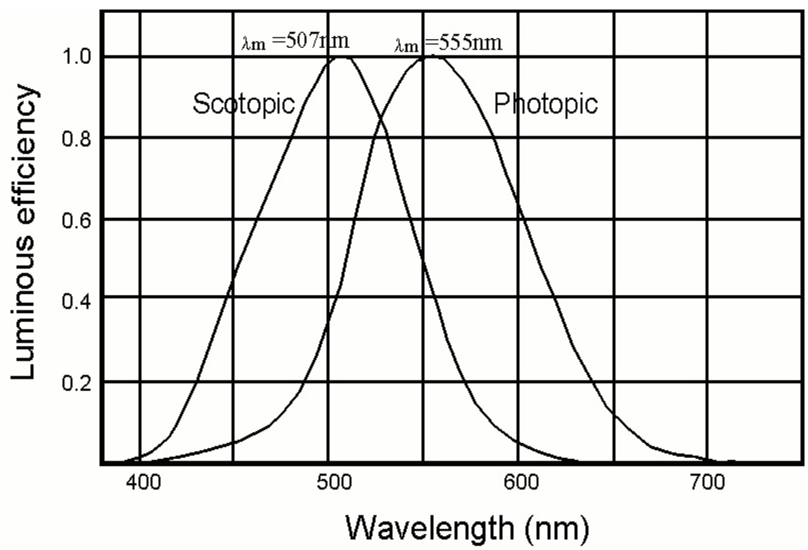


Figure 4-19. The human eye's photopic (day) and scotopic (night) response curves.

The necessity to use an optical filter in the IHADSS P43 CRT means that a proportion of the phosphor's luminous (light) output is wasted. This leads to the second important characteristic of phosphors, luminous efficiency. Luminous efficiency is defined as the ratio of the energy of the visible light output to the energy of the input signal. It is expressed in units of lumens per Watt (lm/W), a ratio of visible light (in lumens) to the input power (in Watts).

Since power consumption is an important concern, the more efficient the image source (with respect to its light production mechanism) is at changing input power into light, the more acceptable the source will be to an HMD design. In addition, the more efficient an image source is, the more light it will produce for a given input power. Therefore, for a given transmission loss in the relay optics, the more efficient the image source will be and the greater the amount of light that will be delivered to the viewer's eye(s).

The persistence of a phosphor, defined as the time required for a phosphor's luminance output to fall to 10% of its maximum, is the major factor in the dynamic or temporal response of the display. In the military aviation environment, the temporal response of the total imaging system (sensor, display, and associated electronics) is especially critical in pilotage and target acquisition tasks (Rash and Verona, 1987). The loss of temporal response will result in a degraded modulation contrast at all spatial frequencies (but with greater losses at higher frequencies) (Rash and Becher, 1982). The consequence of the loss of contrast at the higher frequencies is that fine details (e.g., wires, tree branches) in the scene will not be present in the image viewed by the pilot or by any user in other applications.

Non-emissive displays

As the name implies, non-emissive displays do not generate light by themselves, but rather act as a light valve for an external light source. They may be reflective, in which case the light source is located on the front side of the display, or transmissive, in which case the light source is placed behind the display, or a combination of both (transflective). In each case, the display pixels act as individual (discrete) light switches. For a reflective display, the switch behaves as a mirror, directing the light toward the observer during the ON time and away from the observer during OFF time; for a transmissive display, the light switch becomes a shutter, open (transparent) during the ON time and closed (opaque) during the OFF time.

Examples of reflective displays include liquid crystal on silicon (LCOS) and digital micro-mirror displays (DMD). The optical design for reflective displays is more demanding. This is because during pixel-off-time, light is either scattered, or absorbed, or redirected away from the light path to the eye. Consequently, greater care must be taken in the design in order to prevent stray light from reducing contrast. In terms of advantages, this category of displays presents:

- Increased pixel aperture fill factor - results in smaller pixels and higher density (each pixel drive can be hidden under the pixel itself, behind the reflective layer).
- Increased luminance - reflection coefficient of the order >70%.

Transmissive displays require rear illumination but potentially can provide higher luminance. Their disadvantages are mostly related to their need for a backlight; these include greater power consumption, increased weight and volume, and heat generation. The best known example of this category is the AMLCD.

The example display technologies cited above are just a few of the many available to the HMD designer. All of which will be discussed fully in the following sections.

Pixel method of classification

An alternative method for classifying FPDs is by the number of pixels generated simultaneously (Figure 4-20) (Powell, 1999). Using this approach, the following classifications are used:

- Matrix display – All pixels are generated independently and are directly addressable. These displays usually have a large number of pixels, from several thousand to more than a million. See Figures 4-21 and 4-22 for illustrations of various display designs having a matrix structure.
- Line display – All pixels of one display line (*x*-dimension) are generated independently and are directly addressable; the line is scanned in the *y*-dimension. Some position feedback mechanism is required by the display generator to update the display drive according to the instantaneous location in *y*-direction of the display.
- Single pixel display – Only one pixel (a beam) is generated. Two-dimensional (2-D) scanning mechanisms position the beam in both the *x*-, and *y*- dimensions. As in the line display case, positional feedback mechanism is required by the display generator in order to update the drive according to the instantaneous location of the beam. See Figure 4-23 for illustrations of single pixel structures. A typical CRT display is an example.

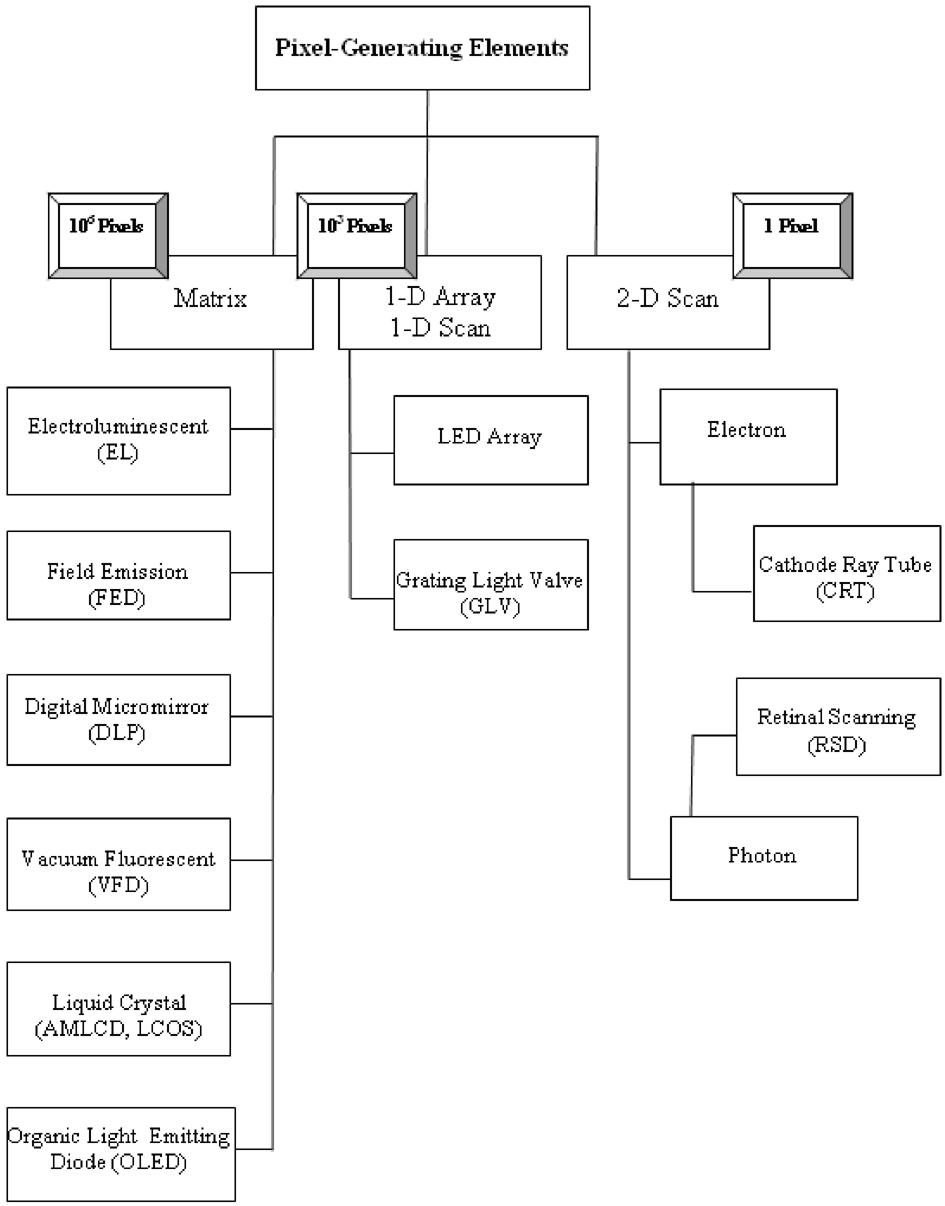


Figure 4-20. Pixel method of display classification (Urey, 1999).

The matrix and the scanning implementation will be discussed in greater detail in following sections. Technologies based on scanning in one dimension use a linear array of about 10^3 pixels and can provide high resolution and good image quality. However, a correlation of pixel variations in the scan direction leads to more stringent luminance matching conditions than for the matrix approach. Successful applications such as fax machines, document scanners, and cameras demonstrate that these problems can be overcome but at the cost of speed and complexity. Consequently, the speed of the transport mechanism and the size of the pixels limit this approach for HMDs.

Light generation method of classification

It is useful and interesting to further investigate the mechanisms used by the various display technologies to generate light. Such an investigation provides a third possible classification approach, effectively a combination of the first two (Ferrin, 1997). For the emissive (self-contained) displays, the mechanisms include phosphorescence (CRT), electroluminescence (AM and AMEL), field emission (FED), fluorescence (VFD), or gas discharge (Plasma). Both the reflective and transmissive displays depend on an external light source that is selected based on system performance requirements. These mechanisms of light generation are summarized in Table 4-4 and more fully discussed in subsequent sections.

Major display technologies

In the following sections, each of the major display technologies is briefly reviewed. The information presented is intended to provide the reader with an overview of the most dominant display technologies available for HMDs. Those described are not all inclusive. Even within each major category, it is difficult to accomplish more than to provide a snapshot of the individual technologies as their development is still in flux. For more in-depth discussions of these technologies, readers are encouraged to consult more dedicated resources (e.g., Castellano, 1992; Keller, 1991; Kalinowski, 2004; Sherr, 1993; Tannas, 1985; Wu, 2001; Wu and Yang, 2006; Yeh, 1999).

Cathode-ray-tubes (CRTs)

The cathode-ray-tube (CRT) was invented by German physicist Karl Ferdinand Braun in 1897. In its simplest form, a CRT is an electron vacuum tube with an electron source (cathode) at one end and a phosphor screen at the other, usually with internal or external means to accelerate and deflect the electrons (Keller, 1991) (Figure 4-24). Figure 4-25 presents a typical CRT electron source, referred to as an *electron gun*. The CRT ranks near the top for luminance, resolution, flexibility in addressability. It ranks at the bottom on size (primarily depth), weight, high anode voltage, power requirements and heat generation. High performance miniature (≤ 1 inch diameter) monochrome CRTs have been developed for HMD applications. Some of the requirements and design trade-offs for an HMD-designed CRT are summarized in Sauerborn, 1995.

Cathodes

Thermionic cathodes use heat to generate electrons from a solid material and come in two main categories:

- Oxide (film) cathodes of the traditional "RCA-design," consisting of a thick (25 μm to 50 μm) film layer of mostly a mixture of barium, calcium and strontium oxide on nickel, operating at 750ºC to 800ºC, or
- Barium oxide (BaO) cathodes deposited on tungsten that operate at slightly higher temperature (900ºC to 1000ºC). The major limitation of oxide cathodes is that average current density is limited to about 1 Ampere per square-centimeter (Amp/cm^2). The anticipated lifetime of a standard oxide cathode when loading increases to 2 Amp/cm^2 drops to less than 10,000 hours (Falce, 1992).

Table 4-4.
Image generating mechanism: Summary table.

	Emissive		**Reflective**			**Transmissive**
Structure	Matrix	Single Pixel (Scanning)	Matrix	Matrix	Single Pixel (Scanning)	Matrix
Technology	AMEL, EL, LED Array, VFD, OLED, Plasma	CRT	DMD, FLCD	LCOS	RSD	Nematic LCD, AMLCD
Light source mechanism	Phosphors	Electron beam	Lamp, LEDs	Lamp, LEDs White light	Laser beam (R, G, B)	Backlight (EL, LED)
Color mechanism	Spatial: Color phosphors	Spatial: Shadow Mask Temporal: LC Shutter	Temporal	3 panels: Spatial 1 Panel: Temporal	Temporal	Spatial: Color Filters

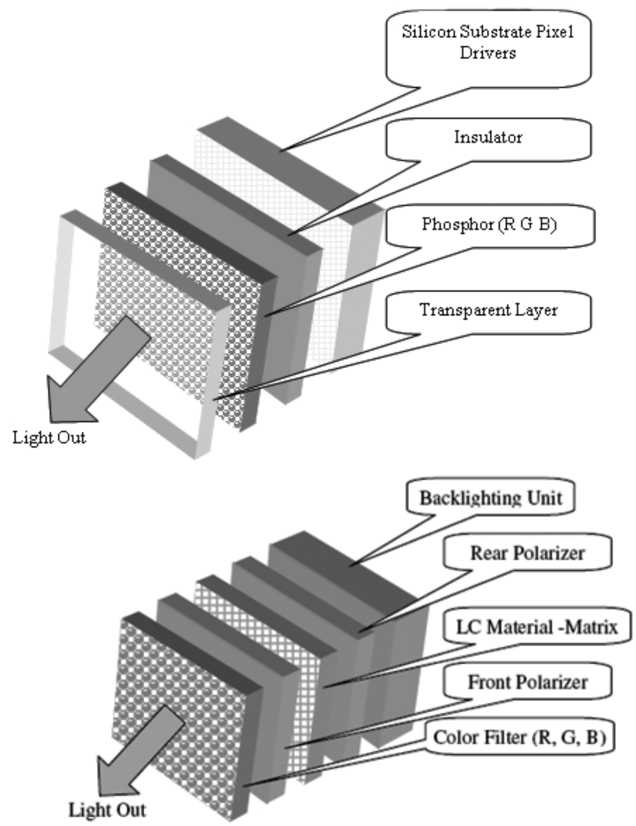
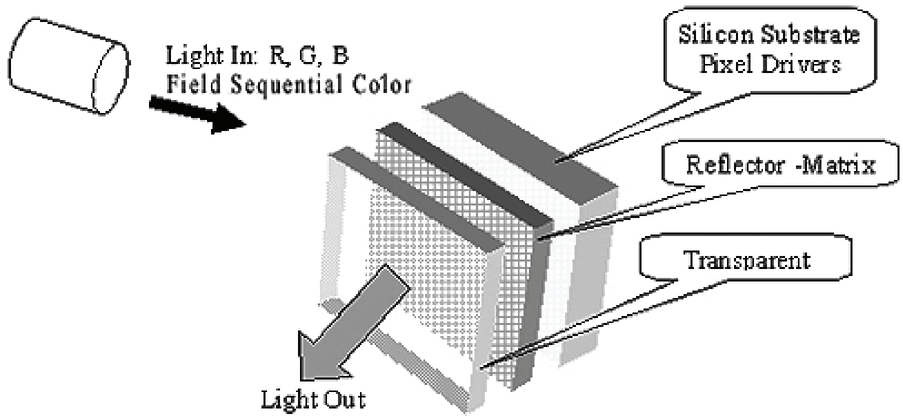


Figure 4-21. Illustrations of various matrix structure displays: Emissive Display: Matrix Structure (OLED, LED Array, VFD, EL, AMEL) (top) and Transmissive Display: Matrix Structure (Nematic LCD, AMLCD) (bottom).

Volumetric cathodes are used when higher average current density emission is needed. Originally developed by Philips in 1940's, the emission mechanism differs significantly from that of oxide cathodes. In this latter case, the extraction mechanism of electrons from the outer orbit of an atom is brute force heat. In the case of a barium-activated metal surface, the positively charged barium and the negatively charged oxygen create an electric dipole acting as an extracting grid assisting with electron extraction (Falce, 1992). Volumetric cathodes come in two designs: dispenser and reservoir.

- Dispenser cathodes employ a porous tungsten matrix and come in two varieties: impregnated and reservoir. Impregnated dispenser cathodes have a barium compound in the pores of the matrix. When the cathode is heated, this barium compound interacts with the tungsten and releases free barium that coats the surface. Typical average current density from an osmium-coated impregnated cathode operating at 980ºC may reach 4-5 Amps/cm^2. For comparison purpose, the anticipated lifetime of a dispenser cathode under 2 Amp/cm^2 load exceeds 50,000 hours (Falce, 1992).
- Reservoir cathodes are more difficult to build, but they last longer and can be pushed to higher emission currents. Current densities of 100 Amps/cm^2 have been achieved in the laboratory. A reservoir cathode has a "reservoir" of barium emission material behind the tungsten matrix. When heated, the barium comes out of the reservoir, infiltrates through the matrix and coats the forward surface.



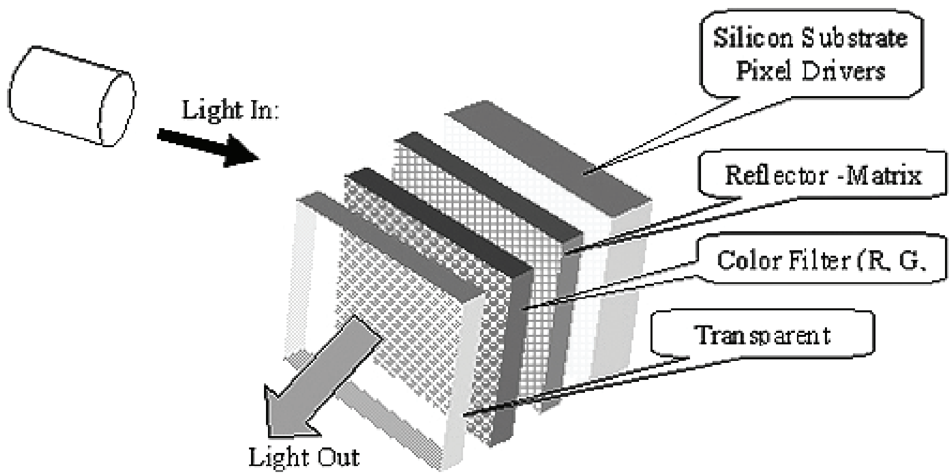


Figure 4-22. Illustrations of various matrix structure displays: Reflective Display: Matrix Structure (DMD, FLCD) (top) and Reflective Display: Matrix Structure (LCOS) (bottom).

The majority of the CRTs used in HMDs employ the dispenser cathode type.

Phosphors

After emission from the cathode, the electron beam is accelerated towards the phosphor screen. The beam is deflected to strike on the desired position on the phosphor screen by a magnetic field. This field is generated by a deflection yoke that has separate sets of coils for horizontal and vertical deflection. The beam deflection amplitude is controlled by the intensity of the magnetic field, which is in turn controlled by the current injected in the coils. When the beam electrons impinge upon the phosphor screen, the phosphors grain (particle) at that particular location emits light by converting the kinetic energy of the electron to photons, i.e., the photoelectric effect.

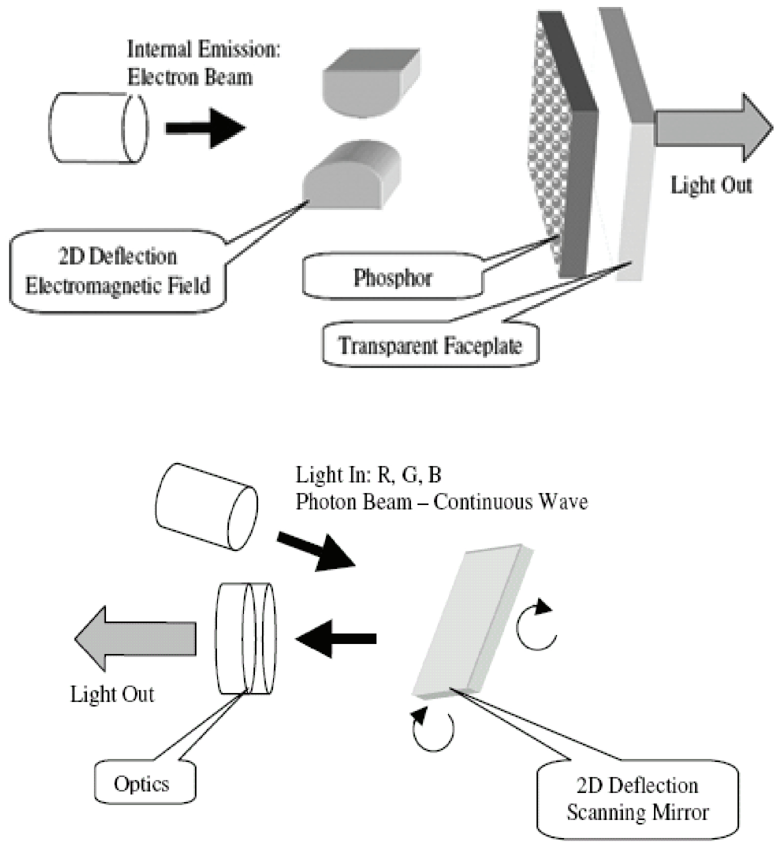


Figure 4-23. Illustrations of various single pixel structure displays: Emissive Display: Single Pixel Structure (Scanning) CRT (top) and Emissive Display: Single Pixel.Structure (Scanning) VR (bottom).

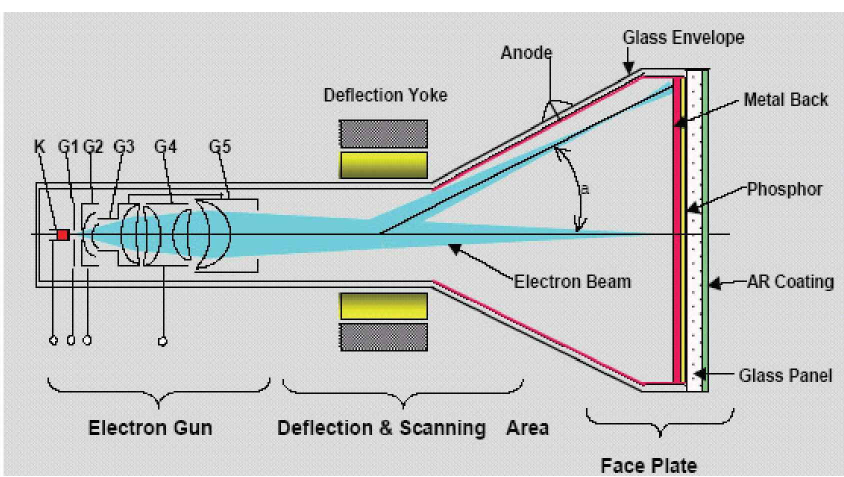


Figure 4-24. Diagram of a typical CRT (Fujioka, 2001).

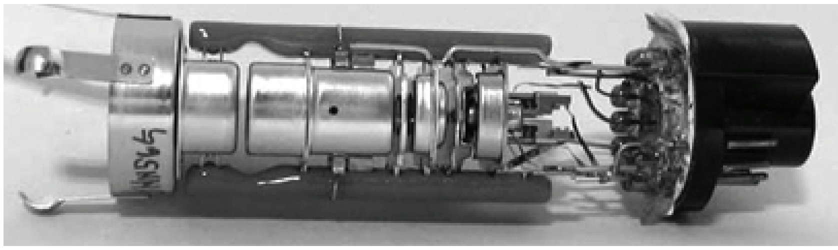


Figure 4-25. Photograph of a CRT Electron Gun. (Source: Wikipedia)

In general the phosphor is an inorganic crystal with grains (particles) of around 7 to 10 μm in size. Characteristics of the major phosphors used for CRTs employed in HMDs are listed in Table 4-5.

Phosphor persistence classification is based on the time required to decay to 10% of peak luminance (Figure 4-26):

- Very long: 1 sec and longer
- Long: 100 ms to 1 sec
- Medium: 1 ms to 100 ms
- Medium short: 10 μsec to 1 ms
- Short: 1 μsec to 10 μsec
- Very short: less than 1 μsec

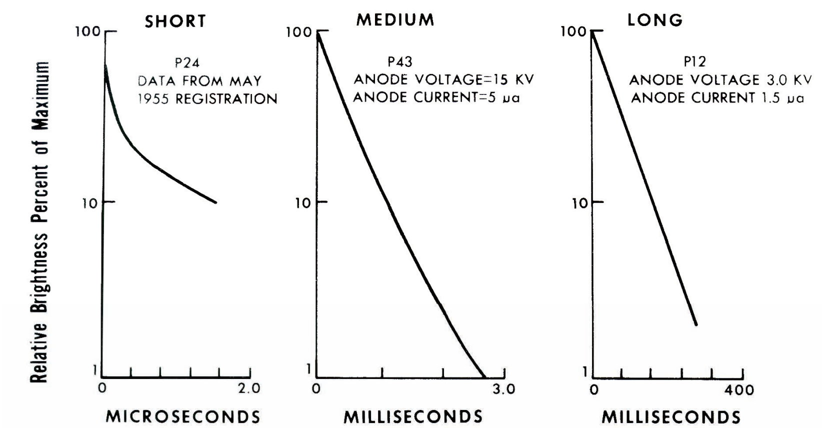


Figure 4-26. Typical decay curves for short, medium, and long persistence phosphors.

Spectral distribution

Spectral distribution refers to the wavelengths for which the phosphor emits energy. Knowledge of this distribution is essential in order to optimize the HMD display for good day-time performance. The photopic response of the human eye peaks at about 555 nm (Figure 4-19). For a phosphor such as P43 and P53 (Figure 4-27) that have the majority (>70%) of their energy concentrated in a narrow band, a matched notch optical filter is needed to allow most of the phosphor light to pass but reject the rest of the visible spectrum thus producing an improvement in the display contrast ratio.

Fiber-optic faceplate

The last surface on the CRT that the light must traverse is known as the faceplate. A plain-glass faceplate on a CRT can cause spurious screen illumination due to internal reflections caused mainly by halation and chromatic aberrations. The halation mechanism is shown in Figure 4-28. When the electron beam strikes the phosphor layer, light rays enter the glass faceplate at various angles. Rays striking the glass above the critical angle are reflected internally back to the phosphor layer generating spurious light. This increases the effective spot size, leading to a reduction of CRT resolution.

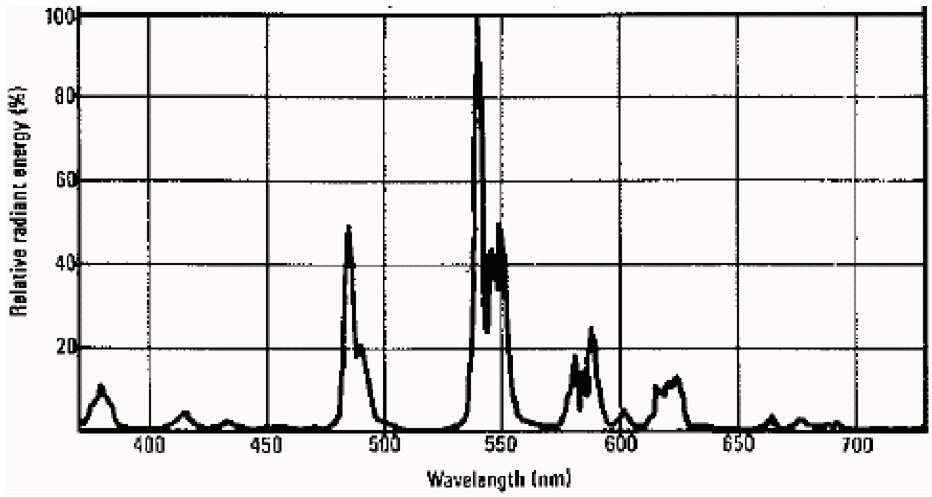


Figure 4-27. P53 Phosphor spectral characteristics.

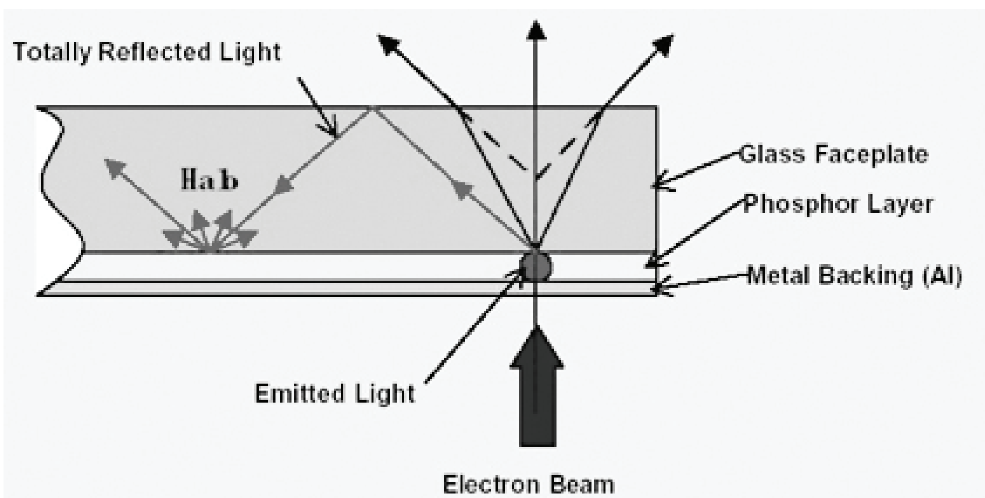


Figure 4-28. Halation in plain-face faceplates in CRTs (Fujioka, 2001).

Table 4-5.
Phosphor characteristics.

Phosphor	P1	P20	P31	P43	P53
Composition	Zn_2SiO_4:Mn	ZnCdS:Ag	ZnS:Cu	Gd_2O_2S:Tb	$Y_3Al_3Ga_2O_2$:Tb
Peak wavelength	525 nm	560 nm	520 nm	543 nm	546 nm
Color	Green	Green-yellow	Green	Yellow-green	Yellow-green
Color Coordinates	x: 0.208 y: 0.704	x: 0.297 y: 0.571	x: y:	x: 0.337 y: 0.561	x: 0.355 y: 0.557
Luminous efficiency	30 lm/W	16 lm/W	40-45 lm/W	40 lm/W	30 lm/W
Persistence	Medium 1-100 msec	Medium-short 60 µsec - 3 msec	Medium-short 100 µsec	Medium 1.3 msec	Medium 6.5 msec
HMD applications	Prototype IHADSS	ANVIS	Used in oscilloscopes	ANVIS/ IHADSS	Selected for HIDSS

Note 1: msec = milliseconds.

Note 2: Over the years, many formulations of the composition of these various phosphors have been developed. The characteristic values presented in this table should be interpreted as guidelines only.

Replacing the solid glass faceplate with a fiber-optic faceplate eliminates both the halation and any chromatic aberrations. A fiber-optic faceplate is a coherent array of millions of optical-fiber waveguides per square inch, each having a diameter of 3 to 10 μm. It acts as an image plane transfer device – an image entering one surface exits as an undistorted digitized image regardless of the shape of the optics itself (Cook and Patterson, 1991). Typically the fiber-optics used have the same coefficient of thermal expansion as the CRT glass, which allows them to be fused directly to the CRT. They are curved on the inside to match the deflection angle of the tube and are flat on the outside. This eliminates the need for dynamic focusing of the electron beam. Fiber-optic faceplates were originally introduced in night vision goggles as the substrate for the phosphor screen at the viewer's end.

Color CRT

The quest for color is fundamental for any display technology. Large-size CRTs achieved full-color capability early during the technology development process using a shadow-mask located in front of the phosphor deposited in a red (R), green (G), and blue (B) pattern, splitting each individual pixel into three *subpixels* placed so closely that the eye cannot distinguish among them. The shadow mask is a metal plate (e.g., invar [a nickel steel alloy]) that effectively ties each of the three electron guns (beams) to one phosphor spot (consisting of three color subpixels) only (Figure 4-29). Driving each color gun with video information pertaining to that particular color for each phosphor spot produces three color pictures in the fundamental colors. The eye spatially integrates the three pictures into one full-color picture.

Currently, the shadow mask technology though is limited to above-medium-size CRTs; also the packaging of the three electron guns and the convergence of three electron beams is difficult to achieve in a CRT smaller than 5 inches (12.7 cm) diagonal (Sherman, 1995).

Field-Sequential Color (FSC) bridges the gap between the capabilities of monochrome CRT and the need for color. Compared to the shadow mask approach, which creates color *spatially*, FSC produces color *temporally*. The video information is generated on a frame-by-frame basis, each frame successively of R, G, B colors, that are displayed in time sequence. If the fields are refreshed fast enough, above the critical flicker frequency of the human visual system (>30 Hz), the viewer integrates the individual fields into a full-color picture. This is the same principle used by the movie industry to create motion from blending a rapid sequence of still images.

Practical implementation consists of a monochrome, white-phosphor CRT with a broad emission spectrum and an electronic-controlled switched color filter on the faceplate. It is interesting to note that earlier color TV designs of the 1940's briefly toyed with a mechanical color-filter wheel rotated in front of the tube – however the commercial implementation was challenging, and eventually the shadow mask won the competition for the large, direct-view color CRTs. Unfortunately the shadow mask approach is unsuitable for miniature CRTs, so that need was not properly addressed. One solution was provided by Tektronix in the 1980's. Tektronix developed a Liquid Crystal Shutter (LCS) based on pi-cells that make use of a nematic LC wave plate (polarization retarder) (Bo, 1984). This provides a totally solid-state solution to the color shutter. Unfortunately the LC Shutter transmittance efficiency is quite low (less than 10%) is typical, which limits the LCS use to low-ambient luminance level.

A second major limiting factor of shutter technology in FSC displays is the presence of visual artifacts. Among these artifacts is flicker sensitivity creating a color break effect associated with rapid head and/or eye movement, which is universally present in military aviation applications. The flicker sensitivity is associated with eye movement. Actual eye movement can be divided into smooth pursuit, with the maximum velocity of 20 to 40 degrees/second, and saccade movements, with the velocity of 300 to 500 degrees/second. Flicker sensitivity was also shown to have a color dependency, with green areas being most sensitive to flicker (at around 150 Hz) and with lower sensitivity for red (around 30 Hz) and blue (around 35 Hz) (Yamada, 2000). A comprehensive overview of flicker sensitivity and other FSC display visual artifacts can be found in Mikoshiba (2000).

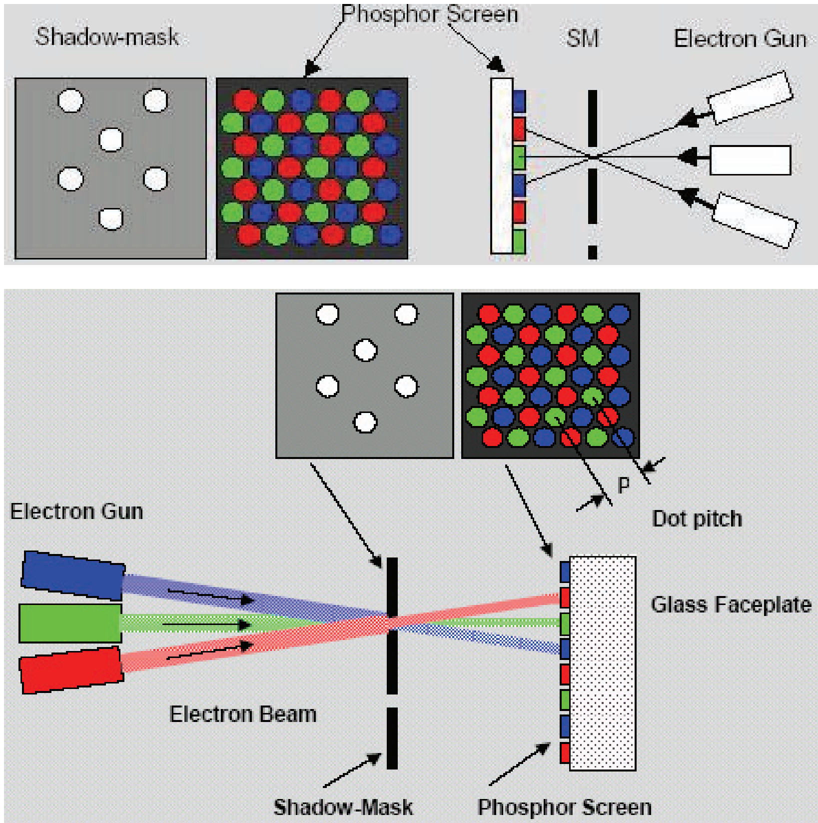


Figure 4-29. Diagram of shadow mask operation (Fujioka, 2001).

Plasma

Plasma display panels (PDPs) are emissive, producing light when an electric field is applied across an envelope of gas. Initially, plasma displays were monochrome, limited to only a few colors. However, in recent years, full-color plasma displays have become rather commonplace.

Color PDPs have a simple construction, basically consisting of two thin sheets of glass separated by a few hundred microns. The space between the sheets of glass is filled with cells containing rare gases (e.g., xenon or neon). Each cell is coated on the bottom in red, green or blue phosphor. Electrodes can be found at the top and bottom of each sheet of glass, or "substratum" (Figure 4-30).

Plasma generates light when an electric field is applied to selected cells (depending on the image) across the gas-filled sachet. Gas atoms are ionized and emit photons when returning to the unexcited state. Plasma technology is most effective for large-area, direct-view displays. It is unlikely that plasma technology will find its way in the HMD application in the near future.

Vacuum fluorescent

Vacuum fluorescent displays (VFDs) (Figure 4-31) are flat vacuum tube devices that use a filament wire, control grid structure and a phosphor-coated anode. They are emissive displays. The monochrome zinc oxide and zinc (ZnO:Zn) phosphor of the vacuum fluorescent displays is very efficient and well proven in automotive applications for both text and graphics. VFDs use a wire filament and a phosphor-coated anode. Active matrix addressing has been demonstrated experimentally.

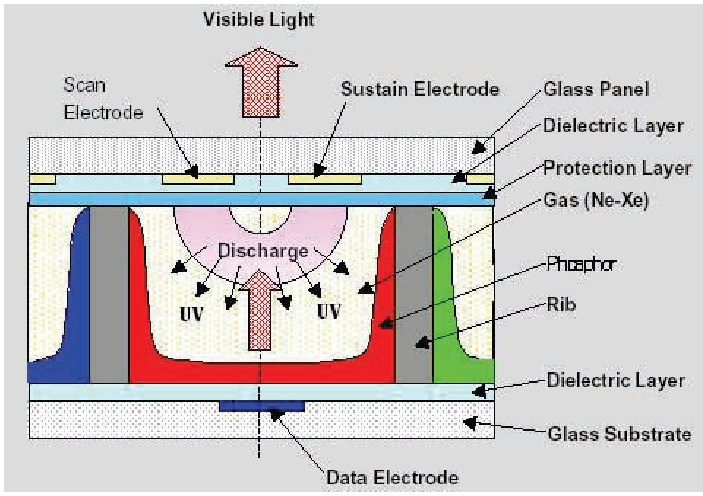


Figure 4-30. Operation of a plasma display (Fujioka, 2001).

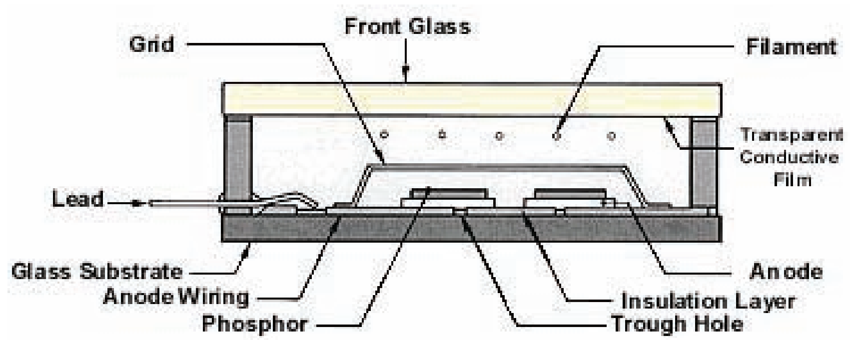


Figure 4-31. Operation of a vacuum fluorescent display (Source: Futaba)

VFD main advantages are:

- Wide temperature range: -40°C to +85°C
- Wide viewing angle with uniform luminance across the display (no hot spots)
- High multiplexing is possible without viewing angle reduction
- Long lifetime and reliability.

However, this technology is mostly applicable to direct-view panels and to date has shown little potential for HMD applications.

Field emission

The emissive Field Emission Display (FED) uses a matrix of point emitters (electron sources) that can be individually addressed (Spindt et al., 1976). Field emission refers to the emission of electrons from the surface of a conductive substrate in a vacuum under the influence of a strong electric field. Light is generated when the electrons strike a phosphor screen. In a sense each pixel acts as a miniature electron gun for its own phosphor dot (Figure 4-32).

Figure 4-32. Operation of a Field Emission Display (Pixtech, Inc.).

However, high luminance is achieved only with an anode voltage in the order of 10kV to allow the use of traditional CRT phosphors; this is one of the remaining fundamental system problems. Both full-gray scale monochrome and full-color FEDs have been developed.

In the late 1990s, this technology seemed destined to succeed big in the marketplace; the thrust on this technology has returned to the research laboratories and is mostly focused on a) improvements in low-voltage, high-efficiency phosphors (Kim, 2000) and b) reliability of the field emission sources, whether from randomly orientated carbon nanotubes (Wang, 1998) or other technology. Another major hurdle for FEDs is the continuing drop in cost of competitive LCDs.

For further information and in-depth research results on phosphors, readers are encouraged to visit the Phosphor Technology Center of Excellence (PTCOE), operating under the Advanced Technology Development Center of Georgia Institute of Technology at the web address: http://www.ptcoe.gatech.edu.

Electroluminescence (EL)

The mechanism of electroluminescence (EL) is the non-thermal conversion of electrical energy (electric current) into luminous energy (light). In EL devices light is generated by impact excitation of a light emitting center (activator) by high energy electrons in materials like ZnS:Mn (Figure 4-33).

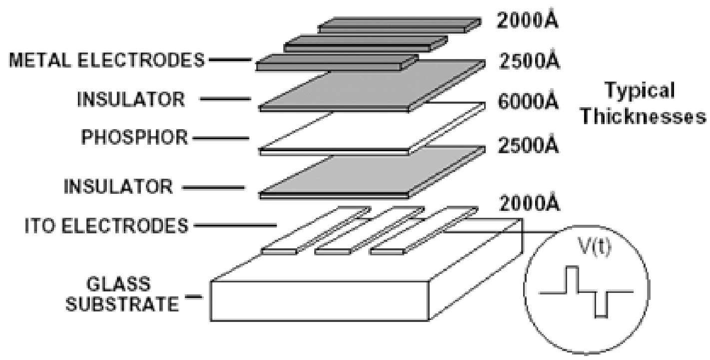


Figure 4-33. Diagram of electroluminescence operation (Source: Planar Systems, 1998).

Due to their compact, self-emissive, low power and weight, and rugged characteristics, EL displays are well suited for HMD applications, in particular for wearable applications. However, the luminance output of these displays is insufficient for avionic applications. To generate higher resolution in a small package the driver electronics was integrated onto the wafer that forms the substrate for the display, with the light-emitting structure on top. The Active Matrix EL (AMEL) thus created overcomes size limitations of the traditional technology. AMEL displays with up to 1000 lines-per-inch (LPI) of resolution have been demonstrated (Khormaei, 1994;

1995). Using silicon on insulator (SOI) wafers to improve driver isolation (80 VAC is required for pixel drive) had enabled fabrication of 2000 LPI test devices (Arbuthnot, 1996).

One of the most challenging tasks for the EL technology is achieving full-color. The blue phosphor in particular has low efficiency; this is still work in progress. EL displays also have been employed as backlights for non-emissive displays, e.g., liquid crystal displays. A comprehensive history of evolution of the EL technology is presented in Krasnov (2003).

Light-emitting diode (LED)

LEDs have been around since the 1950's. Their operation is based upon semiconductors of two types: p-type or n-type, depending upon whether dopants pull electrons out of the crystal, forming "holes", or add electrons, respectively. An LED is formed by p-type and n-type joining the two materials. When a voltage is applied to the junction, electrons flow through the structure into the p-type material, and holes appear to "flow" into the n-type material. An electron-hole combination forms, releasing energy in the form of light. This is a very efficient electricity-to-light conversion mechanism.

LED displays can range from a single status indicator lamp to large-area x-y addressable monolithic arrays. Fabrication of high-density arrays as required for high resolution HMD display panels is challenging; they suffer from optical cross-coupling, mechanical complexity and heat transfer limitations. However, the high light generating efficiency of LEDs makes them very effective as backlights for other non-emissive displays.

Organic light-emitting diode (OLED)

One emissive FP technology that has made great progress in the past decade is the organic light-emitting diode (OLED). This technology uses a wide class of organic compounds, called conjugated organics that have many of the characteristics of semiconductors. They have energy gaps of about the same magnitude, they are poor conductors without dopants, and they can be doped to conduct either by electrons (n-type) or holes (p-type). Initially, these materials were used as photoconductors, to replace inorganic semiconductor photoconductors, such as selenium, in copiers. In the 1980's, it was discovered that, just as with crystalline semiconductors, p-type and n-type organic materials can be combined to make LEDs when an electric current passes through a simple layered structure.

OLEDs are devices that sandwich carbon-based films between two charged electrodes (usually glass), one a metallic cathode and one a transparent anode. When voltage is applied to the OLED cell, the injected positive and negative charges recombine in the emissive layer and generate electroluminescent light.

A typical OLED of the Eastman Kodak Company variety (and practically all OLED manufacturers have licensed Eastman Kodak patents for the technology) is formed by starting with a transparent electrode, which also happens to be a good emitter of holes, e.g., indium-tin oxide (ITO). The ITO electrode is covered with a thin layer of copper phthalocyanine, which passivates the ITO and provides greater stability (Figure 4-34. Then, the p-type material, e.g., naphthaphenylene benzidine (NPB), is deposited, followed by the n-type material, e.g., aluminum hydroxyquinoline (Alq). Finally, a cathode of a magnesium-silver alloy is deposited. All of the films can be applied via evaporation, making fabrication very simple. Electrons and holes recombine at the interface of the n-type and p-type materials and emit, in this example, green light.

One manufacturer committed to the development of active matrix OLED-on-silicon microdisplays is eMagin Corporation, Hopewell, NY (eMagin, 2007). Based on its own patent portfolio as well as licenses from Eastman Kodak, eMagin offers the advantages of integrated silicon chip technology over thin-film transistors – lower weight, higher efficiency, more compact display modules.

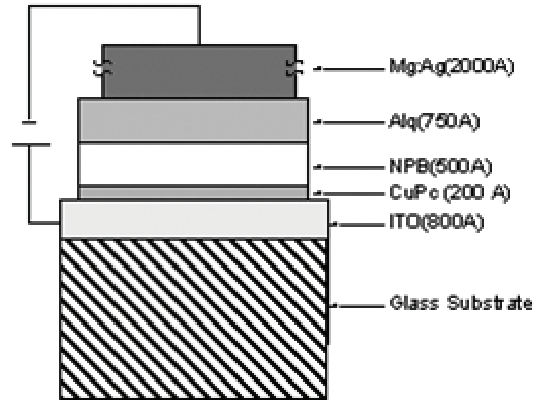


Figure 4-34. Diagram for a typical organic light-emitting diode (Howard, no date).

OLEDs are emissive devices, creating their own light rather than directing light from a second source like liquid crystal-based displays. As a result, OLED devices require less power and can lead to more compact device designs. OLEDs emit light in a Lambertian pattern, appearing equally bright from most forward directions. So moderate pupil movement does not affect brightness or color, and the eye can maintain focus more comfortably, even for extended periods of time.

OLEDs have wide acceptable viewing angles (160° is typical) and are thinner than LCDs (about 1.8 mm [0.07 inches] compared with 6 to 7 mm [0.236 to 0.276 inches] for the LCD). In addition they are low voltage devices; 5-10 Volts is sufficient to cause a very bright emission. This characteristic drives manufacturing costs down, as low voltage circuits are easier and less expensive to fabricate. With no need for backlights and extra heaters or coolers, OLEDs consume less power than other near-eye displays of similar size and resolution.

Other advantages of the technology are:

- High-speed refresh rates – OLEDs are many times faster than LCDs; even faster than CRTs; and can support refresh rates to 85 Hz.
- OLEDs do not require use of polarizers which makes for simpler and more light-efficient optical design.
- Wide operating temperature range – OLEDs turn on instantly and can operate between –55ºC and 130ºC. This is an especially important characteristic for military applications.

The eMagin's OLED display was selected by Rockwell Collins for the initial version of the U.S. Army's Land Warrior HMD program.

Liquid crystal (LC)

Despite the recent "novelty" of LCD products in the market, liquid crystal materials have a long history, dating back as early as the 1880's. Numerous excellent volumes dedicated to LCD's are available to the interested reader (e.g., Kelker, 1988; Tannas, 1985; Wu, 2001). The following is a short list of milestones in the development of LCD:

- 1880's - Liquid crystal phase discovered
    - 1888 Reinitzer, R.
    - 1889 Lehmann, O.

- 1904 - Term “liquid crystals” coined by Otto Lehmann (Sluckin, Dunmur, and Stegemeyer, 2004)
- 1960’s - Electro-optic effect explored
- Early 1970’s - Stable LC materials developed; LC operation modes developed
- Late 1970’s - Ferroelectric effect explored; thin-film transistor (TFT) invented
- 1980’s – Super twisted nematic (STN), ferroelectric liquid crystal (FLC), TFT-LCD demonstrated
- Mid 1980’s - manufacturing infrastructure being built
- 1990’s - Dramatic performance improvements. Dual scan STN. Viable manufacturing yields, LCD monitors overtake CRTs in desktop PC’s. Laptop PCs start the mobile computing era
- 2000’s - Consumer market penetration: High Definition Television (HDTV), mobile communications; plethora of new applications

Liquid crystal is a state of matter intermediate between solid and amorphous liquid. LC molecules are rod-shaped organic compounds with orientation order (like crystals), but lacking positional order (like liquids). LC materials exist in three main classes and are differently arranged in these different phases as defined by the internal molecular structure: *Nematic*, s*mectic* and *cholesteric*. Each have well defined and very different properties (Figure 4-35) (Wu, 2003):

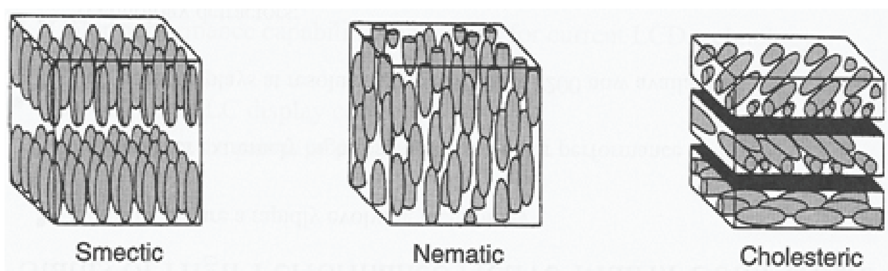


Figure 4-35. LC Diagram of Internal Molecular Structure (Wu, 2003)

- Smectic C (Ferroelectric) LCs (Figure 4-35, left)
  - Layered structure with positional order in one dimension
  - Bistable characteristic, with fast response time (a few µs)
  - Limited gray scale; Thin (<1-µm) cell gap
  - Sensitive to DC voltages

Note: LC materials with smectic A and B structure are too symmetric to allow any vector order, such as ferroelectricity and have not found a display application at this time.

- Nematic LCs (Figure 4-35, middle)
  - Molecules tend to be parallel, but their positions are random
  - Uniaxial; Simple alignment (buffing); Good gray scale;
  - Low drive voltage; Slow (tens to hundreds of ms) response time
  - Mainstream liquid crystal display material
- Cholesteric LCs (Figure 4-35, right)
  - Distorted form of nematic phase in which the orientation undergoes helical rotation
  - Helical structure
  - Bistable memory; very low power displays
  - High luminance efficiency as do not require use of polarizers
  - High driving voltage 20-40V is common

Note: Cholesteric LCs have slow response time and are not usable for real-time video displays. Their market niche is signage, large panel indicators and similar (Figure 4-36).

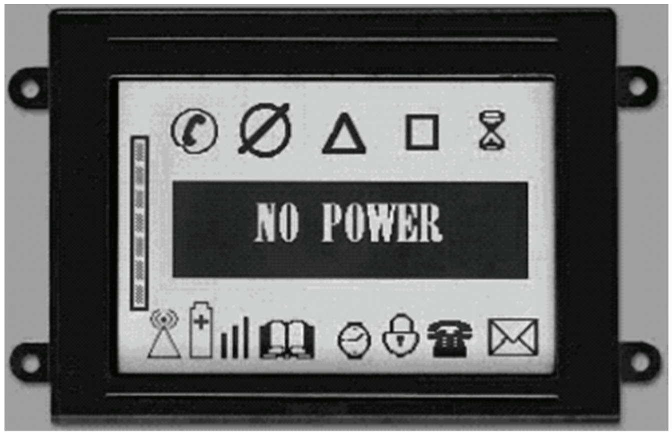


Figure 4-36. Display using cholesteric LC (Source: Kent Display, Inc.)

Properties of LCs are generally anisotropic because of their ordered molecular structure, and ordering leads to anisotropy of mechanical, electrical, magnetic properties, and optical properties (e.g., birefringence).

LCD addressing methods

Display performance is strongly dependent on the addressing method employed (i.e., method of activating individual pixels). The following main options are available for addressing a LC matrix of X columns and Y rows (Figure 4-37):

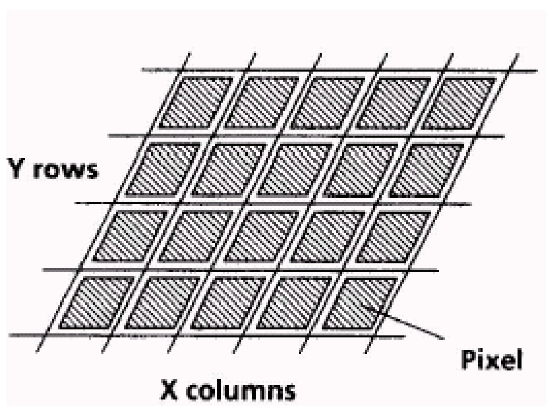


Figure 4-37. An XY matrix display, consisting of X columns and Y rows.

Direct drive

Direct addressing requires X ×Y electrical connections, and each display segment (or cell) is addressed independently. Also each segment requires continuous application of voltage or current to the display element. The approach is simple, low cost, but is limited to low resolution applications, not exceeding approximately 50 pixels/inch. Its use remains largely restricted to segment displays, of the type shown in Figure 4-38.

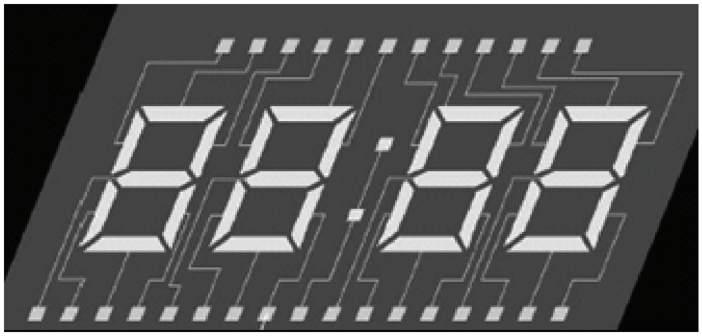


Figure 4-38. Seven-segment LC display.

Passive matrix (PM)

This matrix-type (row and column) addressing has the advantage of minimizing the number of drivers required. It addresses a total of Y (rows) × X (columns) pixels, using only X+Y electrical connections, but at the cost of adding electronic complexity in the drive circuitry. The addressing electrodes are arranged as perpendicular stripe electrodes, which cross each other at each pixel. One row in the matrix is selected by the scanning electrode and the pixels along this line are synchronously addressed by the column signals. In every multiplexing cycle, each row is selected on during 1/Y of the total cycle time T. The driving voltage is defined as the difference between the row and column voltage and is therefore bipolar.

The resolution is limited by the fact that the luminance-drive voltage dependency for LC material is not sharp enough, which severely limits the multiplexing ratio possible (Figure 4-39).

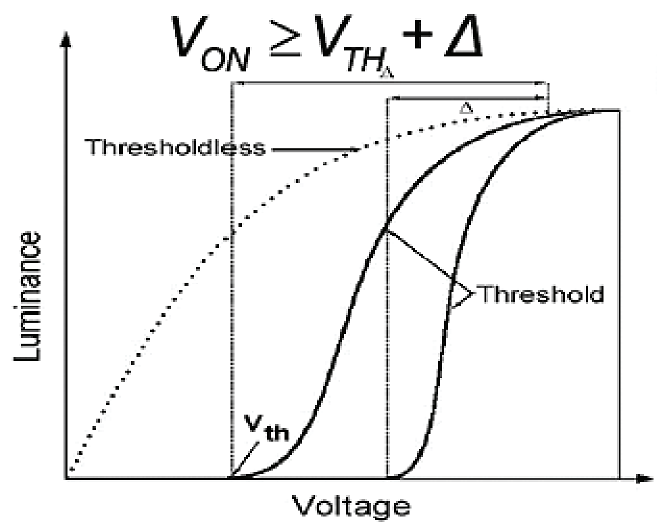


Figure 4-39. Luminance-drive voltage dependency for nematic liquid crystal material.

Active matrix (AM)

The tradeoff between contrast and resolution in PM addressing is a result of requiring the LC to handle both transmission modulation and addressing tasks. Active matrix (AM) addressing provides a way of avoiding this tradeoff. In AM addressing, each individual subpixel (R, G. B) is independently addressed by a thin film transistor (TFT), see Figure 4-40. The highly non-linear switching characteristic of the transistors driving the pixels, eliminates the problems of ghosting and slow response speed. The result is response times of the order of 10-15 ms, minimizing the smear. By controlling the transmission of each individual pixel and doing it independently of all other pixels, AM addressing effectively eliminates pixel crosstalk from limiting the multiplex ratio, enabling large, high-resolution displays. The complete matrix of transistors is produced on a single silicon wafer.

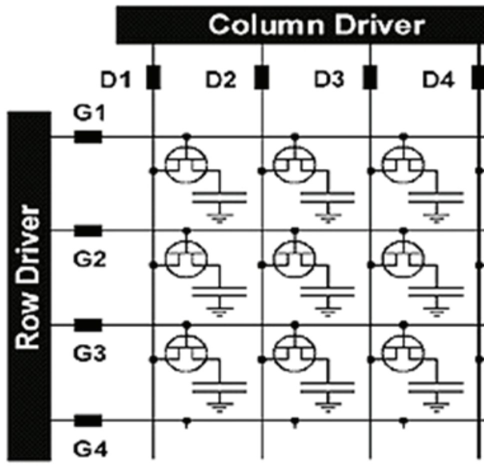


Figure 4-40. Active matrix addressing (Wu, 2003)

In the last decade and a half, as FP technologies have come of age, LCDs have emerged as a major rival to CRTs as the display technology of choice. AMLCDs have become the preferred approach for see-through military HMD applications. LCDs overcome a host of CRT weaknesses. While LCD technology is not without its own disadvantages, its impact on display applications cannot be underestimated

In its most simplistic form, an LCD consists of two substrates that form a "flat bottle" containing the liquid crystal mixture (Wu, 2001). The inside surfaces of the bottle or cell are coated with a polymer that is buffed to align the molecules of liquid crystal. The liquid crystal molecules align on the surfaces in the direction of the buffing. LCDs exist in a variety of configurations, differing primarily by the electro-optical effect the crystal exhibits. For twisted nematic (TN) LCDs, the two surfaces are buffed orthogonal to one another, forming a 90° twist from one surface to the other.

The LCD glass has transparent electrical conductors plated onto each side of the glass in contact with the liquid crystal fluid and they are used as electrodes. These electrodes are made of ITO. When an appropriate drive signal is applied to the cell electrodes, an electric field is set up across the cell. The liquid crystal molecules will rotate in the direction of the electric field (Figure 4-41, top). The incoming linearly polarized light passes through the cell unaffected and is absorbed by the rear analyzer. The observer sees a black character on a sliver gray background. When the electric field is turned off, the molecules relax back to their 90° twist structure (Figure 4-41, bottom). This is referred to as a positive image, reflective viewing mode.

LCDs are non-emissive displays. They produce images by modulating ambient light, which can be either reflected light or transmitted light from a secondary, external source (e.g., a backlight).

One of the latest advances in LCD technology is ferroelectric LCDs (FLCDs). The existence of ferroelectric liquid crystals was first suggested by Meyer in the mid 1970’s (Meyer, 1977). A further refinement of the principle came a few years later (Clark, 1980). FLCDs utilize the intrinsic polarization inherently exhibited by the chiral tilted smectic LC, which is the defining characteristic of ferroelectric materials. These liquid crystal molecules are endowed with a positive or negative polarity in their natural state, even without the application of an electric field. When an electric field is applied, the optical axis assumes a uniform direction throughout the

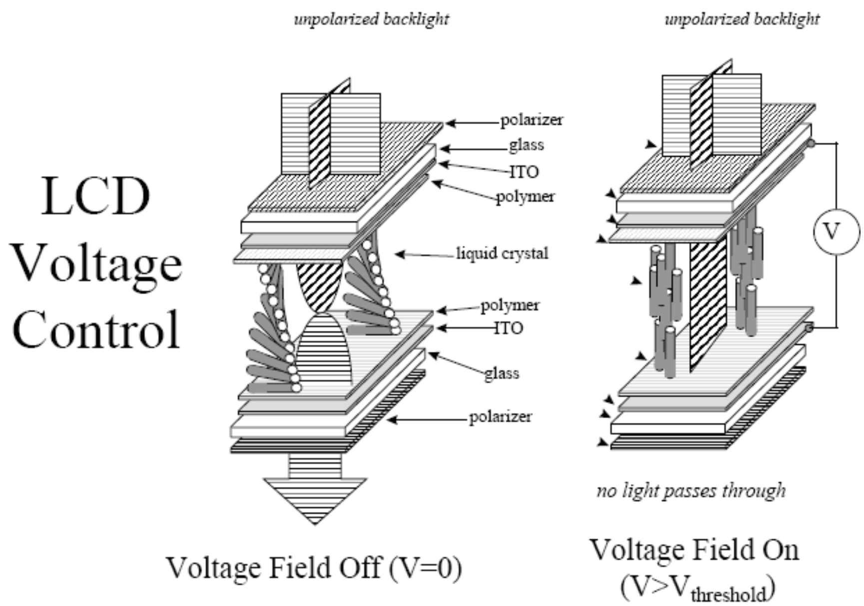


Figure 4-41. Diagram of LCD operation during the application (right) and removal (left) of an electric field (Hel-Or, 2007).

crystal layer. When the polarity of the electric field is reversed, the optic axis rotates 45°. This gives the cell two stable states that are determined by the polarity of the applied electric field. By selecting the appropriate thickness of the FLC layer, it will function as an electrically switchable half-wave plate. This makes FLCs ideally suited to electro-optic applications (Surguy, 1998).

Interest has focused on FLCDs because they offer a number of characteristics that differ from conventional LCDs:

- Memory - Ferroelectric display images are not lost when the power is cut; the image remains intact. Since the arrangement the liquid crystal molecules had when voltage was last applied is retained, the number of scanning lines can be increased without sacrificing contrast quality.
- High response rate - Very high-speed displays are possible. Ferroelectric LC materials show very fast switching time, of the order of 20 to 50 μsec. These speeds are more than 3,000 times faster than TN LCDs.
- Wide viewing angles - Viewing angle limitations are greatly reduced. Since contrast does not change depending on the viewing angle, high resolution, large-scale LCDs are possible.
- Lower cost - Ferroelectric LCDs do not require expensive switching elements like AM drive systems (as TFT LCDs do), making large-scale high-resolution displays with large information capacity possible using simple passive matrix addressing.

Color in FLCDs is achieved using FSC. Using this technique the panel illumination is continuously cycled from red to green to blue rapidly enough so the human eye integrates the three colors sequentially to see full-color on each individual pixel. In contrast, nematic LCD displays achieve color on each pixel by spatially dividing it into 3 subpixels with each subpixel being entirely covered by a red, green, or blue filter. These subpixels are spaced very close together so the human eye integrates the three colors spatially to see full-color.

By using FSC to generate color in the temporal, rather than the spatial domain, FLC displays do not require three color filters for each pixel. This results in improved resolution, light efficiency and reduced costs.

One disadvantage of FLCDs is that the chiral smectic part of the LC is both temperature and mechanical shock sensitive. FLCs tend to revert to their natural helical structure when subjected to mechanical shocks. This problem, although still a concern especially for military applications, has been largely solved for miniature displays. In addition, their temperature operational range is very narrow, 10°C to 40°C, which may restrict their applicability in the military environment.

These displays are still in the research and development stage, but expectations are high that this technology will reveal a dramatic new LCD potential. Future challenges lie in correcting manufacturing difficulties related to improving ruggedization and more effectively controlling spacing between the two substrates. This spacing should not exceed 2 µm maximum in order to remove the helical structure that will otherwise cancel the intrinsic polarization effects, and product development (Mosley, 1994).

The shift from CRT displays to LCD displays greatly changes the nature in which display images are evaluated. One change is in how image defects impact image quality. This is a result of the pixilated nature of LCDs and other FP technologies. Whereas CRTs have one control structure – modulation control in the horizontal scan, but fixed vertical positions of the scans, the LCD has independent control structure for each individual pixel –total control over vertical and horizontal positions. The CRT functionally time-shares its electron gun control but in the process introduces a whole array of geometric and focusing errors as a consequence of its deflection scanning mechanism.

For matrix displays, which eliminate geometrical errors, sufficient gray scale and color rendition are challenges. For LCD the display process is inherently non-linear, involving at its simplest E^2 (square of the local electric field in the LC medium). Actual results also depend on the optical structure, the illumination and viewing angle. The human eye is quite forgiving to full field changes in brightness, color and even contrast – it quickly adapts to the "average" conditions present. But, even minor gray scale and color errors may be objectionable. As they occur possibly closely adjacent in the same image, and the eye does not compensate for them.

Response time is still an LCD problem, which is aggravated at reduced temperatures in field applications. The slower temporal response causes image contrast under dynamic conditions to be lower than corresponding values recorded with static photometric measurements. Rabin (1995) assessed display response time effect on visual acuity by comparing two HMDs: one CRT-based, the second AMLCD. The main conclusion was that at low to moderate rates of visual stimuli presentation, there was no significant difference in dynamic visual performance between the two technologies. However, at higher presentation rates, dynamic visual performance was significantly reduced when the AMLCD was used. Quantitatively the results were expressed as:

- Contrast sensitivity function (defined as 1/ contrast threshold) is the same for temporal frequencies of up to 2 Hz. Beyond 2 Hz, the fall-off rate of the AMLCD is significantly faster - a 2X difference in CRT favor was recorded at 8 Hz and a 3X difference at 16 Hz.
- Target recognition as function of target duration was the same up to 200 ms; below this the AMLCD performance drops almost linearly with target presentation time - reaching 4X in CRT favor, at a target duration of 30 ms.
- For fast moving targets, the AMLCD HMD is even worse – 5X for target velocity of 20° per second in favor of the CRT HMD.

One other major disadvantage for the AMLCDs is their low optical transmission; typically in the range of 8% to 15% for monochrome and only 3% to 5% for color devices. This increases luminance requirements for the backlighting and the optical design, with corresponding increase in electrical power requirements and heat load.

The development of miniature AMLCDs for use in HMDs has been challenging. Seeded by military funding, success has been driven by commercial applications. A major manufacturer of miniature AMLCD displays for both the HMD and commercial communities is Kopin Corporation, (Westborough, MA). Since their development of a class of transmissive LCDs, advertised as CyberDisplay® products, in 1997, Kopin Corporation has shipped more than 20 million displays. These displays have been used in consumer electronics (camcorders, digital cameras) and for advanced night vision goggles and thermal weapon sights programs for the U.S. Army.

Kopin Corporation's CyberDisplay® uses single-crystal silicon transistors that enable pixels typically 15 µm square and of a pixel density exceeding 1600 LPI (Figure 4-42). To construct the transmissive display from opaque silicon, Kopin Corporation uses a patented lift-off process to transfer a very thin IC layer onto a glass plate (Werner, 1993). The success of miniature AMLCD development has depended on thin-film technology that removes the active circuit from the silicon wafer and transfers it to the display glass substrate. One approach has been one pioneered by the Massachusetts Institute of Technology (Cambridge, MA) and commercialized by Kopin Corporation under the trade name Isolated Silicon Epitaxy™ (ISE). This process relies on forming a release layer on the silicon wafer and epitaxially growing the active silicon layer on top of the release layer.

The CyberDisplay® SXGA low-voltage ruggedized (LVR) (Kopin Corporation, 2007) is a full-color SXGA display in a 0.97-inch (24.6-mm) diagonal package for use in targeting, multi-spectral, image fusion, simulation and training, and medical head-mounted systems The LVR's low-voltage architecture results in power consumption of less than 200 mW, which will extend battery life in man-portable applications. Power requirements for display, backlight, application-specific integrated circuit (ASIC) drive electronics and backlight are less that 1W.

Another version is the CyberDisplay® 1280MR, a monochrome SXGA display for thermal imaging applications. This display is available in two versions: the standard twisted nematic (TN) AMLCD and the multi-domain vertical alignment (MVA) display. The MVA display offers a normally black image with high contrast ratio (greater than 300:1) for I^2 and thermal night vision applications.

Digital light processing (DLP®)

The digital light processing (DLP®) display concept, originally known as the Digital Micromirror Display (DMD), was invented in 1987 by Dr. Larry Hornbeck of Texas Instruments, Dallas, TX. The heart of the display is an electronic chip that contains a rectangular array of approximately 2 million hinge-mounted microscopic mirrors; each of these "micromirrors" measures less than one-fifth the width of a human hair (Texas Instruments, 2007). Each mirror corresponds to a single pixel. The display modulates incident light by movement of the individual micromirrors. With an appropriate light source and a projection lens, the display's mirrors reflect the desired image onto a screen or other surface.

Figure 4-43 illustrates the architecture of a single pixel, showing the mirror as semitransparent so that the structure underneath can be observed. The mirrors are held in place on two corners and are free to twist around one axis by ± 10°. When the mirror rotates to its *on* state (+10°), light from a projection source is directed into the pupil of a projection lens, and the pixel appears bright on a projection screen. When the mirror rotates to its *off* state (-10°), light is directed out of the pupil of the projection lens, and the pixel appears dark. Thus, the optical switching function is simply the rapid directing of light into or out of the pupil of the projection lens.

Both grayscale and color are possible with DLP®. Up to 1024 shades of gray can be generated. Color is achieved via a color wheel that filters white light from a lamp source as it travels to the surface of the DLP® chip; converting the white light into red, green, or blue. Specifications for the DLP® chip claim that at least 16.7 million colors can be produced.

It is the human eye's temporal integration time that allows this large color gamut. For example, to produce a purple hue, a mirror would only reflect red and blue light to the projection surface.

DLP/DMDs offer several advantages over other technologies: small volume and weight, high luminance and contrast ratio, and a less visible pixel grid (as compared to LCDs). Based on these advantages, several HMD applications have been suggested (Preston, 2002).

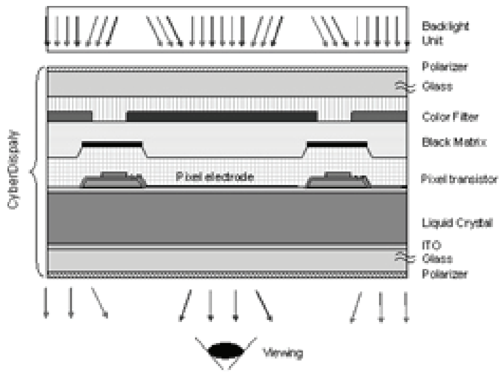


Figure 4-42. Diagram for Kopin Corporation's CyberDisplay® transmissive LCD (Kopin Corporation, 2006).

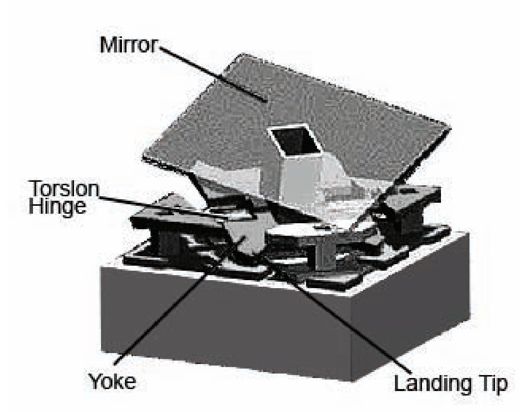


Figure 4-43. Depiction of a single micromirror pixel (Texas Instruments, 1997).

Laser

The highest luminance image source available is the laser. Making use of the persistence of vision characteristic of the eye, lasers are used in a scanning mode to produce an image in the manner of CRTs. Rather than an electron beam, a laser beam is scanned in two dimensions, with the beam intensity modulated at every pixel (Rash, 2001). When scanned at frequencies greater than 60 Hz, a flicker-free image is produced. In addition to high luminance, laser-based displays are capable of a wide color gamut with excellent color saturation.

One of the original versions of these displays is known as a virtual retinal display (VRD). The VRD modulates the scanning laser beam with video information, producing a raster image placed directly onto the retina of the user's eye. The VRD may also include a depth accommodation cue to vary the focus of scanned photons rapidly so as to control the depth perceived by a user for each individual picture element of the virtual image. Further, an eye tracking system may be utilized to sense the position of an entrance pupil of the user's eye, with the detected pupil position being used to move the beam so as to be approximately coincident with the entrance pupil of the eye (Furness and Kollin, 1995).

Also known as the Retinal Scanning Display (RSD), the VRD concept originated at the Human Interface Technology Laboratory at the University of Washington (Furness and Kollin, 1995) and is now being developed and commercialized at Microvision, Inc., Redmond, Washington. The RSD (or VRD) offers high spatial and color resolution and high luminance, fundamentally limited only by eye safety considerations. It does not require the use of a display screen. Color imagery is achieved by the use of low-power red, green, and blue lasers.

Due to optical constraints imposed by inherent design characteristics, the final image in HMDs that use laser sources is not scanned directly onto the viewer's retina. Instead, an intermediate image must be formed and viewed using an eyepiece. This configuration is no longer a true VRD and is better described as a *scanning laser display*.

A functional block diagram of Microvision, Inc., scanning laser HMD developed for the U.S. Army's Aircrew Integrated Helmet System program (AIHS) is presented in Figure 4-44 (Rash and Harding, 2002). While this diagram is useful for the understanding of the operation of the Microvision, Inc., AIHS scanning laser HMD, it may be more interesting to look at the system from the perspective of how the laser light (energy) traverses the optical path from laser source to the eye (Figure 4-45). This diagram is applicable to both channels. The percentage values reflect the transmission at each functional block. As can be seen, this theoretical power analysis predicts that only 0.48% of each laser's initial power reaches each eye. This is an important prediction, because historically Warfighters have assigned a negative connotation with lasers in the battlespace. Warfighters have been taught to look away from potential laser sources due to their ability to harm the eye. With this HMD, laser energy purposefully is being directed into the eye.

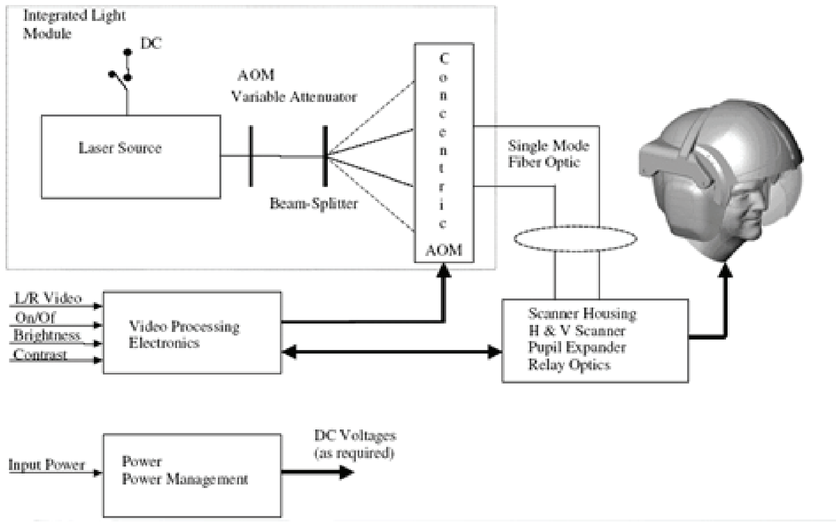


Figure 4-44. Functional block diagram of scanning laser HMD system (Rash and Harding, 2002).

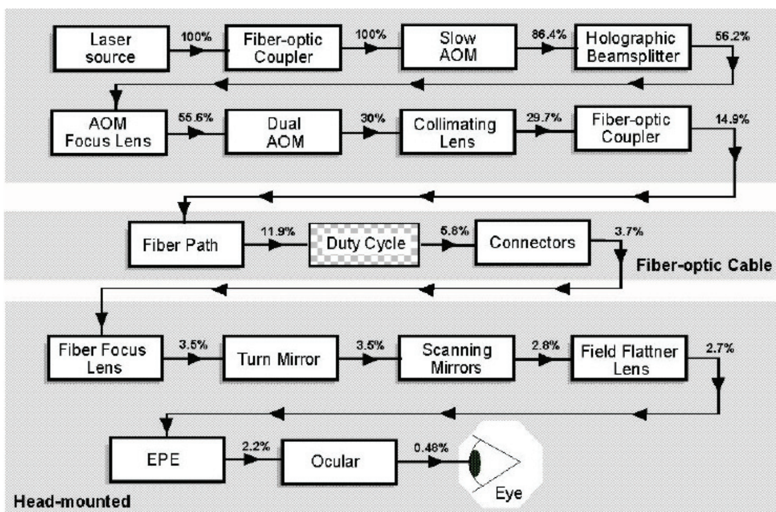


Figure 4-45. Flow diagram for optical path of laser energy (Rash and Harding, 2002).

A 2000s version of this display is presented in Figure 4-46 (top). The predicted high luminance symbology capability of scanning laser source is represented in Figure 4-46 (bottom).

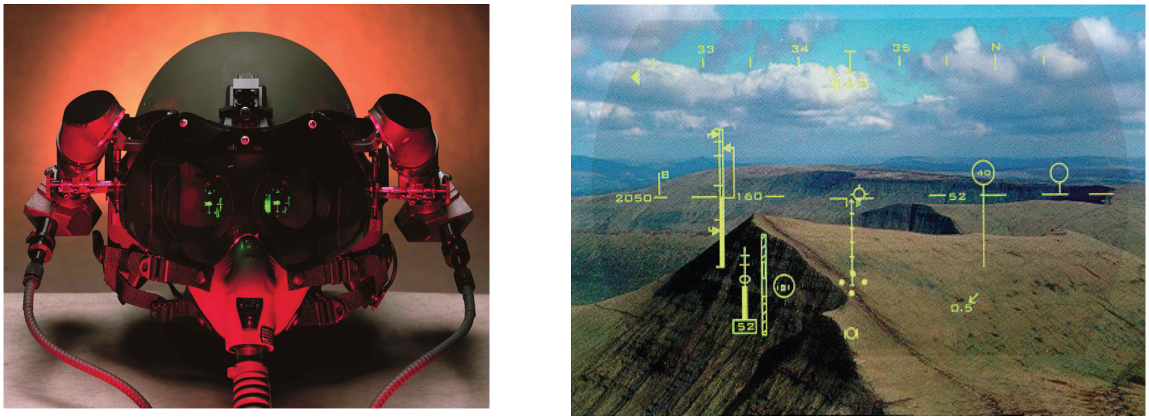


Figure 4-46. Laser scanning HMD (top) and artist's conception depicting the ability of these HMDs to present symbology of sufficient luminance to be seen against daytime backgrounds (bottom) (Source: Microvision, Inc.).

Laser-based systems typically suffer from coherence artifacts. The RSD generates pixels serially, which makes the pixels mutually incoherent; any remaining coherence (e.g., speckle) is typically at subpixel level, hence at high spatial resolution that is beyond the human eye's discerning capability.

The major advantages of using laser sources are high luminance and wide color gamut. The RSD also has the advantage of being an infinitely addressable device, just like the CRT. This allows the option of implementing electronic "imaging warping" to compensate the inherent distortions introduced by the optics for an additional degree of freedom.

Flexible display technologies

Rather than a group of stand-alone technologies, flexible display technologies are new manufacturing approaches to existing FPD technologies (e.g., LCD, OLED). Nonetheless, they are unique enough to warrant their own discussion. Flexible FPD technologies offer many potential advantages including light-weight and robust thin profiles; the ability to flex curve, conform, roll and fold for extreme portability, the ability to be integrated into garments and textiles. Flexible display allow more freedom to design, promise smaller and more rugged devices, and eventually (conceivably) can replace paper (Rash, Harris, and McGilberry, 2005)

An all-encompassing definition of flexible displays is difficult to propose but has been poetically described as "they're like modern art difficult to define but they know one when they see one" (Slikkerver, 2003).

Four different categories have been defined each with its own specific set of requirements for performance and mechanical characteristics. What they all have in common is the replacement of rigid glass substrate by flexible organic or inorganic substrates.

- Flat thin displays: They have the configuration of current FPDs but the thinner substrate will make them thinner and lighter. They are attractive for mobile applications: laptops, cell phones, and Personal Data Assistants (PDAs).
- Curved displays: Curved only once when they are built into a module or device and will maintain the same curvature through their lifetime. They will offer new design freedom – automotive dashboard for instance.
- Displays on flexible devices: They should be at least as flexible as the devices where they will be incorporated - smart cards, or textiles - and should allow frequent bending (Figure 4-47).
- Roll up displays: the quintessential flexible display; requires repeated rolling and unrolling of the display, preferably to a small diameter to allow smaller package.

Figure 4-47. Prototype FOLED (Flexible Organic Light Emitting device) technology, using a flexible substrate (Source: Universal Display Corporation, 2007).

The three more likely technology candidates for flexible displays are the LC, OLED and electrophoresis:

- LCDs are the dominant player in the display market and have already a long history of using plastic substrates. The LC layer is under 10 μm thick, making them suitable for flexible applications. Maintaining the cell gap is the major concern for the flexible LCDs.
- OLEDs, both small molecule and polymer, are possibly the most promising technology. The active layers are typically less than 1 μm thick, which is ideal for flexible displays. Oxygen and water permeation is of particular importance for OLED devices since diffusion of oxygen and moisture through the polymer substrate severely degrades performance and lifetime of OLEDs (Universal Display Corporation, 2007).

On May 24, 2007, Sony Corporation unveiled the world's first flexible, full-color organic electroluminescent display (OLED) built on organic thin-film transistor (TFT) technology (Figure 4-48). OLEDs typically use a glass substrate, but Sony researchers developed a new technology for forming organic TFT on a plastic substrate, enabling them to create a thin, lightweight and flexible full-color display. The 2.5-inch (63.5-mm) prototype display supports 16.8 million colors at a 120 x 160 pixel resolution (80 pixels per inch, 0.318-mm pixel pitch) it is 0.3 mm (0.012 inches) thick and weighs 1.5 grams (0.05 ounces) without the driver (Broadcast Engineering, 2007).

This new 2.5-inch (63.5-mm) OLED display is made of a glass substrate that allows the user to casually bend the screen. Since the display is wafer-thin, one may eventually see these inside magazines as advertisements or perhaps on the back of a cell phone for viewing movies. It uses organic TFT technology to keep clarity intact and to retain its 0.3-mm (0.012-inch) thickness. The screen has a resolution of 120×169 pixels and weighs only 1.5 grams (0.05 ounces). Sony Corporation claims this display will allow for the development of bigger, better, lighter, and "softer" electronics.

- Electrophoresis: Electrophoretic displays rely on a relatively thick optical active layer of about 20-30 microns thick where the liquid with electrostatic particles is encapsulated in a polymer to form a coherent film. The display has a slow response and is not suitable for video but may eventually replace paper (Figure 4-49).

Figure 4-48. OLED flexible display (Source: Sony Corporation).

Figure 4-49. Example of an electrophoretic display (Source: E-Ink Corporation).

Electrophoresis is a phenomenon based on the migration of charged particles when placed under the influence of an electrical field. An electrophoretic display would generally consist of a lower electrode (with protection layers), a layer of charged particles within a medium such as a dielectric fluid, and an upper electrode with protection layers (Rash, Harris, and McGilberry., 2005).

In February 2004, the U.S. Army teamed up with Arizona State University (ASU) to establish the Flexible Display Center (FDC) – a five-year, $43.6 million manufacturing R&D center designed to speed the commercialization of emissive and reflective display and TFT backplane technologies on polymer and metal-foil substrates (http://flexdisplay.asu.edu, 2007). The types of flexible displays the U.S. Army is interested in must be more rugged than those currently demonstrated on glass substrates and require less power. Such displays will be attractive for lightweight, wearable computer applications for use on the battlefield for communication and tactical information access.

Working within the FDC are researchers from a strategically formed team of military, industry and academic partners. Army partners include the U.S. Army Research Laboratory and the Natick Soldier Center. Industry partners include EV Group, Honeywell, Universal Display Corporation, Kent Displays, E Ink, Ito America, General Dynamics, Rockwell Collins, Abbie Gregg Inc. and the U.S. Display Consortium. University collaborators include Cornell University, the University of Texas, and Waterloo University. Additional partners will be added as the center matures.

The agreement has an option for supplementing funding of up to $50 million over a five-year period. The goal of the Army investment in critical issues for flexible displays is to move the timeline for commercial introduction forward and secure flexible technology availability for the Objective Force Warrior.

Dr. John Pellegrino, Chairman DOD Technology Panel on Electron Devices, Director US Army Research Laboratory Sensors and Electron Devices Directorate summarized (USDC Flexible Display Conference, 2003), the technology opportunities FDC is looking to fund:

- Electro-optic materials, emissive/ reflective
    - OLEDs: Full-color, stable materials with low differential color aging
    - OLEDs: Improved Blue emitters
    - OLEDs: Improved thermal stability, operating temperature
    - Electrophoretics: video rates, full-color, stability
- Backplane electronics, Poly-Si, a-Si (n-type only)
    - Deposition, full-color, patterning flexible substrates
    - Roll-to-roll processing - Tools
    - Registration and dimensional control

  - o Process integration
  - o Integrate drivers with flexible active matrix backplane
- Substrates and Barriers: Metal foil/dielectric, flexible glass/plastic, plastic/barrier
  - o Materials/ substrate stability
  - o Barrier coating for substrate
  - o Conformal top encapsulation
  - o Adhesives for flexible top cover
  - o Sustainable under flexing
- Manufacture Integration
  - o Deposition, full-color, patterning flexible substrates
  - o Roll-to-roll processing—Tools
  - o Registration and dimensional control
  - o Process integration
  - o Integrate drivers with flexible active matrix backplane

"Flexible displays are the next revolution in information technology that will enable lighter-weight, lower-power, more-rugged systems for portable and vehicle applications," says Brig. Gen. Roger Nadeau, former Commanding General of the Army's Research, Development and Engineering Command (RDECOM). Flexible displays have a great potential within the military community for almost all direct-view applications. When the flexible technologies will have an impact on microdisplays and, hence, HMDs is not yet defined.

However, it is the large-area displays, not the miniature ones that drive the demand for new displays. Although the revenue per square inch of active display area is higher for microdisplays, the total market for large displays dwarfs that for miniature panels. The explanation for this condition is based on application; there is a greater volume demand for large-area displays (Figure 4-50).

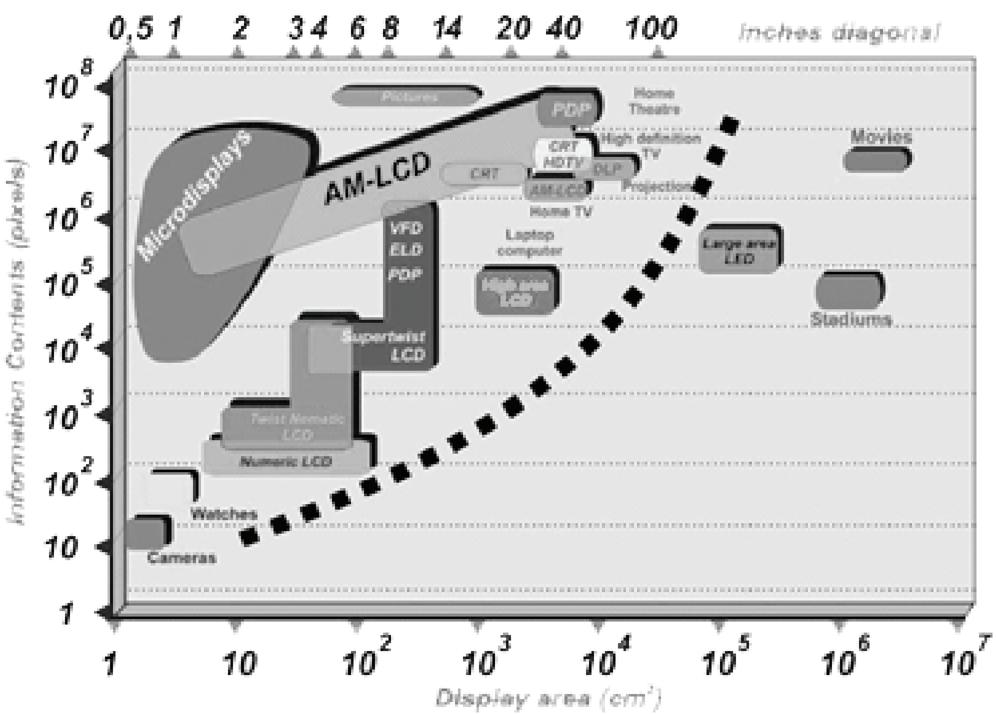


Figure 4-50. Display applications by size (Wu, 2003).

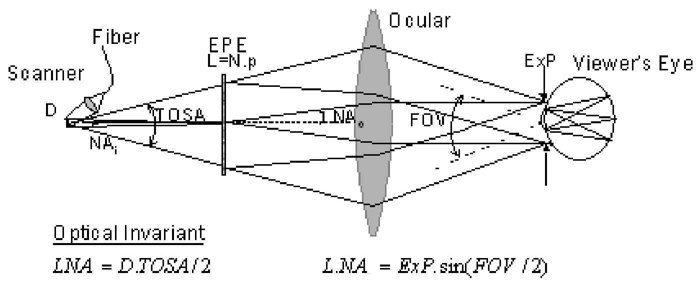
Unique Issues

The integration of image sources into the various optical designs during the development of HMDs has posed a number of unique issues. Two of the most interesting problems, both associated with scanning laser HMD designs, are that of *exit pupil expansion* and *pinch correction*.

Exit pupil expander (EPE)

With scanning laser image sources, a 2-D scanning motion generates an image at an intermediate image plane. In this case, the focusing numerical aperture (NA), formed by the light converging to form the flying spot within the raster, is typically defined by the NA required to form a near-diffraction-limited spot size. As pixel size is typically similar to spot size, the NA exiting from a pixel in the intermediate image plane will have a NA similar to that coming into the intermediate image plane. It is this limited exit NA that results in an exit pupil of approximately 1 to 3 mm. Without the EPE located at the image plane, the beam angles before and after the EPE are equal ($\theta o = \theta i$), hence the exit pupil size (ExP) can be computed using the optical invariant of the system (Figure 4-51):

$$ExP \tan (FOV/2) = D \tan (\theta_{TOSA}/2) \qquad \text{Equation 4-11}$$



where, N = resolution, p = pixel size, and L= image length

Figure 4-51. Ray trace for use of exit pupil expander.

To enlarge the NA of the incoming laser beam to the required exit pupil size (15 mm is standard), an EPE acting as an NA expander is placed at the intermediate image plane between the scanner and the exit pupil. Effectively the EPE divides the optical system into two parts. The function of the EPE is to overcome the limitation imposed by the optical invariant (Melzer, 1998). For HMD systems, once the number of pixels (N), FOV, and exit pupil size requirements are specified, the intermediate image size (L) and the output cone angle (θo) parameters can be computed. The optical invariant can be written separately on either side of the EPE. [Note that the optical invariant before and after the EPE plane does not remain constant in the EPE presence (Urey, 2000a).]

A number of EPE approaches were investigated during the Microvision, Inc. development for a scanning laser HMD for the AIHS program. These include a diffractive (holographic) element and Micro Lens Arrays (MLAs). Figure 4-52 (left) shows a photographic setup for observing the exit pupil for a holographic EPE. The exit pupil appears as a set of beamletts (Figure 4-52, right). Each beamlett contains the entire image (Rash and Harding, 2002). In the AIHS design, a dual MLA approach eventually was employed.

Pinch correction

The adopted scanner architecture is crucial in defining a scanning laser HMD. Scanners for display applications demand high operating frequencies and a large mirror-size x scan-angle product. In addition, the mirror has to remain optically flat during operation under high strain, high acceleration forces, and high thermal loads. The scanning technique usually employed is based on a horizontal scanning (sinusoidal motion) operating at resonance and vertical scanning that is saw-tooth in profile and linearly controlled. The sinusoidal motion of the fast scan combined with the linear motion of the slow scan generates the 2-D raster pattern. Scanner speed non-linearity along the scan line must be corrected electronically. A third scanner is needed to provide raster pinch correction (Powell, 2001; Urey, 2000b; Urey, 2001).

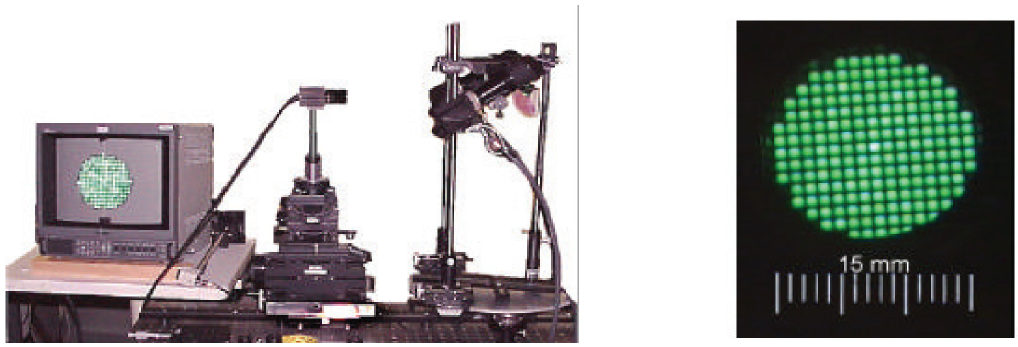


Figure 4-52.The photographic setup (left) and the exit pupil (right) as observed for a holographic exit pupil expander (Rash and Harding, 2002).

In a scanning display (e.g., cathode ray tubes), lines are generally scanned horizontally, and contrast is achieved by increasing or decreasing electron beam intensity as it passes over the display area. The scanning laser HMD is the same with the exception that the scanning area is the retina instead of a phosphor, as in the case of the CRT, and the beam is a photon beam of a laser instead of an electron beam. For each eye, two laser beams are scanned back and forth across the retina. The beams follow a sinusoidal motion, and increasingly diverge from the ideal horizontal raster line as they approach the edge of the raster. Figure 4-53 shows graphs of scanned lines with and without a second-harmonic pinch correction scheme developed by Microvision, Inc. Figure 4-53A shows the case where two lines are being scanned simultaneously without pinch correction. As seen in the figure near the right edge, distance A is shorter than distance B, but line separations are the same in the middle of the display. Also notice that scanned lines cross near the edge where the top line crosses the previously scanned bottom line of the line pair. This crossing reduces the usable active area of the display and thereby reduces system efficiency.

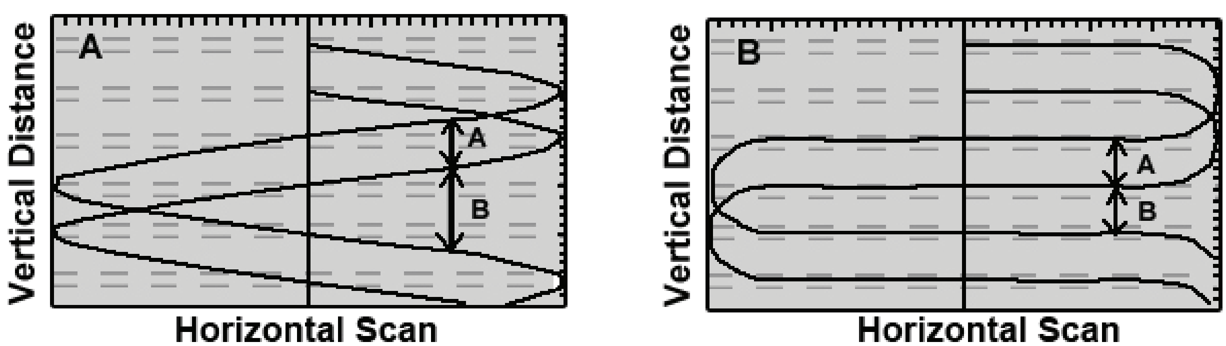


Figure 4-53. Graphs of scanned lines representing dual scans with no pinch correction (A) and dual scans with pinch correction (B). Note the difference between distance A and B in (A), whereas with pinch correction (B), the distances are the same. Original graphs supplied by Microvision, Inc.

Compare this with the pinch correction shown in Figure 4-53B. Here a second harmonic solution has been applied to the scanned lines. Note that near the right edge, distance A and B are now the same. In addition, the line crossing takes place much closer to the edge thereby increasing the usable area of the display thus increasing system efficiency. The full effect of the pinch correction is shown in Figure 4-54.

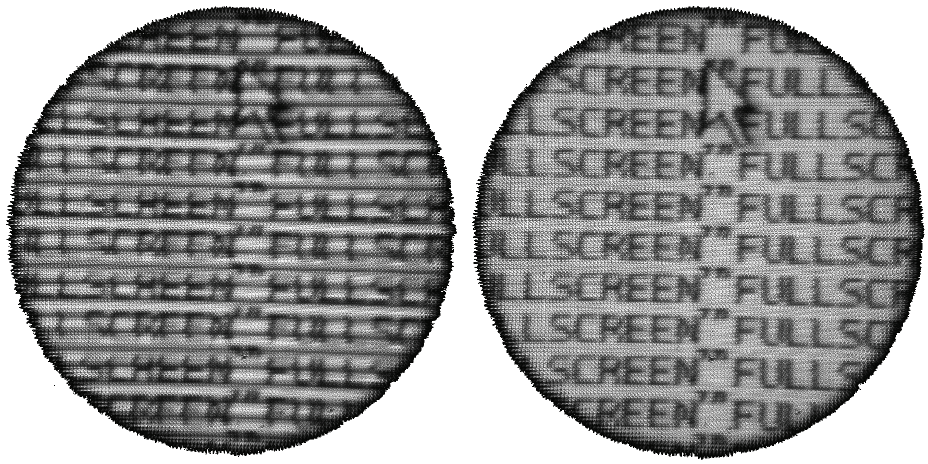


Figure 4-54. Before (left) and after (right) application of pinch correction.

References

Arbuthnot, L., Mendes, J., Sproull, W., Aguilera, M., Aitchison, B., King, C., Dolny, G., Ipri, A., Hsueh, F-L., Stewart, R., Ahmed, S., Solanki, R., Keyser, T., Schlesinger, S., Becker, G., Kagey, D., and Spitzer, M. (1996). A 2000 line per inch active matrix EL display, *Society for Information Display Digest*, 374-377.

Bachman, W.G. (1988). Extended-wear soft and rigid contact lenses: Operational evaluation among Army aviators. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 88-17.

Blackwell, H.R. (1946). Contrast thresholds of the human eye. *Journal of the Optical Society of America*, 36, 624-643.

Blackwell, O.M., and Blackwell, H.R. (1971). Visual performance data for 156 normal observers of various ages. *Journal of the Illuminating Engineering Society (IES) of North America*, October, 3-13.

Bo, P.J. (1984). The Pi-cell: A fast liquid-crystal optical-switching device. *Molecular Crystals and Liquid Crystals*, 113, 329-339.

Braybrook, R. (1998) Looks can kill. *Armed Forces Journal International*, 4, 44.

Broadcast Engineering. (2007). Retrieved July 23, 2007 from http://broadcastengineering.com/ news/sony-flexible-tv-display-0604/.

Buchroeder, R.A. (1987). Helmet-mounted displays, *Tutorial Short Course Notes T2*. SPIE Technical Symposium Southeast on Optics, Electro-optics, and Sensors, Orlando, FL.

Cakmakci, O., and Rolland, J. (2006). Head-worn displays: A review. *Journal of Display Technology*, 2(3), 199-216.

Calves, J.P., and Brun, J. (1978). Colour and brightness requirements for cockpit displays: Proposal to evaluate their characteristics. Presented at the 29th AGARD Avionics Meeting, Paris.

Castellano, J.A. (1992). *Handbook of Display Technology*. New York: Academic Press, Inc.

Clark, N.A., and Lagerwall, S.T. (1980). Submicrosecond bistable electrooptic switching in liquid crystals. *Applied Physics Letters*, 36, 899.

Cook, L., and Patterson, S. (1991). Fiber optics for displays. *Information Display*, 6, 14-16.

Crookes, T.G. (1957). Television images. *Nature*, 179, 1014-1025.

Crowley, J.S. (1991). Human factors of night vision devices: Anecdotes from the field concerning visual illusions and other effects. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 91-15.

Droessler, J.G., and Rotier, D.J. (1989). Tilted Cat helmet mounted display. *Head-Mounted Displays, Proceedings of SPIE,* 1116, 19-26.

eMagin. (2007). Retrieved June 12, 2007 from www.emagin.com

Falce, L. (1992). Dispenser cathodes for CRT's. *Information Display,* 1, 11-14.

Farrell, R.J., and Booth, J.M. (1984). *Design handbook for imagery interpretation equipment.* Seattle, WA: Boeing Aerospace Company.

Ferrin, F. (1997). Selecting new miniature display technologies for head mounted applications. *Proceedings of SPIE, Helmet-Mounted Displays II*, 3058, 115-121.

Fischer, R.E. (1997). Fundamentals of HMD optics. *Head mounted displays: Designing for the user*. New York: McGraw-Hill, 83-116.

Flexible Display Center at Arizona State University. (2007). Retrieved July 1, 2007 from http://flexdisplay.asu.edu/.

Fujioka, T. (2001). Monitor innovation of the future, Presentation October 3, courtesy of Totoku Electric Co., Ltd.

Furness, III, T.A., and Kollin, J.S. (1995). Virtual Retinal Display. US Patent #5,467,104 issued November 14, 1995.

Futaba, (2007). Retrieved July 23, 2007 from http://www.futaba1.de/ index.php?option=com_content&task =view&id=21&Itemid=47

Girolamo, H. (2001). A decade of progress 1991-2001: HMD technology ready for platform integration. *Proceedings of SPIE,* 4361, 43-70.

Godfrey, G.W. (1982). Principles of display illumination techniques for aerospace vehicle crew stations. Aerospace Lighting Institute: Tampa, FL.

Harding, T.H., Martin, J.S., Beasley, H.H., and Rash, C.E. (1996). Figures of merit and performance specifications for the IHADSS and ANVIS. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 96-13.

Hel-Or, H. (2007). Lecture notes: Display technologies. Retrieved June 1,2008 from http://cs.haifa.ac.il/ hagit/courses/ist/Lectures/IST08_Displaysx4.pdf

Howard, E.W. (No date). "Organic Light Emitting Diodes (OLED)" authored by Webster E. Howard Vice-President of Technology of eMagin.

Hur, R., and Pham, T. (2001). White LED. Retrieved June 21, 2007 from http://www.mse.berkeley.edu/classes/matsci102/ F01reports/whiteled.pdf

Illuminating Engineering Society (IES) of North America. (1984). *IES Lighting Handbook*. New York: Illuminating Engineering Society of North America.

International Organization for Standardization (ISO). (1987). Ergonomics of office VDUs: Visual requirements. Geneva: International Organization for Standardization. Draft DP 9241, Part 3.

Kopin Corporation. (2006). About CyberDisplay. Retrieved June 12, 2007 from http://www.kopin.com

Kaiser, P.K., Herzberg, P.A., and Boynton, R.M. (1971). Chromatic border distinctness and its relation to saturation. *Vision Research*, 11, 953.

Kalinowski, J. (2004). *Organic Light-Emitting Diodes: Principles, Characteristics and Processes (Optical Engineering).* New York: Marcel Dekker.

Karim, M.A. (1992). *Electro-optical displays*. New York: Marcel Dekker.

Kelker, H. (1988). Survey of the early history of liquid crystals. *Molecular Crystals and Liquid Crystals,* 165(1), 1-43.

Keller, P. (1991). *The Cathode Ray Tube Technology: History and Applications*. New York: Palisades Press.

Khormaei, R., (1995). A 1280x1024 active matrix EL display. *Society for Information Display Digest*, 891-895.

Khormaei, R. (1994). High resolution active matrix electroluminescent display. *Society for Information Display Digest*, 137-140.

Kim, J.Y., Jeon, D.Y., Yu, I., and Hong-Gun, Y. (2000). A study on correlation of low voltage cathodoluminescent properties with electrical conductivity of In2O3-coated ZnGa2O4:Mn phosphors. *Journal of the Electrochemical Society*, 147(9), 3559-3563.

Klymenko, V., Harding, T.H., Martin, J.S., Beasley, H.H., and Rash, C.E. (1997). Image quality figures of merit for contrast in CRT and flat panel displays. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 97-17.

Klymenko, V., Harding, T.H., Martin, J.S., Beasley, H.H, Rash, C.E., and Rabin, J.C. (1997). Image quality figures of merit for contrast in CRT and flat panel displays. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 97-17.

Krasnov, A.N. (2003). Electroluminescent displays: History and lessons learned. *Displays*, 24, 73-79.

Lattimore, M.R. (1990). Military aviation: A contact lens review. *Aviation, Space, and Environmental Medicine*, 59(2), 125-128.

Lattimore, M.R., and Cornum, R.L. (1992). The use of extended wear contact lenses in the aviation environment. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 92-35.

Lehmann, O. (1889). *Über fliessende Krystalle*, *Zeitschrift für Physikalische Chemie*, 4, 462-72.

Levinsohn, R., and Mason, R. (1997). Advances in helmet mounted display technology. *Head-Mounted Displays II, Proceedings of SPIE*, 3058, 219-230.

Lippert, T.M. (1986). Color-diffe rence predictions of legibility performance for CRT raster imagery. *Society of Information Display Digest*, 86-89.

Markey, W.P. (1988). Preliminary findings - Night vision goggle distortion measurements and specifications. Fort Belvior, VA: Center for Night Vision and Electro-Optics. Memorandum for record.

Martin, W.L., Lee, R.D., Ginsburg, A.P, Genco, L.V., and Task, H.L. (1983). Optical and Human Performance Evaluation of HUD (Head-Up Display) Systems Design. Air Force Aerospace Medical Research Laboratory, Wright-Patterson AFB, OH.

McLean, W.E., and Smith, S. (1987). Developing a wide field of view HMD for simulators. *Display System Optics, Proceedings of SPIE*, 778, 79-82.

Melzer, J.E. (1998). Overcoming the field of view: resolution invariant in Head Mounted Displays. *Head and Helmet Mounted Displays III, Proceedings of SPIE*, 3362, 284-293.

Melzer, J.E., and Moffitt, K. (1997). *Head mounted displays: Designing for the user*. New York: McGraw-Hill.

Meyer, R. (1977). Ferroelectric liquid crystals: A review. *Molecular Crystals and Liquid Crystals*, 40(1), 33-48.

Mikoshiba, S. (2000). Visual artifacts generated in frame-sequential display devices: An overview. *Society of Information Display Digest*, 384-387.

Mosley, A. (1994). Ferroelectric LCD's: The way to the marketplace. *Information Display*, 2, 7-11.

Nelson, S.A. (1994). CVC HMD – next generation high-resolution head-mounted display. *Helmet- and Head-Mounted Displays and Symbology Design Requirements, Proceedings of SPIE*, 2218, 7-16.

Planar Systems. (1998). Electroluminescent Displays, statement by Christopher N. King Executive Vice President / Chief Technical Officer

Post, D.L. (1983). Color contrast metrics for complex images. Doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Powell, K. (2001). Multi-beam bi-directional raster scanning in retinal scanning display. *Proceedings of SPIE*, 4361, 77-88.

Powell, K., Urey, H., and Bayer, M. (2001). Multibeam bidirectional raster scanning in Retinal Scanning Displays. *Helmet and Head-Mounted Display VI. Proceedings of SPIE,* 4361, 77-88.

Pratt, A.B. (1916). Weapons Delivery System for a Marksman. US Patent #1,183,492 issued on May 16, 1916.

Preston, A.R. (2002). Holgraphic display with switchable aspect ratio. US Patent #6377369 issued on April 23, 2002.

Rabin, J. (1995). Dynamic visual performance: Comparison between helmet-mounted CRTs and LCDs. *Journal of Society of Information* Display, 3(3), 97-100.

Rash, C.E. (Editor). (2001) *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft*. Bellingham, WA: SPIE Press.

Rash, C.E., and Becher, J. (1983). Analysis of image smear in CRT displays due to scan rate and phosphor persistence. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 83-5.

Rash, C.E., and Harding, T.H. (2002). A proposed laser safety evaluation for the Microvision, Inc., scanning laser helmet-mounted display system. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2003-01.

Rash, C.E., Harris, E.S, and McGilberry, W.H. (2005). Flexible display technologies…Do they have a role in the cockpit? Fort Rucker, AL: US Army Aeromedical Research Laboratory. USAARL Report No. 2005-07.

Rash, C.E., Kalich, M.E., van de Pol, C., and Reynolds, B.S. (2002). The issue of visual correction compatibility with helmet-mounted displays. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2003-04.

Rash, C.E., McLean, W.E., Mozo, B.T., Licina, J.R., and McEntire, B.J. (1999). Human factors and performance concerns for the design of helmet-mounted displays. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 99-08.

Rash, C.E., Monroe, D.R., and Verona, R.W. (1981). Computer model for the evaluation of symbology contrast in the Integrated Helmet and Display Sighting System. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 81-6.

Rash, C.E., and Verona, R.W. (1987). Temporal aspects of electro-optical imaging systems. *Imaging Sensors and Displays II, Proceedings of SPIE*, 765, 22-25.

Reinitzer, F., (1888). Beiträge zur Kenntniss des Cholesterins, *Monatshefte für Chemie (Wien*). Vol. 9, 421-441.

Rotier, D.J. (1989). Optical approaches to the helmet mounted displays. *Head-Mounted Displays, Proceedings of SPIE*, 1116, 14-18.

Sauerborn, J. (1995). Helmet-mounted tubes for displays. *Information Display*, 6, 10-13.

Schott, J. (1997). *Remote Sensing*. New York: Oxford University Press.

Shenker, M. (1987). Optical design criteria for binocular helmet-mounted displays. *Display System Optics, Proceedings of SPIE*, 778, 70-78.

Sherman, C. (1995). Field-sequential color takes another step, *Information Display*, 3, 12-15.

Sherr, S. (1993). *Electronic Displays*. New York: John Wiley and Sons.

Slikkerver, P. (2003). Bending the Rules, *Information Display*, 3, 20-24.

Snyder, H.L. (1985). Image quality: Measures and visual performance. *Flat-Panel displays and CRTs*. New York: Van Nostrand Reinhold.

Sluckin, T.J., Dunmur, D.A., and Stegemeyer, H. (2004). Crystals that flow. Classic papers from the history of liquid crystals. *Liquid Crystals Book Series*, In: Gray, G.W., Goodby, J.W. and Fukuda, A. (Eds.). London: Taylor and Francis.

Spindt, C.A., Brodie, I., Humphrey, L., and Westerberg, E.R. (1976). Physical properties of thin-film field emission cathodes with molybdenum cones, *Journal of Applied Physics*, 47(12), 5248-5263.

Surguy, P. (1998). How ferroelectric liquid crystal work." *Information Display*, 2, 24-27.

Tannas, L.T. (Editor). (1985). *Flat panels and CRTs*. New York: Van Nostrand Reinhold.

Task, H.T. (1979). An evaluation and comparison of several measures of image quality for television displays. Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory. AMRL-TR-79-7.

Task, H.T., Hartman, R.T., and Zobel, A.R. (1993). New methods for night vision goggle test and evaluation. Wright-Patterson AFB, OH: Armstrong Laboratory. AL/CF-TR-1993-0177.

Task, H.L., Kocian, D.F., and Brindle, J.H. (1980). Helmet mounted displays: Design considerations, *Advancement on Visualization Techniques*, AGARD, No. 255.

Texas Instruments. (2007). How DLP technology works. Retrieved July 11, 2007 from http://www.dlp.com/tech/

Texas Instruments, (1997). Why is the Texas Instruments Digital Micromirror Device™ (DMD™) so reliable? Retrieved May 11, 2008 from www.dlp.com/downloads/default.aspx?&ref=/downloads/white_papers/135_Myth.pdf

Tsou, B.H. (1993). System design considerations for a visually coupled system. *Emerging Systems and Technologies, The Infrared and Electro-Optics Systems Handbook*, 8, 515-540.

Universal Display Corporation. (2007). Retrieved July 3, 2007 from http://www.universaldisplay.com

Urey, H. (1999). Scanner design and resolution trade-offs for miniature scanning displays, *Proceedings SPIE,* 3636, 60-68.

Urey, H. (2000a). Optical advantages of retinal scanning displays, *Proceedings SPIE*, 4021, 20-26.

Urey, H. (2000b). Scanned Display with Pinch Timing and Distortion Correction. US Patent # 6,140,979, issued on October 31, 2000.

US Display Consortium (USDC) Flexible Display Conference. (2003). February 3-4, Phoenix, AZ.

van de Pol, C. Greig, J.L., Estrada, A., Bissette, G.M., and Bower, K.S. (2007). Visual and flight performance recovery after PRK or LASIK in helicopter pilots. *Aviation, space, and Environmental Medicine*. 78(6), 547-553.

Velger, M. (1998). *Helmet-Mounted Displays and Sights*. Boston: Artech House.

Wang, Q.H., Setlur, A.A., Lauerhaas, J.M., Dai, J.Y., Seelig, E.W., and Chang, R.P.H. (1998). A nanotube-based field-emission flat panel display, *Applied Physics Letters*, 72(22), 2912-2913.

Wells, M.H., and Haas, M. (1992). The human factors of helmet-mounted displays and sights. In: Karim, M. (Ed.), *Electro-optical displays*, New York: Marcel Dekker.

Werner, K. (1993). Transferred-silicon AMLCDs, *Information Display,* 10, 21-23.

Widdel, H., and Post, D.L. (1992). *Color in electronic displays*. New York: Plenum Press.

Wikipedia (2007). Electron gun photo. Retrieved July 23, 2007 from http://en.wikipedia.org/wiki/Image:Egun.jpg

Wood, R.B. (1992). Holographic head-up displays. In: Karim, M. (Ed.), *Electro-Optical Displays*, New York: Dekker.

Wu, S. (2003). Liquid crystal materials and devices, School of Optics/CREOL, University of Central Florida, Retrieved from http://lcd.creol.ucf.edu/

Wu, S. and Yang, D. (2001). *Reflective Liquid Crystal Displays*, New York: John Wiley and Sons.

Wu, S. and Yang, D. (2006). *Fundamentals of Liquid Crystal Displays*. New York: John Wiley and Sons.

Yamada, F. (2000). Color sequential LCD based on OCB with an LED backlight, *Society for Information Display Digest*, 1180-1183.

Yeh, P., and Gu, C. (1999). *Optics of Liquid Crystal Displays*. New York: John Wiley and Sons.

Zuckerman, J. (1954). *Perimetry*. Philadelphia: Loppincott.

5 AUDIO HELMET-MOUNTED DISPLAYS

Phuong K. Tran
Bruce E. Amrein
Tomasz R. Letowski

A helmet is a device covering the head and intended to protect the user from hazards to the head. Helmet-mounted display (HMD) systems are generally described as display devices worn on the head as a part of the helmet assembly to provide video information directly in front of the eyes. Sometimes these devices are referred to as head-mounted display systems or head/helmet-mounted display systems (Rash, 2006) since they can be worn on the head both with and without a helmet. In addition, the above display systems are not limited to visual projections and may also include projections of auditory and, potentially, tactile (haptic) signals. Various display systems may be used together providing multimodal information enhancing the user's situation awareness. For example, it has been extensively reported that combined audio/video displays provide a significant increase in visual target detection performance and greatly reduce visual search time (Bergault, Wenzel and Lathrop, 1997; Bolia, D'Angelo and McKinley, 1999; Nelson et al, 1998; Pinedo, Yound and Esken, 2006).

Signals are variable quantities by which information is transmitted from a source to a receiver (Isaacs, 1996). Auditory signals are the acoustic and vibratory signals that create an auditory image of the external world; whereas tactile signals are mechanical pressure signals that are perceived as pressure on the skin (see Chapter 18, *Exploring the Tactile Modality HMDs*).

Auditory signals can arrive to a listener from natural sound sources surrounding the listener or from electroacoustic transducers converting recorded or synthesized electric signals into acoustic waves. The electric signals that are being converted to acoustic signals are called audio signals. The audio signals can be generally defined as audible acoustic signals recorded or generated in an electric form and emitted to the environment as acoustic signals by electroacoustic transducers (audio sources). The audio signals can be as simple as a beep or a spoken word or they can create complex immersive environments such as auditory virtual reality. The system of audio sources projecting auditory signals is called audio display system or, in short, audio display. In other words, an audio display system is a system converting audio signals into acoustic or mechanical (vibration) signals that elicit auditory sensations. Complex auditory sensations result in a perceptual representation of acoustic environment acting on the auditory system of the listener. This perceptual representation is called the auditory image of the environment. Acoustic or vibratory signals radiated by an audio display system can create their own acoustic environment or they can augment an existing auditory image of the natural environment. Numerous applications of audio display technology extend from radio communication, auditory navigation, hearing enhancement (e.g., hearing aids), and music, to fully immersive auditory virtual reality used for training and entertainment purposes. There are also some medical devices (e.g., cochlear implants) that can directly stimulate the auditory nerve and create auditory images but they are not considered in this book.

Audio display systems need to be differentiated from its product, auditory display, in the same way as video display systems need to be differentiated from its product, visual display. Visual displays can be produced by video display systems or can be a result of a specific visible behavior or arrangement of visible objects in the natural environment. In the same way an auditory display can be produced by an audio display system or it can be an arrangement of natural sounds entering the ear (Letowski et al., 2001). In other words, an auditory display is the sum of acoustic signals generating a perception of a particular acoustic environment. The auditory display may consist of a variety of intentional or unintentional sounds, such as speech communications; natural and synthetic sound effects; music, combat-related sounds and urban and vehicle sounds, as well as ambient noise.

The ambient noise is a sum of all unwanted continuous and repetitive sounds that blend together and create an all-encompassing acoustic background at a given location (ANSI, 1994).

The audio HMD system is an audio interface worn as a part of the helmet assembly and providing auditory stimulation. Although the name implicates that the system is helmet-mounted, in reality it is more often a head-mounted system (also abbreviated HMD system), which typically rests on the head of the user even if the system is fully integrated with the helmet. In addition, many military and civilian operations require uninterrupted access to a head-worn audio HMD system even when no helmet is worn. Therefore, in a large number of cases the audio HMD system is not integrated with the helmet and can be worn with or without the helmet as a part of the modular headgear.

A fully-featured audio HMD system must fulfill three major functions providing (1) audio display for radio communication and for other audio-supported functions, (2) hearing protection against harmful high intensity sounds, and (3) means for preserving an effective auditory awareness of the environment. In addition, an audio HMD system is usually equipped with a head-worn or boom microphone as well as a head tracker (Rash, 2006), which need to be incorporated in the design. Thus, the design of an audio HMD system needs to consider appropriate input and output transducers, wiring, connectors, switching systems, signal processing devices, electric impedance matching, padding, isolation issues, and low-power interfaces to the equipment processing the stimuli. Note: An audio head-mounted display (HMD) system is an audio communication system worn on the head or mounted in the helmet of the user. The system may or may not be equipped with a speech communication microphone.

The above requirements for audio HMD systems primarily address the needs of the dismounted Warfighter operating in constantly changing conditions. Mounted Warfighters and aviators may not need as much auditory situation awareness as dismounted Warfighters and their operational (encapsulating) helmets provide some degree of hearing protection. Therefore, the specific focus on individual features of audio HMD systems should depend on the military platform (e.g., dismounted operations, tank, or helicopter) and operational environment in which the system is intended to be used. It is however, important to stress that this system always needs to provide adequate hearing protection. Note that hearing loss is the most common disability of military personnel and the number of hearing loss cases rises rapidly during military conflicts. In addition, under emergency conditions auditory situation awareness may become equally important to all users regardless of the platform and always needs to be taken into account in audio HMD system design. It has to be recognized that all audio HMD systems that protect (cover) the ears introduce some degree of uncertainty in localizing outside sound sources and are detrimental to natural speech communication. Thus, determining the optimal balance between the three main functionalities of the HMD systems discussed above is the most challenging task facing design and selection of HMD systems for specific operations.

This chapter defines the acoustic environment and the concept of an auditory signal together with descriptions of various modes and techniques of audio signal delivery. The background information is followed by a discussion of technical and operational factors affecting design and selection of audio HMD systems including hearing protection and auditory awareness issues. This discussion is combined with an analysis of system requirements for military applications. Advantages, disadvantages, and salient characteristics of each of the audio display design options are discussed to help the reader understand the trade-offs involved in creating or selecting a functional, effective, and reliable audio HMD system that serves its intended purpose and works in concert with the visual HMD and other headgear. The anatomy and physiology of the hearing organ and psychoacoustic of sound perception are not discussed in this chapter since they are addressed in-depth in Chapters 8 (*Basic Anatomy and Structures of the Human Ear*), 9 (*Auditory Function*) and 11 (*Auditory Perception and Cognitive Performance*).

Acoustic Environment

The auditory system is the sensory system responding to a mechanical disturbance of the elastic medium that propagates though the medium as a longitudinal wave. This wave is called an acoustic wave and is perceived as a sound. The term sound is also used in the literature to describe an auditory sensation created by an acoustic wave or mechanical vibration. Therefore, the term sound has dual formal definitions and refers to both the acoustic wave and the auditory sensation (ANSI, 1994).

Opposition of the medium to wave (sound) propagation is called acoustic impedance. Acoustic impedance relates two most fundamental properties of the acoustic wave: acoustic pressure and particle velocity. This relation can be written as:

$$p = Z \times v \,, \qquad \text{Equation 5-1}$$

where p, Z, and v indicate acoustic pressure, acoustic impedance, and particle velocity, respectively. Acoustic pressure is a change in the atmospheric pressure due to a mechanical disturbance of the medium. Particle velocity is the velocity of the oscillatory movement of a particle caused by wave propagation. The product of acoustic (sound) pressure and particle velocity is called sound intensity (I) and it defines the acoustic power of a vibrating particle. Since acoustic pressure and particle velocity are related according to Equation 5-1, sound intensity is proportional to the square of acoustic pressure.

Sound intensity of everyday sounds varies over several magnitudes and therefore it is customary to express its values on the logarithmic scale called sound intensity level. Sound intensity level is defined as:

$$i = 10 \log_{10}\left(\frac{I}{I_0}\right), \qquad \text{Equation 5-2}$$

where i and I mean sound intensity level and sound intensity respectively. A unit of a sound intensity level is the decibel (dB) (ANSI, 1995b). A 10-times change in sound intensity results in an increase of sound intensity level by 10 dB. The value I_0 is the reference point in relation to which sound intensity level is calculated. In many applications of acoustics and audio the value of I_0 is standardized and equal 10^{-16} Watts/centimeter^2 or 10^{-12} Watts/meter^2. This value is also called the zero level and corresponds roughly to the threshold of human hearing at 1000 Hz. Since sound intensity is proportional to the square of acoustic pressure, the Equation 5-2 can be also written as

$$i = 20 \log_{10}\left(\frac{p}{p_0}\right), \qquad \text{Equation 5-3}$$

where p is actual acoustic pressure and p_0 is the reference acoustic pressure. The standardized value of the reference acoustic pressure p_0 corresponding to $I_0 = 10^{-12}$ W/m^2 is equal to 2×10^{-5} Pa. When the sound level is calculated in reference to $p_0 = 2 \times 10^{-5}$ Pa, this level is called the sound pressure level (SPL) and is written as dB SPL (ANSI, 1995b).

Sounds can physically differ in a number of parameters including sound intensity, spectrum, and sound duration. One common classification of sounds based on their duration divides them into continuous (steady-state) sounds and impulse sounds. Continuous sounds are stationary or slightly varying sounds that are longer than a period of observation. Examples of such sounds are sounds of power generators, moving vehicles, and waterfalls. Relatively uniform traffic noise and cafeteria noise can be also considered continuous sounds. Impulse sounds are short sounds that have rapid onset and decay. Such sounds include explosions, weapon fire, and door slams.

Obviously, these two classes of sounds are just the extreme points of a physical continuum encompassing all the sounds, which normally include both stationary and impulse components.

Many sound sources, especially low-frequency sound sources, radiate sound in all directions. Such sources are called omnidirectional sound sources. The sources that radiate most of their energy in one or few distinct directions are called directional sound sources. Examples of such sources are unidirectional and dipole sound sources having beam-like and figure-of-eight radiation patterns, respectively.

Acoustic waves propagating though a medium are absorbed, reflected, dispersed (diffused), and diffracted by space boundaries and various objects located within the medium (space). The distribution of sound energy emitted by sound sources located in the space and modified by boundary effects of the space is called the sound field. A sound field observed at a specific point in the space is called an acoustic image of the field. The acoustic image acting on the listener's ears is the auditory display. The properties of the sound field greatly depend on the amount and distribution of reflected energy and its rate of decay after termination of sound source activity. This rate is called reverberation and is usually expressed as reverberation time (RT) defined as the time needed for the sound energy to decrease by 60 dB.

Sound pressure measurements are usually reported as the sound pressure levels in dB SPL. However, a specific sound pressure level does not necessarily mean that the sound is loud or even perceived by human hearing. Acoustic waves of very low and very high frequencies that fall outside of the range of the human hearing do not contribute to the loudness of the sound. Therefore, if someone wants to assess perceptual effects of sound, the measurement needs to take into account the properties of the human ear (see Chapter 8, *Basic Anatomy and Structure of the Human Ear*). There are several weighting curves that when applied to dB SPL data provide information about potential auditory effects of the specific sound. Most commonly used weighting curves are A-, B-, and C-weighting. These curves are mirror images of average frequency-dependent equal-loudness curves of human hearing in the 0-40, 40-70, and 70-120 phon range, respectively. They mainly represent the way in which the frequencies below 1000 Hz are filtered by the ear at different SPLs. The SPL data processed with these weightings are written as dB (A), dB (B), and dB (C) (ANSI, 1995b).

Auditory Signals and Display Formats

The process of perceiving sound is called hearing or audition (see Chapter 9, *Auditory Function*). The sensation of sound can be created by acoustic waves arriving at the ears of the listener (air conduction) or by direct vibration applied to the head (bone conduction). The auditory system acquires, interprets, selects, and organizes simple and complex auditory stimuli and creates an auditory image of the physical environment surrounding the listener. The field of science devoted to the human perception of sound is called psychoacoustics. In order to understand how humans perceive sounds one must know what the human hears, and which portions of the perceived sounds are considered to be useful information (signal) and which portions are considered to be distracting background (noise). While the sound pressure levels presented to the human ear may be precisely measured, it is difficult to determine exactly an auditory effect of the stimulation. These auditory effects may depend on a person's expectations, attention, health, and multi-faceted environmental conditions. They also depend on the relative importance assigned to the specific sounds by the listener. For example, Warfighters rely heavily on auditory information carried by environmental sounds when they are on patrol or on search missions and on sound signatures of weapons, helicopters, and vehicles when they are in a combat situation. The importance of auditory information increases many-fold when visual information is obscured by smoke, fog, or darkness.

Auditory signals can be generally defined as an acoustic or vibratory stimulus received by the hearing system and converted into auditory information. Both intentional sound messages and unintentional sounds can be signals. If specific auditory information is not considered useful and degrades perception of auditory signals it becomes an interfering noise. Auditory noise may have internal (physiological) and external (acoustic) origins. The effect of noise on the perceived signal is usually quantified as a signal-to-noise ratio (SNR). The SNR is the ratio of some measured aspect of a signal to a similar measure of a concurrent noise expressed in a logarithmic form (Letowski et al., 2001).

Auditory signals can be projected by distal and proximal display systems. The distal auditory display systems are those where the actual sound sources are located away from the listener's ears. Examples of distal audio display systems are all real-world environments and various loudspeaker-based sound projection systems. Proximal audio display systems are display systems located close to the listener's ears. All audio HMD systems are proximal audio display systems since they are mounted to the listener's head at or close to the listener's ears. A small loudspeaker mounted on the shoulder strap is usually sufficiently far away from the listener's ears to consider it a distal display.

When a listener is placed in an acoustic environment that contains several sound sources surrounding the listener, all sounds arrive at the both ears of the listener regardless of the location of the sources. This situation is shown in Figure 5-1. The sounds may arrive at different times and with different intensities and they may arrive directly from the sound sources (Figure 5-1) or after being reflected from surfaces in the surrounding environment. Regardless of the specific pathways, the sounds from each sound source will arrive at both ears of the listener.

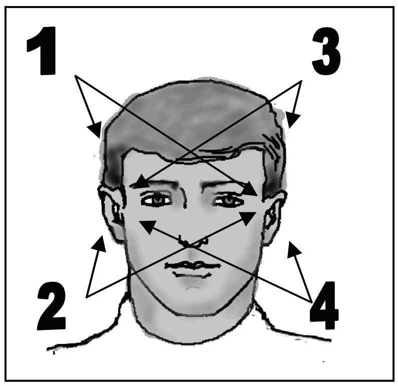


Figure 5-1. Auditory display created by distal sound sources (1 through 4).

Auditory signals may be presented to the human in various forms and by various techniques. The main classification of the auditory displays (Letowski et al. 2001) is shown in Table 5-1.

Table 5-1.
Classification of auditory displays created by various types of input signals and signal projection systems (natural sound sources and/or audio display systems).

Input Signal	**Signal Projection**	
	One Ear (Monaural Listening)	**Two Ears (Binaural Listening)**
One Channel (Monophonic Signal)	Monotic	Diotic (Biaural)
Many Channels (Multi-channel Signal) (Stereophonic Signal)	Monotic (Monomic)	Dichotic (Spatial)

A monophonic signal is a single channel signal delivered to one or many transducers of the audio display system. Multi-channel (or stereophonic) signals are a group of uncorrelated (or correlated) signals delivered to individual transducers of the audio display system. Regardless of the type of the signal, the audio system can create an audio display that can be projected to one (monaural listening) or both (binaural listening) ears of the listener.

A monotic auditory display is created by auditory signals delivered only to one of the listener's ears. This type of display is also frequently described as a monaural display. In the case when we want to stress that several signals are combined together and delivered to a single ear, the monotic display can be called a monomic display. Figure 5-2 shows monotic or monaural sound presentation to the listener's left ear.

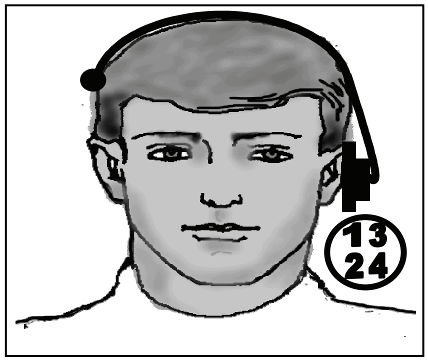


Figure 5-2. Monotic or monaural auditory display. All (1, 2, 3, and 4) sound sources are presented to a single ear of the listener.

When the same signal is presented to both ears of the listener the auditory display is called diotic or biaural. The biaural (diotic) display causes the image of the sound sources to appear within the listener's head instead of being located to the side of the head as in the case of monotic displays. Biaural listening improves speech intelligibility, especially in a noisy environment due to the increased perceived loudness provided by the biaural presentation and better spatial separation of the phantom signal source from extraneous noise sources. Figure 5-3 shows the concept of the biaural presentation method.

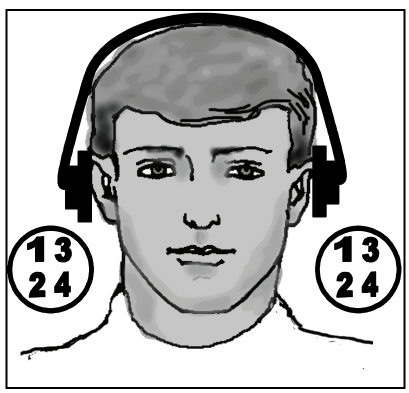


Figure 5-3. Diotic (biaural) auditory display. The same four sound sources are presented to each ear of the listener.

When different signals are delivered to each ear of the listener such an auditory display is called a dichotic or spatial display. Biaural (diotic) and dichotic (spatial) presentations are two forms of two-ear listening called binaural listening. The dichotic display is produced when independent (uncorrelated) auditory stimuli are delivered to the listener's right and left ear. This type of display takes place when the listener monitors two or more radio networks at the same time. Figure 5-4 shows the situation when uncorrelated stimuli are presented to each ear. Stimuli 1 to 4 are heard in one ear and stimuli a to d are heard in the other ear of the listener. When the signals delivered to the left and right ear of the listener are different but correlated they create a spatial auditory display in which phantom sound sources are distributed in space.

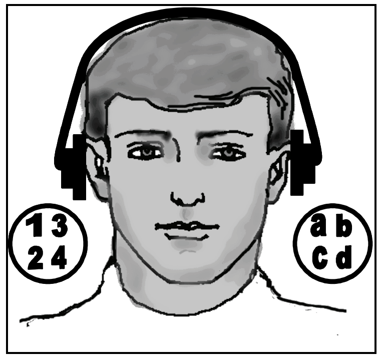


Figure 5-4. Dichotic auditory display. Different sets of sound sources are presented to each ear of the listener.

The display formats listed in Table 5-1 are the basic formats of audio displays. However, various combinations of the basic formats are possible through signal processing and audio switching techniques. For instance, two or more channels may be presented to one or both ears, and when a high priority signal requiring immediate attention arrives, it may be directed to one ear at a higher intensity level than the less important signals. With the proper control of the incoming signals the listener may cause the sound intensity of the selected channel (right or left) to be presented at a higher level than the other channel.

Auditory signals delivered though the audio HMD systems may provide different degrees of spatial information (i.e.; information to permit the listener to determine where the sound source is located in space) to the listener. Monophonic signals and uncorrelated multi-channel signals are localized as originating either in the ear or from within the listener's head. Such phantom location of the sound source is called internalization. If the monophonic signal delivered to the left and right ear has the same intensity the phantom sound source is located in the center of the listener's head. By changing the intensity ratio between the signals arriving at the left and right ears the phantom location of the sound source can be moved anywhere on the arc connecting left and right ears. Such displacement of the phantom sound source from its central position in the head is called lateralization. Figure 5-5 shows two examples of biaural auditory displays with the phantom sound source lateralized to one ear (Figure 5-5a) and internalized in the center of the head (Figure 5-5b).

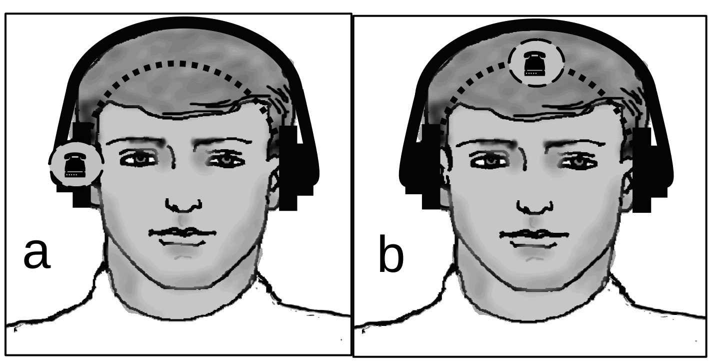


Figure 5-5. Biaural display with the phantom sound source completely lateralized to the right (a) and located in the center of the head (b). The dashed line shows an arc along which the phantom image lateralizes.

When the auditory signals that are delivered to both ears of the listener are stereophonic (correlated) signals such an audio display is frequently called a spatial, binaural, or stereophonic display. The phantom sound sources created by the stereophonic signals are internalized within the listener's head similarly to the diotic signals. However, specific location of the phantom sound source within the head is not only determined by the difference in the intensity of sounds delivered to the left and right ear but also by interaural cross correlation (ICC) associated with signals received by the two ears (White, 1987). The values of ICC can vary from -1 to +1, where -1 indicates identical signals are presented to both ears, but they are 180° out of phase; 0 indicates the two signals are unrelated to each other (dichotic) ; and +1 indicates the two signals are equal and in phase (diotic). Examples of phantom sound source locations dependent on the ICC values are shown in Figure 5-6. Note that when left and right signals are 180° out of phase the location of the sound source in the head is smeared and less defined (Figure 5-6a).

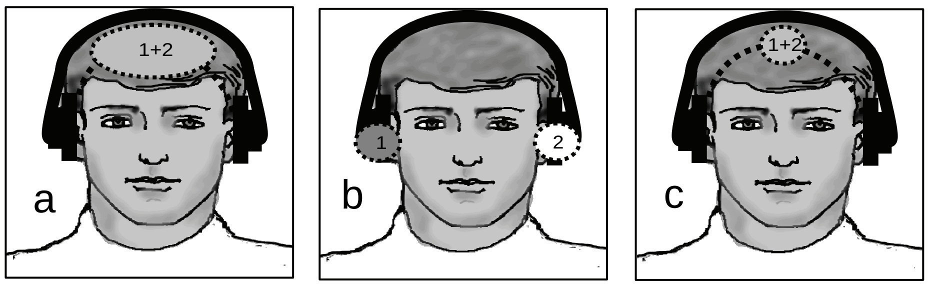


Figure 5-6. Phantom sound source locations created by different amounts of interaural cross correlation (ICC) between signals 1 and 2 delivered to the right and left ear: (a) ICC= -1, (b) ICC= 0, and (c) ICC= 1.

Head-Related Transfer Function (HRTF)

The human auditory system determines the location of the distal sound source in space by making use of binaural and monaural localization cues. In addition, familiarization with sound sources and head movements can enhance a person's ability to localize the direction of incoming sound.

Binaural cues are the differences in the intensity and the time of arrival of the sound wave from a particular sound source arriving at the left and right ear of the listener. These cues are called the interaural intensity difference (IID) and the interaural time difference (ITD). Binaural cues operate in the horizontal plane and allow the listener to determine the azimuth of the incoming acoustic signal. The amounts of IID and ITD are dependent on the size of the listener's head and for the signal source located on the side of the listener's head they can be as large as 15-20 dB and 0.6-0.8 ms, respectively. The amount of IID is frequency dependent whereas the amount of ITD is not. However, the ITD can be converted into an equivalent interaural phase difference (IPD), which is frequency dependent. Humans appear to use the IPD cues to localize sources of low frequency sounds and the IID cues to localize sources of high frequency sounds.

The monaural cues are related to the geometry of the listener's head and the shape and location of the pinna. These cues are used both to determine the azimuth and elevation of the sound source and to differentiate between the front and back directions of the incoming signals. Different monaural cues operate in different frequency ranges. For example, due to the small dimensions of pinnae as compared to the wavelengths of acoustic energy heard by humans, the pinna cues are most effective for localization of high frequency sound sources at and above 3000 Hz (Musicant and Butler, 1984).

The shape and intricacies of the human head and the upper torso generating monaural localization cues constitute a complex directional acoustic filter (equalizer) that relates the locations of the sound sources in the space to the specific characteristics of the signals arriving at the ears of the listener. This filter is called the head-related transfer function (HRTF). An HRTF is the ratio of the sound pressure at the ear of the listener to the sound pressure that would exist at this point if the listener was not present expressed as a function of frequency. Each location of the sound source in space is coded into a pair (left and right) of HRTFs that allow the listener to identify this location. The shape of the HRTF does not change with the distance from the sound source to the listener except for very short distances of less than 1 meter when the proximity effects and the sound bouncing between the head and the sound source need to be taken into consideration. A small set of HRTFs obtained for a group of listeners is shown in Figure 5-7. Please note that the shape of HRTF differs quite substantially among people and our listening experience is affected by the peaks and valleys of our individual HRTFs.

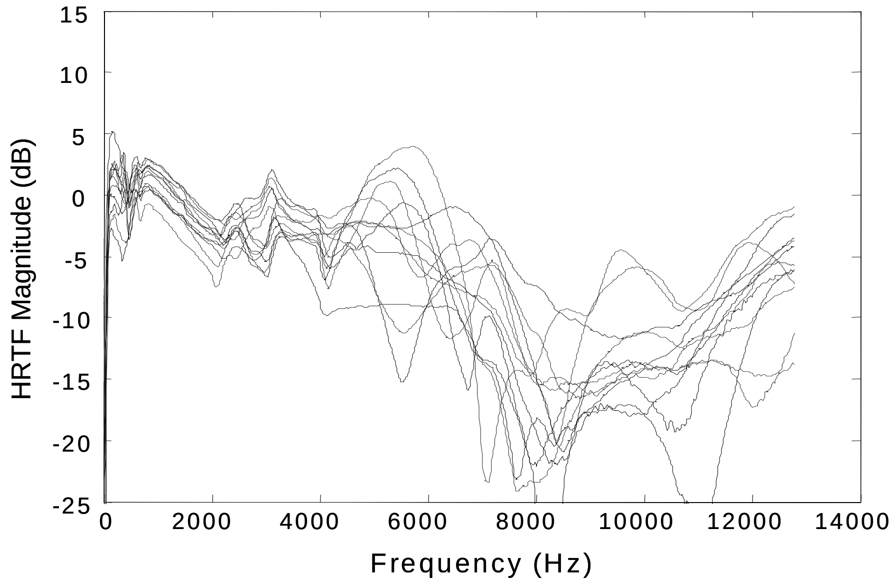


Figure 5-7. A set of HRTFs (magnitude) obtained for a small population of listeners at one specific position of a sound source (Vaudrey and Sachindar, 2003, used with permission).

The natural human ability to localize sound sources in space is lost when the auditory signals are presented directly to the human head by the proximal display systems, such as earphones. In order to restore an impression that the sound sources are located in space outside of the human head (externalized), the specific HRTFs of the listener need to be captured and incorporated into the signal delivered to the display system. A pair of HRTFs for the left and right ear represents the anatomical capabilities of a specific human that are used in identifying the specific direction toward a sound source. A set of HRTFs for various angles of incidence captures all binaural and monaural localization cues characterizing a specific individual and can be used to synthesize spatial perception in a virtual environment created by audio HMD system. Therefore, a monophonic sound recording convolved with a matched pair of HRTFs and played through earphones will result in the impression the sound source is located in space and outside of the head of the listener. In other words, when the recording of the natural distal environment is convoluted with a pair of matched HRTFs and played through earphones (proximal displays), the listener experiences externalized auditory images (Hartman and Wittenberg, 1996). Such audio displays are called 3D-audio displays or spatial audio displays. Figure 5-8 presents the differences between the display of real-world stimuli, the biaural display, and the spatial biaural display created with HRTFs.

In order for the listener to have natural externalization of the sound sources, the auditory signals must be convoluted with the listener's own HRTFs. Such HRTFs are called individualized HRTFs. Attempts to create average (non-individualized) HRTFs that can be effectively used for many listeners were unsuccessful and resulted in poor accuracy of the sound field recreation (Wenzel, Arruda, Kistler and Wightman, 1993).

The HRTFs can be recorded in a number of ways, but the most common method is to place miniature microphones in the openings of a listener's ear canals and to make multiple recordings of a standard test signal presented from various azimuths and elevations around the listener (Wightman and Kistler, 1989ab). The test signal can be a frequency-swept sine wave or a standard impulse signal such as maximum-length sequence (MLS) or a Golay code (Zahorik, 2000). Figure 5-9 shows a typical HRFT measurement system (left) and the microphone placement in the listener's ear (right). In order to determine the HRTFs for specific listeners and directions, the recordings of the standard signal are made in the manner shown in Figure 5-9 and compared to the similar recordings made with a single microphone located at the point corresponding to the center of the listener's head. The differences between these recordings for specific sound source locations constitute the set of HRTFs for a given listener. Such HRTFs can be applied to any acoustic signal through a process called convolution and delivered to the Audio HMD systems to create an impression of auditory stimuli arriving from the surrounding space.

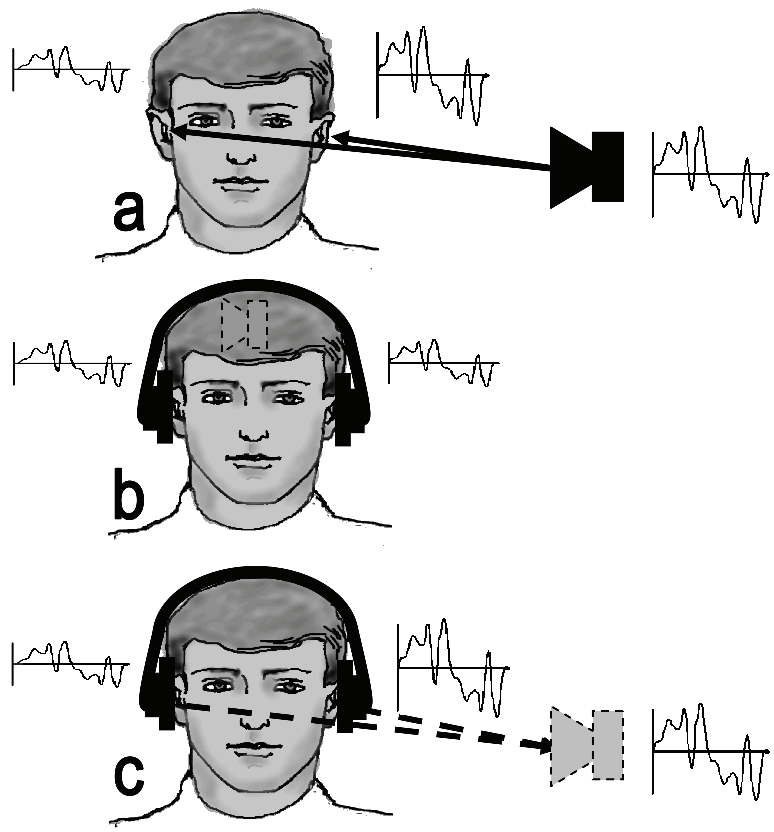


Figure 5-8. Natural hearing and spatial audio display created by using HRTFs. Part a shows spatial position of the distant sound source perceived by natural hearing. Part b shows a biaural display that creates the image of the sound source inside the head. Part c shows spatial position of the sound source faithfully recreated when the earphone-reproduced signals are filtered with an appropriate head related transfer function (HRTF). Signal waveforms shown in individual panels represent the signals emitted by the real loudspeaker (panel a), associated with phantom loudspeaker (panel c), and the signals heard by the listener (symbols next to the head in all three panels).

Figure 5-9. Measurement of HRTFs using the KEMAR manikin and a loudspeaker mounted on a robotic arm (left) and the HRTF microphone placement in the listener's ear (right). (U.S. Army Research Laboratory (left) and Vaudrey and Sachindar, 2003 (right) (used with permission).

An alternate method of recording HRTFs involves placing the sound source in the ear canal and placing an array of microphones around the listener (Zotkin et al., 2006). Such a recording configuration is shown in Figure 5-10. This method is based on the reciprocity principle with microphone and loudspeakers reversing their positions in space in comparison to the previous method. The advantage of the reversed setup is much shorter measurement time. However, such a system requires many identical calibrated microphones and is both expensive and difficult to maintain.

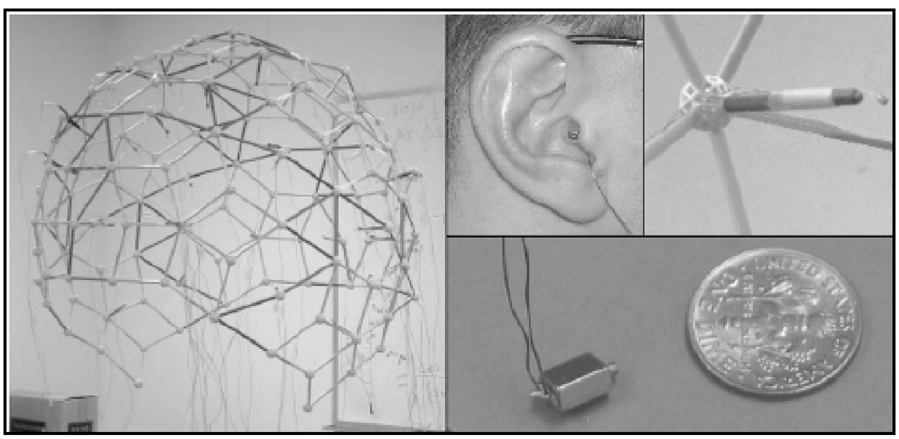


Figure 5-10. Left: The measurement mesh consisting of 32 microphones. Bottom right: The miniature loudspeaker. Top middle: The miniature loudspeaker inserted into the listener's ear. Top right: An enlargement of the one node of the measurement mesh (adapted from Zotkin et al., 2006).

Despite the natural differences in specific techniques used to record HRTFs in various laboratories, it seems that the properly measured HRTF sets obtained in various laboratories on the same people are operationally equivalent. A study conducted at the U.S. Army Research Laboratory at Aberdeen Proving Ground (MacDonald and Tran, 2008) compared the effects of four sets of KEMAR HRTFs on localization accuracy of the listeners in virtual auditory environments. The four sets were: (1) in-house KEMAR HRTFs measured with MLS signal and Tucker Davis Technology hardware (RoboArm), (2) HeadZap HRTFs from AuSim Inc., (3) Massachusetts Institute of technology (MIT) HRTFs from Media Lab at MIT, and (4) CIPIC HRTFs from CIPIC (Center for

Image Processing and Integrated Computing) Interface Lab at the University of California at Davis. These four sets were compared by a group of 16 listeners in a localization task using spatial stimuli presented over headphones. The task was limited to the horizontal plane and eight discrete phantom sound source locations. The mean absolute localization errors observed at each of the eight phantom source locations for the four sets of HRTFs are shown in Figure 5-11.

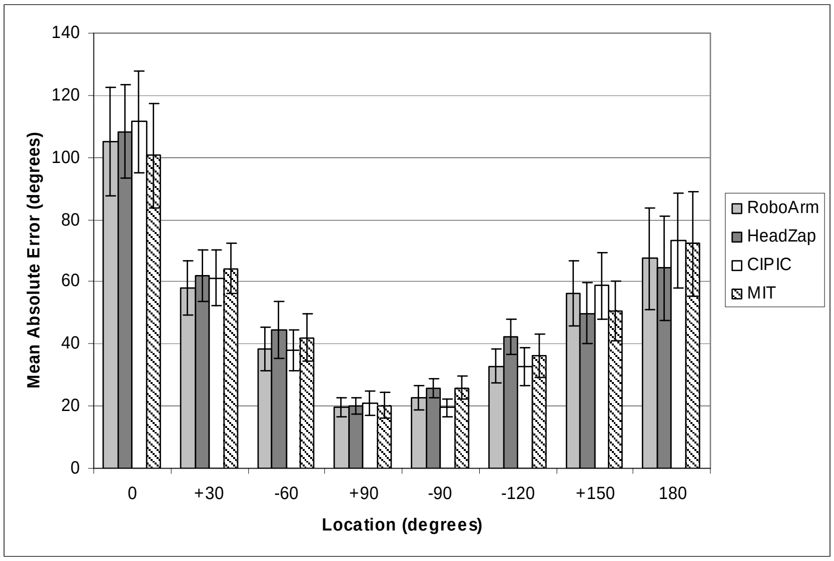


Figure 5-11. Result of the HRTF comparison study. Mean absolute errors for eight phantom sound source locations obtained with four HRTF sets (U.S. Army Research Laboratory).

The data in Figure 5-11 show that all four HRTF datasets resulted in functionally identical spatial audio simulations despite the differences in recording technologies, locations, and recording personnel for each of the datasets. Considered in terms of the localization performance of the listeners using the generalized (non-individualized) HRTFs, the four sets were nearly indistinguishable from one another.

Audio HMD System Specifications

The technical specifications of an audio HMD system depend on the complexity and multi-purpose character of the system and can substantially differ from one system to another. However, the minimum specifications of such a system should include the type of device, available operational modes, audio performance data, hearing protection data, weight, and the type of electric interface to the supporting platform. It is also assumed that each of these devices meets standard military requirements for ruggedness and required range of operational conditions. These requirements are specified in MIL-STD 810F *Department of Defense Test Method Standard for Environmental Engineering Considerations and Laboratory Tests* (DOD, 2001). All these parameters are important for proper operation of the system but the core requirements that characterize the quality of the audio interface and affect its potential range of applications are the audio performance data. These data should represent a number of specific electroacoustic characteristic of the system on the basis of which two or more systems can be compared. The list of basic electroacoustic characteristics of audio HMD systems is provided in Table 5-2. More information about specific transducers and other elements of the audio HMD systems is provided in the later sections of this chapter.

Table 5-2.
Basic operational characteristics of audio HMD systems.

Operational Characteristic	Definition
Sensitivity	System sensitivity is the effectiveness of the system (audio transducer) to convert input signal into the output signal. Three basic audio sensitivities are earphone sensitivity, loudspeaker sensitivity, and microphone sensitivity. Earphone sensitivity – or earphone efficiency - is the sound pressure level in dB SPL produced by an earphone in response to a 1 mW signal in a standardized coupler with a built-in microphone (IEC 60268-7). Unit: dB SPL/mW (dB SPL per milliwatt). Loudspeaker sensitivity – or loudspeaker efficiency - is the sound pressure level in dB SPL produced by a loudspeaker at 1 m distance in response to a 1 W signal (IEC 60268-5). Unit: dB SPL/W (dB SPL per watt). Microphone sensitivity is the voltage output in mV produced by a microphone in response to a 94 dB SPL signal (1 Pa) (IEC 60268-4). Unit: mV/Pa (milivolts per Pascal).
Frequency Bandwidth	Frequency bandwidth is the frequency range within which the system sensitivity does not change by more than a specific number of dB, usually 3 dB, from its nominal sensitivity level. Unit: Hz (Hertz). In digital systems the bandwidth is defined as the data transfer rate and measured in bits per second. Unit: kb / s (kilobits per second).
Nonlinear Distortions	Nonlinear distortions are new frequency components in the output signal that do not exist in the input signal but result from the presence of the input signal. Nonlinear distortions are typically measured as a percent of the total signal. The most common type of nonlinear distortions is harmonic distortions and intermodulation distortions. Unit: % (percent).
Maximum Input Power	Maximum input power (rated power) is the highest continuous power that the device can handle without producing excessive sound distortion or being damaged. Maximum input power is usually defined by a specific percent of nonlinear distortions that the system cannot exceed under normal operational conditions. Unit: mW (milliwatts) or W (watts).

Table 5-2. (Continued)
Basic operational characteristics of audio HMD systems.

Dynamic Range	The ratio of the highest undistorted input signal to the smallest input signal that produces output signal that is discernible from system noise. Unit: dB (decibels).
Headroom	Headroom is the level difference between the typical operating level of the device and the maximum operating level defined by the maximum input power or the onset of signal clipping. Unit: dB (decibels).
Electric Impedance	Electric impedance is the opposition of a system to the flow of the electric current. Unit: Ω (ohms). Electric impedance can be either the input impedance (earphone, loudspeaker) or the output impedance (microphone) of a system. Effective transmission of power requires that input impedance of the next system matches the output impedance of the previous system.

Audio HMD Transmitter Systems

An audio transducer is a device that converts electric energy into acoustic energy or acoustic energy into electric energy. The former type of transducer is called a transmitter and the latter is called a receiver. Common examples of audio transmitters and audio receivers are the loudspeaker and the microphone, respectively. Audio transmitters are the main elements of audio HMD systems converting audio signals into auditory stimuli. In addition, the system may be equipped with an audio receiver (boom microphone) for converting user's speech into audio signals that can be transmitted through radio communication equipment. The topic of audio receivers is briefly addressed in several places in this chapter but is beyond the scope of the chapter.

Common sound delivery methods used in audio HMD systems can be divided into two basic classes based on the sound transmission interface: air conduction transducers (earphones) and bone conduction transducers (bone vibrators). Each of these has their own distinct advantages and limitations and the selection of one or another needs to be dictated by specific operational requirements.

Earphone systems

The terms earphones, headphones and headsets are frequently used interchangeably but to the audio community each of these terms defines a unique system of technologies. The relation between these systems is shown in Figure 5-12.

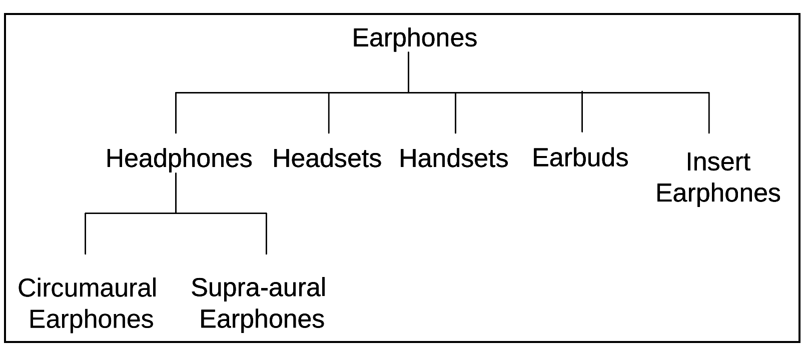


Figure 5-12. Types of earphones.

Earphones are all audio transducers that are directly coupled to the ear of the listener in such a manner that the reproduced sound waves are delivered to the eardrum through the air in the ear canal. In some professional literature (e.g., telephony, audiology) earphones are sometimes referred to as earpieces. Headphones are the earphones applied outside of the ear and supported by a headband placed over the head whereas insert earphones are the earphones inserted directly into the ear canal. The earphones mounted in the concha (earbuds) are not insert earphones and are considered a separate class of the earphones. Headsets are the headphones equipped with an additional microphone for speech communication and handsets are earphone-microphone combinations applied to the head by hand such as the telephone handsets.

The headphones are generally divided into circumaural and supra-aural headphones (details below). They are usually referred to as the circumaural and supra-aural headphones when considered as a system together with the supporting headband. Otherwise, they may be simply referred to as circumaural and supra-aural earphones.

Circumaural earphones are large earphones with earcups (earmuffs) that surround the outer ear resting against the head with little or no contact with the pinna. The audio transducer is loosely coupled to the ear with a relatively large volume of air under the earcup (ANSI, 1995). An important characteristic of circumaural earphones is that their earcup encloses and extends the cavity of the ear canal lowering natural resonances of the canal and adding new resonances of the earcup volume.

The circumaural earphones can be either closed-air (closed- back) or open-air (open-back) earphones. The closed-air earphones have earcups that completely separate the output from both sides of the earphone membrane from the external sound environment. This type of earcup attenuates external noise and minimizes sound leakage from the earphone. The noise attenuation is typically 10 dB to 15 dB at mid-frequencies. The closed back of the earphone results in extended and more resonant (boomy) low-frequency response of the earphones due to more pronounced resonances of the earcup space and earphone enclosures. Open-air earphones have an open grille at the back and in some cases the openings on the side of the ear cup. Modern open-back earphones use a semi-open design in which the open grille at the back of the earcup is replaced by several small, well defined openings that are used to tune the frequency response of the earphone. These earphones are often referred to as the semi-open earphones. An example of such an earphone is the AKG K240 headphones. The open-air and semi-open earphones do not isolate the listener from the external environment as much as the closed-air earphones and they leak the sound to the environment. However, open back design reduces the effects of earcup and earphone resonances making the frequency response smoother and reproduced sounds more natural. Examples of closed-air and open-air circumaural headphones are shown in Figure 5-13.

Figure 5-13. Examples of circumaural earphones: Sennheiser Model HD201 (closed-air [left]) and Sennheiser Model HD 595 (open-air [right]) (Courtesy of Sennheiser USA).

In general, circumaural headphones produce high sound quality and provide some hearing protection against external noise. They are used for studio listening and when some degree of isolation from the environment is required. However, they are bulky and expensive when compared to other styles and may become uncomfortable after prolonged use due to the lack of air circulation under the earmuffs. In military operations the closed-air circumaural earphones are frequently mounted in the aviator or tanker helmets, which need to provide protection against the platform noise and where environmental hearing (auditory awareness of the environment) is not critical. Figure 5-14 shows an example of a circumaural audio HMD system designed for use in the U.S. Army's combat vehicles.

Supra-aural earphones rest on the external ears pressing against the pinnae. An audio transducer is coupled to the ear through a foam (soft) or rubber (hard) cushion. The earphones provide virtually no isolation from external noise and leak some sound to the environment. In some cases the audio transducer is specially located at a distance from the ear to provide a more uniform sound. Examples of such earphones are the Sennheiser HD 414 that is separated from the ear by a thick foam cushion covering the ear and the AKG 1000 that features a pair of loudspeakers radiating into the ears without any physical coupling.

Figure 5-14. Closed-air circumaural earphones mounted in the Combat Vehicle Crewman (CVC) Helmet used in the U.S. Army combat vehicles. The CVC Helmet is shown with protective ballistic shell attached at its top (Courtesy of the Bose Corporation).

Similarly to the circumaural earphones, the supra-aural earphones can be open-air (open-back) or closed-air (closed-back) systems providing various trade-offs of sound quality, comfort, and sound isolation. They are smaller and typically much lighter than circumaural earphones. Traditionally, supra-aural earphones are open-air types; however, here are also some supra-aural earphones that are the closed-air type (e.g., Sony MDR V700). Two examples of supra-aural headphones are shown in Figure 5-15.

Earbuds are small bowl-like earphones with or without small earcups that are placed in the concha or at the entrance of the ear canal without fully covering it. They are usually open-air type. Earbuds can be secured in place with a headband or ear clips that clip on the outer ears. This is the most common type of earphone used with portable devices such as cellular phones, iPods®, and MP3 players. In general, earbuds are inexpensive and lightweight (usually weighting less than 10 grams) but are not always comfortable and in general provide an audio display of lower quality than other types of earphones. However, despite their small size they are designed to produce quite high sound levels comparable to those of high power circumaural earphones in order to compensate for their lack of isolation from external noise. For example, the AKG K12P earbud-type earphones are rated as producing 127 dB SPL/V. Two examples of earbud-type earphones are shown in Figure 5-16.

Figure 5-15. Lightweight Supra-aural Earphones: Sony Model MDR 410LP (left) (Photo courtesy of the Sony Corporation); Beyer Dynamic Model DT 231PRO (right) (Courtesy of Beyer Dynamic Corp).

Figure 5-16. Sony MDR A34L earbud earphones with a headband (left) and Sony MDR EX71SL earbuds (right) (Courtesy of the Sony Corporation).

The last basic type of air conduction audio devices used in audio HMD systems are insert earphones, also known as canal earphones, canalphones, in-the-ear monitors, or in-the-ear earphones. Insert earphones use very small audio transmitters that fit inside the ear canal or are coupled to the ear by acoustic tubing ending with eartips (ear adapters). The eartips may be either disposable foam plugs or differently-sized permanent tips. Some manufacturers also offer custom molded insert earphones. Such earphones offer high noise isolation and have the potential for perfect fit to the shape of the ear canal. However, custom fitted canal earphones are expensive and cumbersome to replace.

The degree of comfort depends greatly on how the earphones fit into the ear canals of the listener. If the earphones fit perfectly, they provide excellent ambient noise attenuation and are comfortable for long term use. Their sound quality is usually very good and comparable to that of supra-aural and circumaural earphones. However, in some cases the users can notice the effect of occlusion due to modified acoustic properties of the ear canal blocked by the earphone. The occlusion effect is an increased audibility of low-frequency bone conducted sounds. This effect results as an additional amplification of the talker’s own voice, which is heard by the talker as louder and stronger in low frequency energy (darker) than normal voice. Some occlusion effect is also present when the ear is covered by closed circumaural earphones. Some people with sensitive skin can also experience skin irritation if the insert earphones are worn for an extended period of time. Another disadvantage of insert earphones is their high maintenance; special care is needed due to progressive accumulation of earwax in the eartips of the earphones. The earwax closes the output port of the earphone and degrades the quality of the sound. To minimize this effect some of the insert earphones, e.g., Etymotic ER 6, use a special replaceable filter at the tip of the earphone. Finally, insert earphones, especially those that are not custom made, have a tendency to move in the ear when the user is running or moving heavily. These movements produce some noise in the ear canal and changes in the quality of sound. An external view of Shure E 4G insert earphone is shown in Figure 5-17.

Figure 5-17. Shure E4 insert earphone (right); shown with a custom molded silicon-gel sleeve (left) (Courtesy of Sensaphonics, Inc.).

A special application of an audio display system incorporating insert earphones is hearing aids that are designed to compensate for various types of hearing impairment. The primary difference between an insert earphone and a hearing aid is electronic circuitry of the hearing aid that includes built-in microphone, amplifier, some signal processing circuitry, and a battery to power the system. Hearing aids are self contained in a very limited space. Depending on the location of the system circuitry, hearing aids are classified as (1) behind-the-ear (BTE), (2) in-the-ear (ITE), (3) in-the-canal (ITC), and (4) completely in-the-canal (CIC) hearing aids. The BTE hearing aid is placed behind the pinna and is coupled to the ear though acoustic tubing terminated with an earplug. Conversely, the CIC hearing aid is deeply inserted in the canal and completely hidden from a casual observer in the shadow of the canal. Figure 5-18 shows typical shapes of various types of hearing aids.

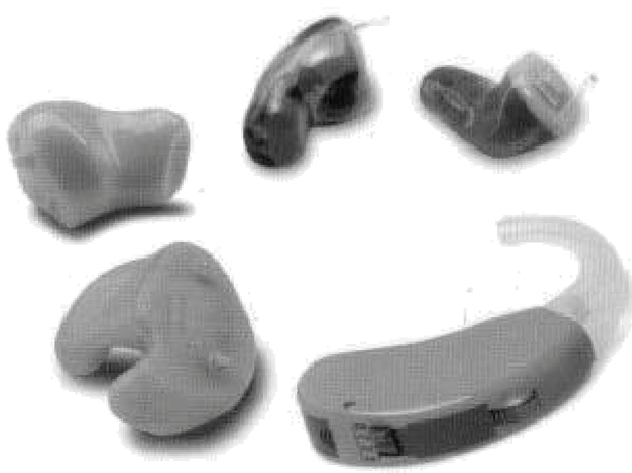


Figure 5-18. Various types of hearing aids (clockwise from the top: left and right completely-in-the-canal (CIC) hearing aids, behind-the-ear (BTE) hearing aid, in-the-ear (ITE) hearing aid, in-the-canal (ITC) hearing aid (Courtesy of Beckner Hearing Aides, Inc.).

An audio HMD system can incorporate any type of the audio transmitters described above. For example, both the closed-air circumaural earphones and the insert earphones have been incorporated in U.S. Army aviator helmets (e.g., HGU-56/P and SPH-4). Some proposed military helmet assemblies included small supra-aural earphones or loudspeakers built into the helmet's ballistic shell (e.g., Land Warrior Integrated Helmet Assembly Subsystem [IHAS]). Similar systems are also offered for the motorcycle helmets. The advantages and disadvantages of various earphones types used for audio displays are summarized in Table 5-3.

Table 5-3.
Advantages and disadvantages of various types of earphones used in audio HMD systems.

Earphone type	Advantages	Disadvantages	Examples
Closed-air circumaural headphones 	-Comfortable -Excellent acoustic isolation	-Bulky -Accumulate heat -Low frequency resonances	-AKG K271 -Beyer DT 250 -Beyer DT 770 -Bose QuietComfort 2 -Sennheiser HD 265/280 Sony MDR 7506
Open-air circumaural headphones 	-Comfortable -Sound natural	-Bulky -Low acoustic isolation	-AKG K240 -AKG K701 -Beyer DT 990 -Sennheiser HD 595/600/650
Closed-air supra-aural headphones 	-Less bulky, lighter weight compared to circumaural -Good acoustic isolation	-Accumulate heat -Resonance	-AKG K171 -Beyer DT 231 -Bose QuietComfort 3 -Sennheiser HD 25 -Sennheiser PX 200 -Sony MDR V700 -Telephonics TDH 39
Open-air supra-aural head-phones 	-Light weight -No acoustic isolation	-No acoustic isolation	-AKG K141 -AKG K1000 -Beyer DTX 30 -Grado RS1 -Grado SR 325 -Sennheiser HD 414/465 -Sony MDR 410

Table 5-3. (Cont)
Advantages and disadvantages of various types of earphones used in audio HMD systems.

Earbuds 	-Low cost -Very light weight -No acoustic isolation	-Low sound quality -No acoustic isolation	-AKG K14P -Altec Lansing AHP 131 -Beyer DTX 20 -Sony MDR A34L -Sony MDR EX81 -Sony MDR E82
Insert earphones 	-Very light weight -Good acoustic isolation	-Need good fit -High maintenance -May cause skin Irritation and occlusion effect -Good earphones are expensive	-Bose TriPort IE -Etymotic Research ER 3A/ER4/ER6 -Shure E3/E4 -Westone UM2

One problem caused by earphone-based audio display systems is that such systems occlude the ear and adversely affect auditory awareness of the environment. In many cases, such as air or mounted operations such awareness is not essential. However, during dismounted operations in quiet or in urban environments such awareness is critical to Warfighter's safety and effectiveness. In these cases earphone-based systems need to be equipped with environmental microphones and combine the function of a traditional audio display (audio communication, GPS-based navigation) with audio-processed monitoring of the nearby environment (talk-through system). Such systems are commonly called Communication and Hearing Protection Systems (C&HPSs). Several C&HPSs, such as the Nacre QuietPro, Bose ITH, ATI QuietCom, and Silynx QuietOps™ are available for both military and civilian applications. However they are quite expensive and can each cost several hundred dollars or more. In addition, all of them share some operational limitations including poor directional properties that affect the user's ability to identify the direction of incoming sound and a scaling of the distance to the sound source due to the amplification of external sounds. Despite these limitations, they effectively combine three required functions of audio HMD systems and will be discussed in-depth after the section on hearing protection devices.

Bone conduction systems

An alternative to earphone-based audio displays are bone conduction displays that transmit audio information through the bones of the skull without occluding the ears. Bone conduction systems utilize mechanical vibrators that deliver the audio signals to the listener by vibrating the bones of the skull. When pressed against the skull, the vibrator excites the bones and soft tissues of the head, transmitting the auditory signals through the mechanical pathways of the head into the cochleae. In addition to bone conduction transmitters, audio HMD systems can also use bone conduction receivers (microphones) to convert skull vibrations produced during speech emission into audio signals. Some advantages of the bone conduction microphones over air (boom) microphones are that they are not sensitive to environmental noise and can be placed inconspicuously at almost any place on the head.

Bone conduction audio HMD systems differ from earphone-based audio HMD systems only in the use of bone vibrators in place of earphones as the audio transmitters projecting auditory signals. Such systems can serve as separate audio communication systems or can be built into the helmet. In either case they need to be used with some form of noise protection system when used in noisy environments. Bone conduction systems are very effective audio displays when they operate in quiet and in moderate noise levels (up to approximately 80 dB (A)) without compromising the wearer's auditory awareness of the environment. At higher intensity noises they can be worn with an independent hearing protection system without affecting its operation. The use of hearing protection extends the operational range of bone conduction systems to approximately 110 dB (A). Bone conduction systems can be effectively used by both dismounted and mounted Warfighters because with proper mounting on the Warfighter's head they are not sensitive to vehicle vibrations (Henry and Mermagen, 2004). In addition, when transmitting signals convoluted with the wearer's HRTFs, they can provide directional resolution similar to that of the earphone-based display systems (MacDonald, Henry, and Letowski, 2006). As such they are a viable alternative to other types of audio HMD systems when auditory awareness is of critical concern.

Bone conduction displays, when used in quiet and moderately noisy environments, are inconspicuous, easy to hide, and have minimal effect on situation awareness of the surrounding acoustic environment. Bone vibrators and bone microphones can be used in situations where the listener must monitor acoustic activity in the surrounding environment and does not want anyone to be aware of the use of the communication system. The primary disadvantage of using bone conduction audio display systems in quiet environments is that they can produce some amounts of aerial leakage or excite the device to which they are mounted (i.e., a helmet). Fortunately, aerial leakage is preventable and the designs of military bone conduction audio displays need to significantly reduce or eliminate this leakage.

In high-noise environments, noise has less of an effect on the perception of bone-conducted messages than on the perception of messages emitted through a distal audio display (Knudsen and Jones, 1931). When the auditory signal and the noise are both emitted in the surrounding space their auditory images overlap spatially. When the auditory signal is transmitted through the bones, its phantom source is located inside the head and spatially separated from the noise sources located outside of the head. Spatial separation between the signal and noise sources improves the detection and clarity of the signal.

When a bone conduction system is used in a high noise environment its real value lies in its ability to be worn without interfering with the use of hearing protection devices (hearing protectors). In fact, the presence of hearing protection causes the sounds transmitted by bone conduction to be perceived as louder than when the ears are open. This effect is due to the sound amplification by the cavity of the external ear when it is closed by a hearing protector (Henry and Letowski, 2007). Thus, bone conduction systems worn with hearing protection devices can be used to communicate in noise levels up to approximately 100-110 dB (A) (Letowski, Henry and Mermagen, 2005; Letowski et al., 2004).

The quality of bone conduction displays greatly depends on where and how the vibrators are coupled to the bones of the skull. In general, locating the vibrator close to the cochlea improves the cochlea's response to stimulation (Stenfelt, Håkansson and Tjellström, 2000). However, the effectiveness of stimulation is also affected by the orientation of the axis of stimulation and the interconnections between bones and cartilages of the skull. McBride, Tran, and Letowski (2005) examined eleven locations around the skull to determine the most sensitive head locations for bone vibrator placement. Among the locations tested, the condyle (the bony portion directly in front of the opening to the ear canal) was found to be the most sensitive, followed by the jaw angle, mastoid bone, vertex (top of the head), and temple locations. Figure 5-19 shows relative average differences in the head sensitivity.

In addition to selecting an appropriate location for the vibrator placement it is also important to make a secure and stable contact between the vibrator and the head. The skin lies fairly loosely over the bones of the skull and provides some damping of the skull vibration caused by the bone conduction vibrator. The same damping effect is caused by hair and body fat on the head. In addition, large vibration magnitude and the curved surface of the contact area may cause intermittent and weak contact between the vibrator and the skull. Therefore, the vibrator needs to be pressed against the head with some minimal static force to transmit its vibrations effectively.

The greatest effect of skin damping on vibration transmission is at low frequencies and for the same static force pressing the vibrator against the head the damping effect decreases with an increase in frequency of stimulation. Békésy (1939) reported that at low frequencies around 200 Hz a force of 250G[1] applied over a contact area of 0.5 cm^2 is sufficient to transmit vibrations through the skin without an excessive loss of the transmitted signal. Békésy (1939) also reported that vibration transmission at frequencies above 7000 Hz is no longer affected by the static pressure as long as the pressure exceeds a certain minimum value.

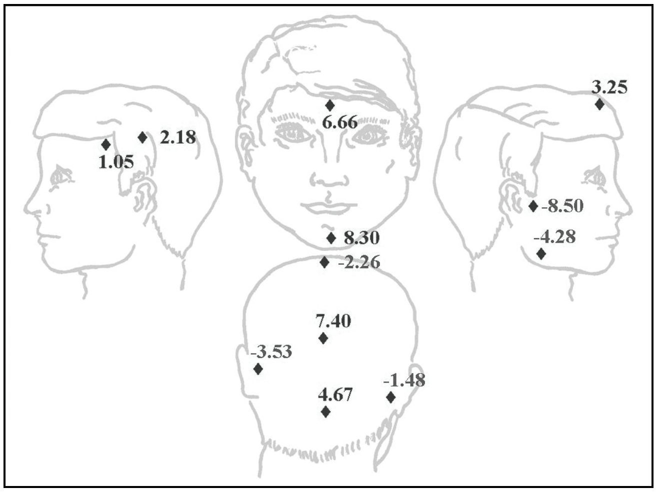


Figure 5-19. Average differences in coupling efficiency of a bone conduction vibrator at various locations on the human head. The locations with negative values are recommended for future applications. Units are in dB (HL) (adapted from McBride, Tran, and Letowski, 2005).

The effect of static force on transmission loss of a 2500 Hz tone is shown in Figure 5-20. For static forces above 500G and a skin thickness of 2.5 mm, the skin attenuates vibration by only approximately 2 dB for frequencies up to approximately 10 kHz. According to data provided by Békésy (1960) the static force of 250-300G is adequate for proper operation of bone conduction display systems, Other authors have recommended the use of similar or slightly higher forces: 200 to 400G (Harris, Haines, and Myers, 1953; Watson, 1938), 300 to 600G (Goodhill and Holcomb, 1955), 350 to 750G (Whittle, 1965). Although large static forces may be desirable for reliable and repeatable coupling of the transducer to the scull, forces exceeding 400-500G can cause physical discomfort for the listener and are therefore not practical for long term use and the force of 250-300G seems to be an acceptable compromise between quality of display and comfort of use.

The area of contact between the vibrator and the skull is another factor affecting effectiveness of bone conduction transmission. Khanna, Tonndorf and Quellar (1976) reported that the perception of vibrations improves with an increase in the area of contact. However, Goodhill and Holcomb (1955) observed better reliability of the threshold data with a vibrator having a contact area of 1 cm^2 than with a comparative vibrator having a contact area of 3.2 cm^2. Thus it seems that the optimal contact area is dependent on the stimulation location. The effect of contact area is also dependent on the signal frequency. Watson (1938) and Nilo (1968)

[1] Letter **G**, used to describe the amounts of static force in this chapter, stands for "Gram-of-force" as opposed to letter "**g**" meaning "gram-of-mass". Gram-of-force (G) is a standard SI metric unit of force. Please note that in this context letter "**G**" does not mean gravity.

observed that the changes in the contact area from $1.1cm^2$ to $4.5cm^2$ had only minimal effect on hearing thresholds at low frequencies but hearing thresholds improved with larger contact areas for frequencies above 2000 Hz (2000 to 7000 Hz). Watson (1938) also noted that a smaller area, on the order of 0.5 cm^2, was uncomfortable to the wearer even with a relatively small (375G) amount of contact force. The concentration of pressure on a smaller area increases the wearer's discomfort and needs to be avoided.

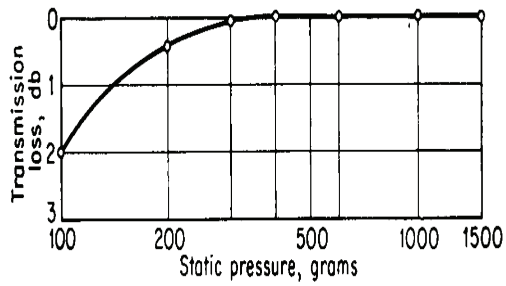


Figure 5-20. The effect of static force (in grams) on transmission loss (dB) of 2500 Hz vibrations transmitted through the skin (Békésy, 1939).

From an operational point of view bone vibrators can be divided into three main categories: head-mounted, in-the-ear, and dental transducers. Head-mounted vibrators are designed to be placed on the surface of the head and secured by a headband or other headgear-type retention system. They are commercially produced by a handful of companies (e.g., Percom, Temco, Sensory Devices, and Oiido) and may have various shapes and sizes. Examples of head-mounted bone vibrators are the Percom 31MIT and Teardrop vibrators shown in Figure 5-21.

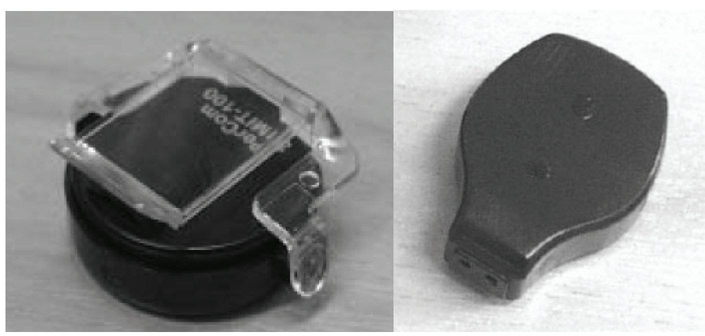


Figure 5-21. Percom 31MIT (left) and teardrop (right) head-mounted vibrators (Courtesy of Percom, Inc.).

Bone conduction systems can also be designed to operate in the ear canal (in-the-ear vibrators). An example of in-the-ear bone conduction display system is the TransEar, developed by Ear Technology Corporation, which has a bone vibrator embedded in the earmold. Such devices are similarly unobtrusive as ITC or CIC hearing aids and may be the devices of choice for creation of spatial bone conduction displays since their point of ear stimulation coincides with the entrance to the ear canal. However, they occlude the ear which negates the primary advantage of bone conduction systems over the earphone-based systems.

The dental vibrator is specially designed for placement in direct contact with the user's teeth. The vibrator can be attached to the listener's tooth or made to be clamped between the teeth as shown in Figure 5-22. Bone vibrators clamped between the teeth have been used as audio display systems by Navy Seals and recreational divers. Mounting a vibrator on a tooth is a challenging operation requiring dental skills and vibrating the teeth for an extended period of time may be harmful to the dental structure. In addition such wired systems are not operationally practical because of the cable connection running through the mouth. However, the concept of a dental vibrator with a wireless connection appears attractive for some special applications (e.g., stealth operations). This technology is still in its early stages of development.

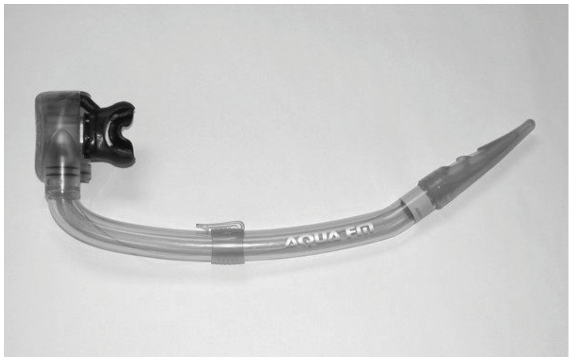


Figure 5-22. Vibrator embedded into the mouthpiece of snorkel (Aqua FM snorkeling system) (Courtesy of AMPHICOM®).

Various forms of bone conduction systems are used by hearing impaired people and serve as alarm devices, hearing aids, and assistive listening devices. Their use as general purpose audio displays is still in its infancy although several types of bone conduction audio HMD systems have been developed for use by firefighters, law enforcement agencies, and special operations forces. The main limitations of the existing systems are excessive nonlinear distortions at high operational levels and lack of optimized head interfaces. The list of major manufacturers of bone conduction devices and their main products is shown in Table 5-4.

A comparison of the main advantages and disadvantages of air conduction (earphone-based) and bone conduction audio systems is provided in Table 5-5. It is evident from Table 5-4 that the general problem in selecting an audio HMD system is the proper balance between auditory awareness of the environment and needed hearing protection. Since this balance is affected by the type of hearing protection device incorporated in the audio display system, it is important to consider all available options.

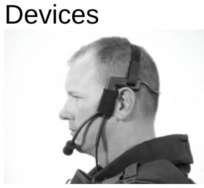
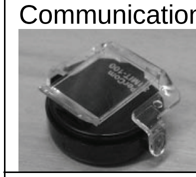
Hearing Protection Devices

Susceptibility to noise varies among people, but exposure to high levels of noise can cause permanent hearing impairment (see Chapter 11, *Auditory Perception and Cognitive Performance*).

Exposure to continuous noise at levels exceeding 85 dB (A) for 8 hours or more causes noise-induced hearing loss (NIHL). Such continuous noise sources include aircraft, tracked vehicles, and power generators. The resulting hearing loss is initially temporary and becomes gradually permanent after prolonged exposure. Similarly, exposure to impulse noise, such as weapon fire or bomb explosion, with peak sound pressure level exceeding 140 dB can cause permanent damage (acoustic trauma) to the hearing system. Very high impulse levels can also mechanically damage the tympanic membrane and soft tissue organs such as the lung or liver.

According to Department of the Army Pamphlet 40-501 *Hearing Conservation* (DA PAM 40-501, 1998), the first indication of the early state of noise induced hearing loss is decreased sensitivity in the range of frequencies above 2 kHz. Other symptoms include tinnitus (ringing in the ear), temporary muffling of sound, and a feeling of fullness in the ear, stress, and fatigue. More information on hearing loss is included in Chapter 11. The reader is also referred to the U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM) website (USACHPPM, 2006) that is an excellent source of information on hearing conservation.

Table 5-4.
Commercially available bone conduction communication systems.

Company	Product Name	Uses	Contact Information
Temco 	-HG-16 (boom microphone) -HG-17 (bone microphone) -HG-21 -HG-21D -FM-200 (gas mask) -Band-aid BC System	Stealth communication Gas masks Construction workers	www.temco-j.co.jp
Radioear / Sensory Devices 	-Radioear B-70/ 71/72 -BC System RE-1	Hearing testing Hearing aids Special forces Law enforcement	www.sensorydevices.com
Aliph 	-Jawbone	Mobile wireless communication (Bluetooth)	www.jawbone.com
PerCom / Audio Communications 	-Teardrop MIT -31MIT -17 MIT	Firefighters Law enforcement	www.audiocomms.co.nz
Vonia, Pegaso 	-BH-10/20/80 -EZ-500 -EZ-2000 -EZ-2000 -EZ-3000 -EZ-4200 -EZ-10/20/80	Telemarketing Entertainment Stealth communication Underwater communication	www.vonia.co.kr www.pegaso.co.jp

Military environments are predominantly high noise environments. Common military vehicles and equipment generate very high noise levels requiring hearing protection as specified by Department of Defense Instruction 6055.12 (DOD, 1991) and DA PAM 40-501 (1998). Continuous noise levels in armored personnel carriers (M113 and Bradley Fighting Vehicle) are the highest noise levels among Army vehicles. Internal noise in an idling vehicle is approximately 92 dB (A). Noise levels in these vehicles increase with speed and can reach 118 dB (A) at 40 mph. Similarly, noise levels inside military helicopters are higher than 100 dB (A). In helicopters, such as the Blackhawk and Apache, noise levels at the pilot's position can reach 106 and 104 dB (A), respectively.

Table 5-4. (Continued)
Commercially available bone conduction communication systems.

Company	Product Name	Uses	Contact Information
IntriCon 	-Bone conduction headset LV23	Warehouse operations Train operations Safety personnel Audio production crew	http://www.intricon.com.sg
Oiido Equipment 	-Bone conduction headset	Law enforcement Telemarketing	http://www.oiido.com/
Atlantic Signal LLC 	-Tactical Headset System MH180-H -Tactical Headset System MH180-S -Tactical Headset System MH-3	Military Law enforcement Gas masks	http://www.mhseriestacticalheadsets.com/index2.html
Tactical Command Industries 	-Tactical Assault Bone Conduction (TABC) Headset -SPEC-OPS II Binaural Bone Conduction Headset	Law enforcement Special forces	www.merchantmanager.com/tactical/headset_products.htm

Table 5-5.
Advantages and disadvantages of air conduction earphones and bone conduction vibrators.

Audio Display Systems	Advantages	Disadvantages
Air conduction systems	• Mature technology • Wide range of styles, quality, and prices	• Occlude ear canal • Interfere with hearing protection devices and must provide proper noise attenuation • Some systems provide insufficient ear ventilation and/or irritate the ear
Bone conduction systems	• Do not occlude ear canal • Inconspicuous to use (easy to hide) • Less susceptible to ambient noise effects than distal display systems; • Do not interfere with hearing protection devices	• Not quite mature technology, • Aerial sound leakage • Excessive static pressure may cause discomfort • Require a separate hearing protection system

Military personnel are also exposed to extremely high impulse noise levels from weapon firings. For example, USACHPPM (2006) reports at the gunner's position the peak sound pressure level for the Multi-Role Anti-Armor Anti-Personnel Weapon System (MAAWS) recoilless rifle is 190 dB SPL and for the Light Antitank Weapon M72A3 it reaches 182 dB SPL. Even small arms weapons like the M16 rifle and M9 pistol produce impulse noise levels reaching 157 dB (peak) at the shooter's ear, far above the hazardous level of 140 dB (peak). The methods of measuring impulse noise levels are the subject of the ANSI standard S12.7 (ANSI, 1986).

Military Standard MIL-STD 1474D, *Design Criteria Standard: Noise Limits* (DOD, 1997) is the governing noise control document for military materiel used by the U.S. Department of Defense. It specifies noise limits to equipment designers and manufacturers. It is intended to cover typical operational conditions. Required noise limits must not be exceeded if the materiel is to be acceptable. This standard is based upon requirements for hearing damage-risk, speech intelligibility, aural detection, state-of-the-art of noise reduction, and government legislation. The standard includes requirements for: steady-state noise, aural non-detectability, community annoyance, impulse noise, shipboard equipment noise, and aircraft noise (DOD, 1997).

Recent studies indicate that current noise exposure standards and design guidelines, as described in requirement 4 of MIL-STD 1474D (DOD, 1997) for impulse generating weapons are seriously in error. To overcome these limitations, the Army Research Laboratory (ARL) developed a mathematical model of the human auditory system which predicts the hazard from any free-field pressure and provides a visual display of the damage process as it is occurring. The model – Auditory Hazard Assessment Algorithm for the Human Ear (AHAAH) – is a powerful design tool which shows the specific parts of the waveform that need to be addressed in machinery and weapon design. This unique model is the only method of assessing noise hazard for the entire range of impulses relevant to the Army. This mathematical model calculates stress in the inner ear based on head orientation, hearing protection [manikin or Real-Ear-Attenuation-at-Threshold (REAT) measurements], aural reflex, and stapes displacement limitation. Risk is calculated based upon a hypothesis that damage to the hair cells in the cochlea correlates to a mathematical function of the number of and amplitude of basilar membrane displacements in a manner analogous to mechanical fatigue of solid materials. (Price, 2005; 2007) Additional information about the AHAAH is available at the AHAAH Website (http://www.arl.army.mil/ahaah/).

The primary means to protect Warfighters from harmful noise levels are hearing protection devices (HPDs). Communication earphones covering the ear or occluding the ear canal protect the user to some degree against the harmful effects of external noise. However, generally, the resulting level of protection is not satisfactory to prevent hearing loss in high level noise environments. In addition, in many operational situations the user may not have a communication system covering the ears. Therefore, in designing audio HMDs it is important to focus on hearing protection offered by both the communication systems and by the dedicated hearing protection devices alone.

Various HPDs and classifications of HPDs are available, but the two most important dichotomies of all HPDs divide them into passive and active devices and into linear and non-linear devices.

Passive linear HPDs are sound barriers that reduce the overall noise level by covering the entire ear or by insertion into the ear canal. They are generally in the form of an earmuff surrounding the ear or an earplug that blocks the ear canal and typically provide 25 to 35 dB of noise reduction, if worn correctly. They can also be shallow conic earplugs connected by a headband and providing some ear occlusion at the entrance to the ear canal. Such devices are called semi-aural HPDs or ear canal caps.

Earmuff cups (shells) may have various depths, shapes, sizes, and weights. The earmuff system consists of two earmuff cups supported by an over-the-head headband, attached to a safety hardhat, or mounted in a helmet. The positive features of the earmuff-type HPDs are their reliable and repeatable level of protection, easy fit, and comfort of wear. They are also easy to don and doff. However, they interfere with wearing glasses, protective masks, and other safety equipment (Wagstaff, Tvete and Ludvigsen, 1996).

Examples of passive earplug devices are foam earplugs, preformed flanged rubber earplugs, and custom-made rubber plugs. Each type of these devices has its own advantages and disadvantages. They are supplied in various sizes and have different sound attenuation characteristics. Foam earplugs may have different sizes and shapes (e.g., cylinder, cone). They are generally disposable HPDs and therefore require constant resupply. The flanged earplugs are soft rubber earplugs that may have one to four flanges (see Table 5-6). In addition, some of the flange earplugs may have built-in filters to equalize their frequency response (e.g., Earlove Earplugs, Musicians Earplugs ER-20) and can be custom-made for individual users (e.g., Musicians Earplugs ER-9, ER-15, and ER-25).

The semi-aural devices, or canal caps, offer less protection at medium frequencies than either earmuffs or earplugs and similar protection at high frequencies. Their typical average attenuation of noise is 10-15 dB. They are supported by a light headband that can be worn under the chin or behind the head, permitting semi-aural devices to be used together with various types of safety equipment. To some degree they facilitate speech communication in noise but are uncomfortable when used for long time periods due to the pressure exerted at the entrance to the ear canal.

Attenuation characteristics of linear passive HPDs are frequency dependent. Their attenuation increases with frequency due to increasing transmission loss through the solid material of the HPD and the elimination of ear canal resonances when earplugs are worn. The HPDs that attempt to provide uniform attenuation across a wide frequency range normally achieve it by some reduction of attenuation at middle and high frequencies. Typical passive earmuff-type HPDs provide attenuation of less than 15-20 dB at frequencies below 200 Hz and approximately 40-50 dB at frequencies above 3000 Hz for well-fitted HPDs. Passive earplug-type HPDs, especially foam earplugs, provide greater and more uniform attenuation at low frequencies but attenuation similar to earmuffs at high frequencies. It must be stressed that noise attenuation provided by earplug-type HPDs varies dramatically with the quality of fit and the depth of insertion. For example, poorly fitted flange earplugs may provide almost no attenuation at low frequencies. The overall protection provided by single or double hearing protectors cannot exceed 50-60 dB due to bone conduction pathways that bypass the ear protection.

Numerous field evaluations of HPDs suggest that the actual noise protection offered by HPDs in real operational environments is much less than the manufacturers’ published data. This difference is mainly due to improper fit of the devices, wear and tear, and inappropriate size. Thus, selection of any HPD to be used as part of an audio HMD system must take into considerations both the official manufacturer's data and the field reports. Various types of passive HPDs, together with their advantages and disadvantages, are shown in Table 5-6.

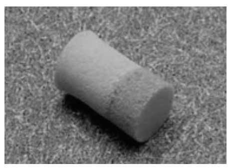
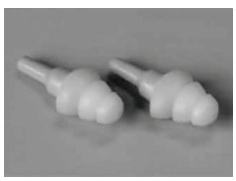
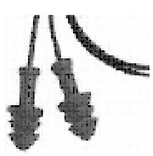
In general, passive linear HPDs are inexpensive (except for custom-made earplugs), easy to use, and effective if fitted correctly. They can be used separately or built into a helmet. For example, most military helmets with integrated communication and audio display systems, such as the HGU-56/P or SPH-4B flight helmets and combat vehicle crewman's (CVC) helmet, provide hearing protection with a closed-air (closed-back) circumaural earcup style of earphone. Additional protection against high levels of noise can be achieved by wearing a combination of earmuff and earplug devices (double hearing protection). Similarly, when using an earphone-based audio display, additional noise protection can be added by wearing the earplugs in combination with the earphones but such use will also reduce the level of the auditory signals generated by the earphones.

As described previously the major deficiency of earmuff-type HPDs is insufficient attenuation of low frequency industrial and military noise levels. Conversely, both earmuff-type and earplug-type HPDs provide relatively high attenuation of high frequencies, which adversely affects speech communication in quiet and in low levels of noise. In general, all passive linear HPDs interfere with speech communication and prevent detection of low-level sounds in the surrounding environment, thereby compromising situation awareness. This deficiency is addressed by different types of level-dependent HPDs. Level-dependent HPDs are a class of non-linear HPDs that significantly attenuate hazardous high intensity impulse sounds while passing low intensity sounds, such as conversational speech, with minimal attenuation. Level-dependent reduction of noise levels can be achieved by either passive or active reduction techniques.

Passive nonlinear HPDs are vented devices that have small orifices, diaphragms, or valves built into the HPD. These increase the protection provided against impulse noise as the noise level exceeds a certain threshold, usually 120 dB SPL (Shaw, 1982). Above this threshold, high noise levels result in a turbulent flow of acoustic energy through the non-linear element of the protector effectively closing the vent. At noise levels below this threshold, the protector acts as a regular vented earmold (earcup) providing usually less than 20 dB noise attenuation at high frequencies and none or very little attenuation at low and middle frequencies below 1000 Hz (normally less than 5 dB). Such protection characteristics of level-dependent HPDs facilitate speech communication and awareness of the environmental sounds in quiet and moderately noisy environments while protecting the Warfighters from high intensity impulse sounds from their own and enemy weapon fire. An additional feature of some passive level-dependent HPD is that they function as a pressure valve that slows down rapid changes of atmospheric pressure typically experienced during take-off and landing of aircraft. Examples of passive level-dependent HPDs include V-51R (American Optical), Bilsom ISL 655 (Bilsom), Ear Defender EP3 (EarPro), EarGuard (Cirrus), and the Combat Arms Earplug (Aearo).

The Combat Arms Earplug (CAE) is a level-dependent device designed for military operations. The earplug is produced in both single-end and dual-end versions shown in Figure 5-23. The dual-end version can be used as either a linear (olive drab plug) or non-linear (yellow plug) HPD. A small mechanical filter with a calibrated orifice is embedded in the non-linear end of the plug. When this end is inserted into the ear canal, the CAE passes the low-intensity, low-frequency sounds with as little as 5-8 dB attenuation and allows the user to hear normal conversation, footsteps, or vehicle noise while to some degree attenuating high frequency energy of the sounds. The attenuation of the plug rapidly increases at high noise levels starting at approximately 120 dBP SPL and reaches full peak attenuation of 25 dB at approximately 190 dBP (Dancer and Hamery, 1998), providing wideband hearing protection against the dangerous high-level energy of weapons fire and explosives. The linear portion of the CAE is used for hearing protection from high level steady-state noise environments such as those created by armored vehicles or aircraft where situation awareness is not a priority. It provides approximately 35 dB insertion loss at low and middle frequencies with insertion loss gradually increasing above 1000 Hz. At very high sound intensity levels exceeding 170 dB (peak) both types of the CAE earplug provide similar attenuation for frequencies above 250 Hz.

Table 5-6.
Advantages and disadvantages of various types of passive HPDs.

HPD	Advantages	Disadvantages
Earmuffs 	-Easy to fit -Good noise reduction -Comfortable to wear if light and properly adjusted -May protect pinnae from exposure to adverse environments and burns, e.g., from improvised explosive devices (IEDs).	-Bulky and heavy -Interfere with other headgear
Foam Earplug 	-Inexpensive -Good noise attenuation -Compatible with other headgear	-May cause skin irritation -Easily become dirty -Hard to fit
Musicians Earplug 	-Provide relatively uniform attenuation across wide frequency range -Provide protection without adversely affecting sound quality -Comfortable, especially if custom fit -Compatible with other headgear	-Expensive -Hard to replace -High maintenance (require earwax cleaning from inserted filter)
Single-Flange Earplug 	-Reusable -Inexpensive -Compatible with other headgear	-Poor noise reduction -May cause skin irritation -Difficult to insert correctly -May move with normal jaw movement
Triple-Flange Earplug 	-Good noise reduction -Reusable -Inexpensive -Compatible with other headgear	-Difficult to insert correctly -May cause skin irritation -May move with normal jaw movement

The right panel of Figure 5-23 shows a new version of CAE developed by Aearo in cooperation with the US Army Research Laboratory (ARL). Its main advantage over the old version is a rotating mechanical switch to open and close the filter permitting change from one type of hearing protection to another using a single-end plug.

Active level-dependent HPDs use an external microphone and an internal loudspeaker (earphone) to bypass passive attenuation offered by the HPD when environmental noise is below a certain threshold. Such an electroacoustic system enables the wearer to hear environmental sounds and spoken messages while wearing HPDs. When the threshold level is exceeded, the bypass pathway is closed and the HPD operates as a passive HPD. Most of the amplifiers built into active level-dependent HPDs have a volume control that allows the user to adjust the amplification of the bypass system. The system may also have a built-in sound compressor or automatic gain control (AGC) circuitry that automatically decreases amplification of the external sounds as the level increases. Examples of such devices are the Peltor Tactical 7, BBP Ltd. EP171, and Bilsom 707 Impact earmuffs and the Communications Earplug (CEP). When a radio or other external audio input is added to an active level-dependent HPD it becomes a C&HPS mentioned earlier in this chapter and described more extensively below.

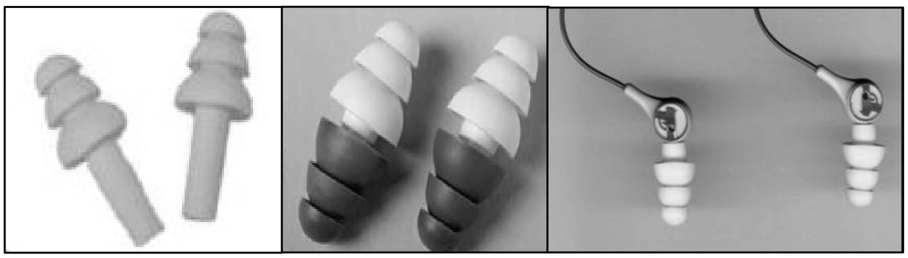


Figure 5-23. Combat Arms Earplug (CAE). Single-end version (left), dual-end version (middle), current single-end version (right) (U.S. Army Research Laboratory photos).

The advantages of a passive level-dependent HPD, such as CAE, when compared to similar active devices, are their low cost, light weight, ruggedness, and relatively easy maintenance. In addition, there is no battery requirement and no tangled wires. However, there are some issues with comfort of use of these devices if they are not available in various sizes (Scharine, Henry, and Binseel, 2005). The level-dependent passive HPD effectively complements bone conduction HMDs and together they are a viable alternative to C&HPS. Selection of a specific solution depends on the military or civilian system that is being supported and specific requirements of the missions to be performed.

Active noise reduction (ANR) devices are another class of non-linear HPDs. ANR is a method of reducing the level of environmental noise by phase cancellation. In the ANR system the surrounding noise is monitored by an environmental microphone, reversed in phase, and emitted back to the listener in an attempt to reduce the overall noise level. Such a noise-canceling scheme reduces noise only in selected locations but is a very good solution for HPDs. A general concept of an ANR system applied to an audio HMD is shown in Figure 5-24.

In the system shown in Figure 5-24, environmental noise is monitored by the external microphone (Noise Reference Mic) mounted outside of the passive HPD system. Captured noise is reversed in phase, signal processed, and emitted under the HPD by an audio transmitter (Headpone Transducer). The internal microphone (Error Mic) located close to the entrance to the ear canal monitors the overall noise level under the earmuff (earcup) and provides a differential signal that controls the amount of out-of-phase noise needed to minimize the overall noise level. The HPDs and audio HMD systems using ANR systems are frequently referred to as noise-canceling earphones or active noise-canceling earphones. Examples of audio displays using ANR technology are the Aiwa HP-CN6, AKG K-28NC, Bose QuietComfort 2, Phillips HN-110, Sennheiser PCX 250, and Sony MDR-NC10 systems. Similar ANR schemes are also incorporated into earbuds and earplugs (e.g., Etymotic ER-61C, Philips HN 060, Sennheiser CX-300, and Sony MDR-NC11).

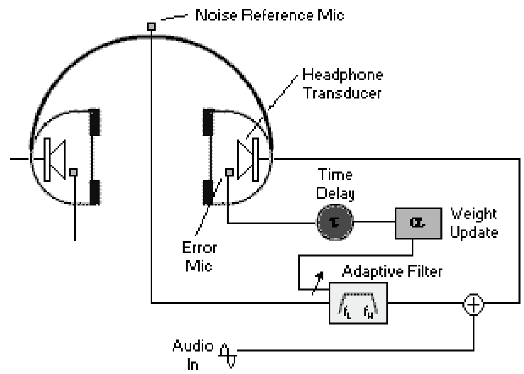


Figure 5-24. Active-noise reduction system incorporated in an audio HMD (Moy, 2001b).

The signal processing circuitry built into the ANR system usually consists of a phase inverter, filter, summer, and time delay unit. Due to the fact that environmental noise usually changes considerably both in time and in space, the filter is typically an adaptive system that self-selects optimum listening conditions. The most common form of adaptive filter used in noise-canceling earphones and hearing protectors is a finite impulse response (FIR) filter with a least mean square (LMS) algorithm. To describe such an audio HMD system the term "adaptive noise-canceling system" is frequently used.

Active Noise Reduction systems are very effective at low frequencies, providing up to 20-25 dB of noise reduction (Gowers and Casali, 1994; Nixon, McKinley, and Steuver, 1992). They are much less effective at high frequencies since reducing high frequencies requires more expensive computation. However, a combination of a passive earmuff with an ANR system provides relatively uniform high noise attenuation across a relatively wide frequency range. Such systems are becoming very common and popular. The ANR system of such devices provides the primary defense against low-frequency noises generated by engines, fans, and motors while the passive system provides the main defense against mid- and high-frequency noises generated by such devices as gas valves, pneumatic devices, saws, and power tools. It should be noted, however, that while ANR earmuffs offer a significant advantage over passive earmuffs, their total noise attenuation is similar to that offered by custom-molded earplugs (Christian, 2000). Therefore, for noise protection purposes both types of devices offer similar protection.

Traditionally, ANR HPDs have been found only in the earmuff-type of audio HMD systems. Today, however, ANR systems can also be found in supra-aural (e.g., Bose QuietComfort3) and in-the-ear devices (e.g., Panasonic RP-HC50E-A and Philips SHN 2500). However, both types of ANR systems integrated with audio HMD systems have some shortfalls. Earmuff-type ANR systems degrade both hearing protection and speech intelligibility when worn with glasses or CB (chemical/biological agent) mask (Mozo and Murphy, 1997). In some designs, ANR systems can affect the operation of the audio transducer distorting the desired signal or introducing additional high frequency noise (Wikipedia, 2007).

One of the problems with comparing various HPD systems is that noise attenuation provided by linear HPD must be measured using a standardized threshold shift method (ANSI, 1997a). However, attenuation provided by ANR systems is measured using the microphone-in-real-ear method (ANSI, 1995a) because of the low-level wideband noise normally produced by ANR systems (Mozo, 2001). Similarly, the passive nonlinear hearing protectors must be measured with the microphone method. Conversely, this method is not suitable for linear earplugs due to the difficulty with insertion of a microphone into the ear canal without compromising earplug attenuation. Unfortunately, for the devices for which both measurement methods can be used, the differences between measured attenuations can exceed 10 dB and are not uniform across the frequency range (Lancaster and Casali, 2004; Neitzel, Somers, and Seixas, 2006).

Communication and Hearing Protection Systems

Communication and Hearing Protection Systems (C&HPS) are modular audio HMD systems that can be worn with or without the helmet. As discussed above, they can also be permanently installed in helmets such as the aviation helmets (e.g., SPH-4) or tanker helmets (e.g., CVC). They typically offer approximately 20 dB of noise reduction. The level of noise reduction can be further increased if a communication helmet is combined with another kind of hearing protector (e.g. EAR earplug) or communications plug (e.g., CEP). For example, the noise level under an SPH-4B helmet earcup is approximately 91 dBA in a UH-60 helicopter flying at a speed of 120 knots and drops to approximately 69 dBA when an EAR™ foam earplug is added (DA FM 1-301, 2000). However, from the operational standpoint, the properties of C&HPS are very similar whether the system is built into the helmet or worn with the same helmet as a separate system.

There are several commercial off-the-shelf (COTS) C&HPS available offering different solutions to hearing protection and auditory awareness requirements. Available devices utilize both circumaural-earphone and insert-earphone types of design. Advantages and disadvantages of both types of hearing protection systems are similar to those listed in Table 5-5. Examples of the circumaural-earphone systems are the Bose ITH and MSA/Sordin Gen II systems. Examples of the insert-earphone systems are the CEP and Nacre QuietPro.

The Bose Improved Tactical Headset (ITH) is an earmuff-type communication system designed to protect Warfighters' hearing (up to 95-plus dB) while allowing them to communicate in the high noise of the M1114 up-armored HMMWVs (High-Mobility Multipurpose Wheeled Vehicles) and other light tactical vehicles being used by the U.S. Army. It can be secured on the head using an over-the-head headband and/or behind-the-head mounting strip. The ITH has two (left and right) forward-facing pass-through microphones and is designed to fit under the Advanced Combat Helmet (ACH). It provides hearing protection through both active and passive noise reduction of approximately 25 dB.

The MSA/Sordin Gen II is an earmuff-type headset that provides noise attenuation of approximately 25 dB. The headset has talk-through capability with a volume control and two (left and right) forward-oriented pass-through microphones. Both the Bose ITH and Sordin Gen II are shown in Figure 5-25.

Figure 5-25. Bose ITH (left) and Sordin Gen II (right) communication and hearing protection systems (Courtesy of Bose Corporation and MSA).

Another company offering earmuff-based C&HPS is Sennheiser. Its WACH 900 (Warrior Advanced Capability Headset) is an earmuff-type ANR and stereo talk-through system. It provides 15-20 dB wideband attenuation with the ANR system turned on. The system is designed for dismounted infantry and can be worn under a helmet or with a respirator. The stereo talk-through capability provides situation awareness. Another earmuff-based C&HPS system offered by Sennheiser is the SNG 100. The SNG 100 is designed for combat vehicle crewmembers and provides noise attenuation of 25-40 dB with ANR. Sennheiser also offers the SLC 110 system, which is an in-the-ear, passive, non-linear, noise reduction, militarized headset. It comes with earplug and concha tips that lock the headset in place. Situation awareness is enhanced by opening an acoustic port that bypasses the attenuation of the earplug. With the port open, impulse and gun-blast noise is attenuated using a non-linear filter that passes low intensity sound with a small loss but attenuates impulse noises by up to 30 dB. All these systems are shown in Figure 5-26.

The CEP was developed at the U.S. Army Aeromedical Research Laboratory (USAARL) for use with the aviator helmet. An expanding foam earplug is attached to a threaded hollow tube extending from the transducer. While the foam attenuates the ambient noise (Noise Reduction Rating - NRR = 29.5 dB), the tube transmits the sound from the transducer to the ear canal. When used on its own it provides noise attenuation from approximately 30 dB at low frequencies to approximately 45 dB in the 4000 Hz to 8000 Hz range. When used in combination with the aviator helmet, the earplug adds approximately 10 dB to the hearing protection provided by the helmet while improving radio communication clarity. Since the signal does not compete with the environmental noise, less audio gain is required to hear voice communication.

Figure 5-26. Sennheiser communication and hearing protection systems (C&HPS) (Courtesy of Sennheiser USA).

The Communication Enhancement and Protection System (CEPS) is an improved CEP with an additional capability of situation awareness, developed for the dismounted Warfighter or for use with the aviation helmet. The microphone providing the input signal to the radio for communication is also used to provide ambient sounds to the user. The microphone permits the Warfighter to hear environmental sounds and voice communication during dismounted operations. The system allows the user to control the level of sound from the external microphone with up to 36 dB of gain. With the lowest gain setting, the sound level to the user is limited to 95 dB (A). When impulse sound levels exceed 128 dB (peak), circuitry in the CEPS automatically disables the microphone to prevent any harmful amplified sound from reaching the ears (Mozo, 2004). The device is powered by two AAA batteries and weights only approximately 2.2 ounces (62 grams). Both the CEP and the CEPS are shown in Figure 5-27.

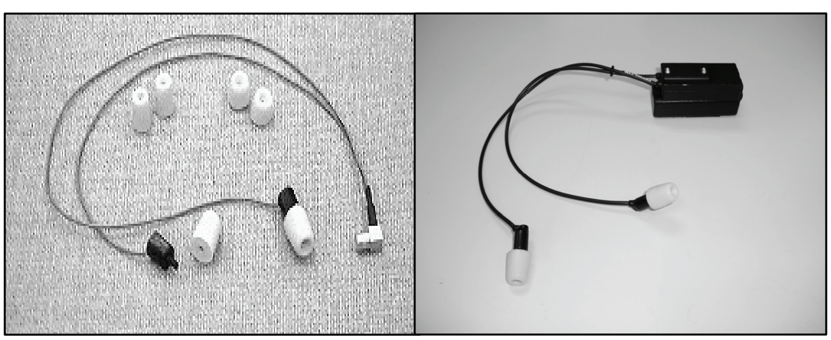


Figure 5-27. CEP (left) and CEPS (right) (Courtesy of USAARL).

The U.S. Air Force Research Laboratory (AFRL) developed a custom-molded earplug-based C&HPS called the Attenuating Custom Communication Earpiece System (ACCES). The system is shown in Figure 5-28. A small speaker is embedded in a silicon earpiece which is made from impressions of each individual user's ear canals. Its custom fit provides better comfort and higher noise attenuation compared to generic insert earphones. The ACCES can be worn alone as an earplug or plugged into the flight helmet (aircrew) or headset (ground support crew) for double protection and intercom communication.

QuietPro is an in-the-ear non-linear ANR system developed by Nacre, AS (http://www.nacre.no/). It includes one outer microphone, one inner microphone, and one miniature loudspeaker for each ear. All three elements are embedded in an earplug. The QuietPro provides binaural talk-thru, radio communication and hearing protection. Specifications for the device indicate that the active noise reduction system attenuates approximately 14 dB, targeted at frequencies between 63 to 500 Hz. Overall passive attenuation is 34 dB; and the loudspeaker is capable of reproducing a sound pressure level >125dB when combining digital ANR, talk-through audio and communication. The QuietPro system is shown in Figure 5-29.

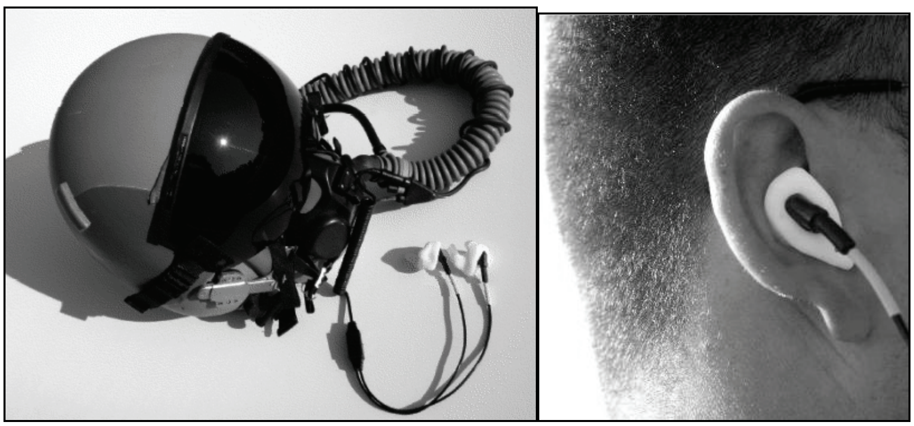


Figure 5-28. ACCES, attached to HGU-55P flight helmet (left) and in the ear (right) (Courtesy of U.S. AFRL).

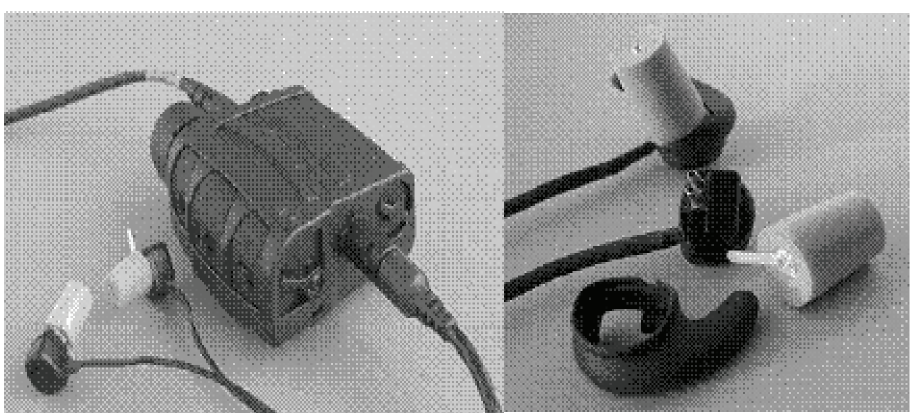


Figure 5-29. QuietPro (Courtesy of Nacre AS).

Earplug-based C&HPS have several advantages over the earphone-based systems. Proper fitting of in-the-ear style earphones provides both comfort and maximum noise isolation. These properties result in good intelligibility of transmitted speech and long-term user satisfaction. However, in-the-ear devices create hygiene-related problems in harsh environments and require careful fitting. There is also a psychological factor - not everybody wants to put something in their ears.

Acoustically Transparent Helmet

In modern warfare, military personnel require protection not only from kinetic threats, directed energy, and loud noises, but also from chemical and biological weapons. Therefore in early the 2000s the U.S. Army considered development of an encapsulating helmet for the Objective Force Warrior. This helmet design was intended to integrate with the biochemical protective suit and provide whole body chemical and biological protection. However, the encapsulating helmet creates a profound acoustic challenge. It would greatly attenuate sounds and distort auditory cues or even completely prevent the sounds from reaching the Warfighter's ears.

In an effort to find solutions to these negative acoustic affects of total encapsulation, various private companies, universities, and government agencies conducted research studies to develop an acoustically transparent helmet. The general concept of the acoustically transparent helmet is to place a network of microphones on the helmet shell and deliver captured spatial sound to the Warfighter's ears in order to restore natural hearing. The U.S. Army Natick Soldier Research, Development, and Engineering Center (NSRDEC) and the U.S. Air Force Research Laboratory jointly funded a project entitled *Concept and technology exploration for transparent hearing systems* that was executed by the Scorpion Audio Team, comprised of representatives of AuSIM, Inc., Fakespace Laboratories, Sensimetrics Corporation, and Boston University (Chapin et al., 2003). Figure 5-30 shows the Natick "Scorpion R2" helmet design, with potential microphone locations to capture the directional characteristics of external sounds.

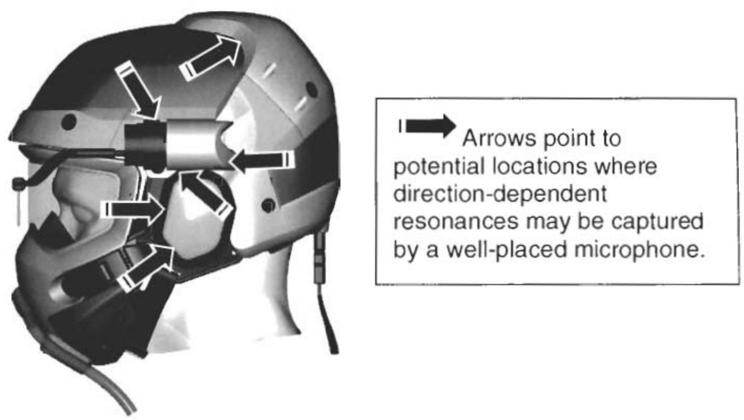


Figure 5-30. Natick "Scorpion R2" helmet design (Courtesy of Scorpion Audio Team, 2003).

Fig 5-31 shows the concept of the AuSIM 3D audio headset which is commercially available from the company. Using information captured by the microphone array, head tracker and HRTFs, AuSIM used their proprietary AuSim3D software to process the incoming sound and re-introduce its spatial cues according to head orientation. This system was used by AuSIM and Sennheiser Government Systems group in an attempt to build an encapsulating helmet providing Warfighters with restored natural situation awareness.

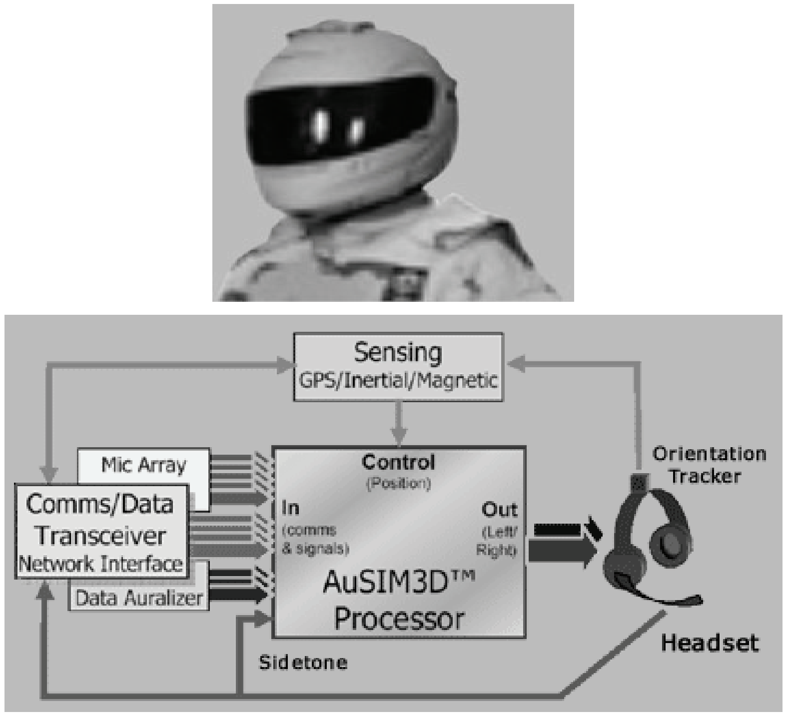


Figure 5-31. AuSIM 3D audio system for situation awareness and Sennheiser's conceptual transparent hearing helmet (Courtesy of AuSIM and Sennheiser Government Systems).

In another effort the ARL and Adaptive Technologies, Inc. (ATI) investigated natural hearing restoration solutions using a motorcycle helmet. Figure 5-32 shows prototypes of motorcycle helmets with mounted microphones built by ATI and ARL. All these efforts were terminated due to lack of funding caused by a change in the U.S. Army strategic vision.

Figure 5-32. Motorcycle helmet with microphone array built at ATI (left) and at ARL (right) for natural hearing restoration research (Courtesy of ATI and U.S. Army Research Laboratory).

Audio HMD System Design Issues

Design and selection of audio HMD systems needs to conform to general rules of human-centered design principles. Human-centered design treats the user as the final element of the HMD system rather than the screen of the monitor, membrane of the earphone, or moving plunger of the bone vibrator. Therefore, the engineering details of the display system needs to be specified not in terms of technology-based sensory stimulation parameters but in terms of perceptual and cognitive demands of the user and worked backwards toward sensory stimulation specifications.

General requirements

Human ability to respond to changing environments and to carry on required tasks requires free and effortless head movements that are minimally impeded by the additional weight of the headgear. This requirement is especially important for users that are moving on their feet and are not supported by a moving platform such as vehicle or aircraft. Therefore the mass of the audio HMD system should be made as low as possible and should not exceed 1.2 lb (545 grams) including all required batteries and not including the mass of the interface cables (Program Manager [PM] Soldier Warrior, 2007). The size and shape of the audio system should not interfere with the mission of the user including driving, crawling, parachute jumping, shooting, and the use of other headgear (e.g., video HMD or protective headgear).

An audio HMD system should be designed to fit easily in the ear or over the user's head without the need for extensive adjustment. The user should be able to don and doff the audio system in less than 10 seconds and without taking off or disengaging other personal equipment or taking off the gloves. The system should be ergonomically designed to self-set and stay stable in the operational position for the length of the user's mission. The parts of the system touching the user's skin should not create any adverse skin reaction or cause health hazards when used in operational environments. All basic mechanical, chemical, and electrical operational safety requirements for personal equipment should be met for the specified temperature, humidity, and atmospheric pressure ranges.

Discussion of various types of audio displays and hearing protection systems conducted in the previous sections clearly indicates that the design of an audio HMD system meeting all three basic operational requirements described in the first part of the chapter is a challenging effort. Further, in selecting one of many available audio systems for specific applications, there are many technical nuances and design compromises that need to be considered in order to develop a cost effective and ergonomically correct solution. Among the technical decisions that need to be made are audio transmitter technology, system interface, comfort, fit, weight, durability, mounting techniques, audio-visual integration, and compatibility with other equipment.

Audio transmitters

Both earphone-based and bone conduction audio display systems can utilize the same type of electroacoustic transducers (e.g., dynamic, piezoelectric, and electret transducers) that convert electric energy to mechanical energy and, subsequently, to acoustic energy. The main operational difference between the audio transmitters used in these two types of display is the difference in load impedance exerted upon the transmitter by air and bones of the head of the wearer. This difference, however, has a huge impact on the technical requirements for both types of transmitters. Thus, despite some physical and operational similarities between the transmitters used in both these types of systems, they differ substantially as specific technical solutions.

Magnetoelectric transducers

Most transducers used in audio HMD systems use either moving coil technology or the piezoelectric principle. Transducers using moving coil technology are called magnetoelectric or dynamic transducers. A dynamic transducer consists of a diaphragm connected to a coil of wires moving in the air gap of a permanent magnet. When an alternating electric current representing an audio signal is applied to the coil it creates an alternating magnetic flux. This magnetic flux interacts with the magnetic field of the permanent magnet pushing or pulling the coil and the attached diaphragm depending on the direction of the resulting magnetic force. The movement of the diaphragm creates changes of air pressure resulting in sound being projected to the ear. A drawing of the cross-section of simple dynamic earphone is shown in Figure 5-33.

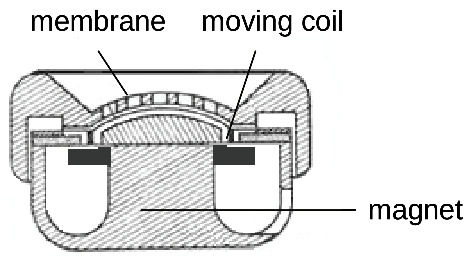


Figure 5-33. A cross-section of a simple dynamic earphone (adapted from (Kacprowski, 1956).

The diaphragm of the dynamic transmitters used in audio HMD systems is typically made of light-weight and stiff foil which requires a large radiating surface to reproduce low frequency signals. This large radiating surface also projects high frequency energy very efficiently. The magnitude of the movement of the diaphragm determines the loudness of the reproduced sound. The permanent magnets used in modern dynamic transmitters are usually neodymium and ferrite magnets. In order to improve the sound quality and the life of the transducer, some companies damp unwanted resonant frequencies and reduce the heat from the moving coil by introducing ferrofluid into the air gap of the transducer. Ferrofluids have the fluid properties of a liquid and the magnetic properties of a solid. A picture of a magnetoelectric transducer used in the earphones is shown in Figure 5-34.

There are two basic types of dynamic transducers: orthodynamic and isodynamic transducers. In an isodynamic transducer, the coil is embedded in the diaphragm in such a way that the resulting magnetic force applied to the diaphragm is equally distributed on the entire diaphragm surface. In an orthodynamic transducer, the force is applied to the diaphragm at only one point. The advantage of the isodynamic transducer is that it reproduces sound more accurately when compared to the orthodynamic transducer. However, the orthodynamic transducer is more efficient in the sense that it can produce louder sound with a given input voltage.

Figure 5-34. A frontal view of magnetoelectric (dynamic) earphone transducer (http://en.wikipedia.org/ wiki/Headphones).

Electromagnetic transducers

Electromagnetic transducers are similar to magnetoelectric transducers except that the coil is stationary and wound around an electromagnet and a metal membrane (or a dielectric membrane with an attached piece of magnetic material) is the moving element. Therefore, the electromagnetic transducers are sometimes called moving magnet transducers. In comparison to the magnetoelectric transmitters, the electromagnetic transmitters are generally smaller and more efficient but have a narrower frequency response. When they are incorporated in the insert earphones they require a good seal to the ear canal to provide reasonably wide frequency response. A view of electromagnetic earphone transducer is shown in Figure 5-35.

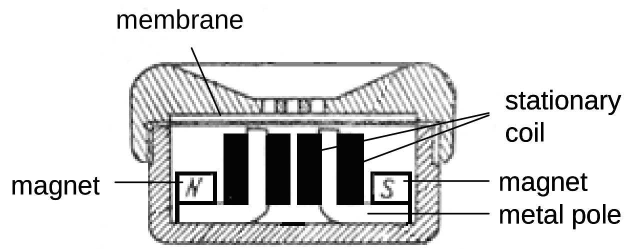


Figure 5-35. A cross-section of an electromagnetic earphone (adapted from Kacprowski, 1956).

Electromagnetic transducers are the transmitters of choice for in-the-ear devices because of their high efficiency. The most common type of electromagnetic transducer used in hearing aids and insert earphones is the magnetic balanced armature transducer in which armature is symmetrically balanced to minimize non-linear distortions of the system. The armature is ferrous material attached to the magnet and excited by an alternating magnetic field created by an audio current passing through a stationary coil surrounding the armature. The armature is attached to a plate or a membrane that vibrates and produces the sound. Typically the armature, coil, and magnetic structure are centered on the axis of the cylindrical construction, and motion of the armature in the axial direction is transmitted to the diaphragm by a pin coinciding with the axis of the cylinder. The whole assembly is supported by a shock-absorbing system. Examples of insert earphones that use balanced armature technology are the Etymotic ER6, and Westone UM1. Another type of electromagnetic transducer is the rocking armature transducer shown in Figure 5-36.

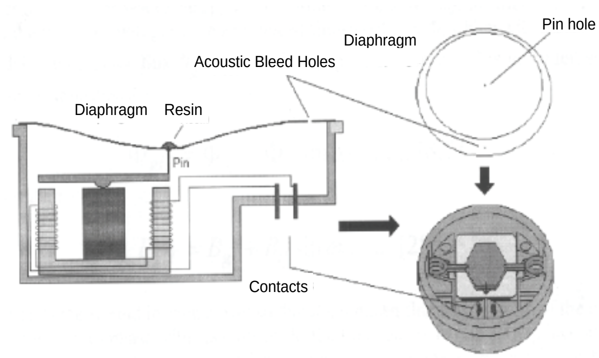


Figure 5-36. Electromagnetic transducer with rocking armature element (Moulton, 2004).

In some insert earphones the electromagnetic transmitter works together with a dynamic transducer to provide both efficiency and wide bandwidth with a wide frequency range. An example of hybrid canal earphones is the Ultimate Ear UE-10Pro that features two loaded armature electromagnetic transducers and one magnetoelectric transducer.

The construction of an electromagnetic transducer used in bone conduction vibrators from RadioEar is shown in figure 5-37. The magnet and coil assembly is attached to a lead block to increase its mass. The spring with spacers maintains the air gap, a critical separation between the permanent magnet poles and the armature. Consider the mass of the enclosure and armature as one part and the mass of the magnet, coils, and the lead block as another part of the system. Both parts of the system are connected together by the spring. When an ac signal is applied to the coil, the varying magnetic force in the air gap causes the metal armature to vibrate vertically. The mass of the magnet assembly is large enough make it appear fixed so that the armature and the enclosure move away and toward the magnet assembly creating mechanical vibrations of the transducer.

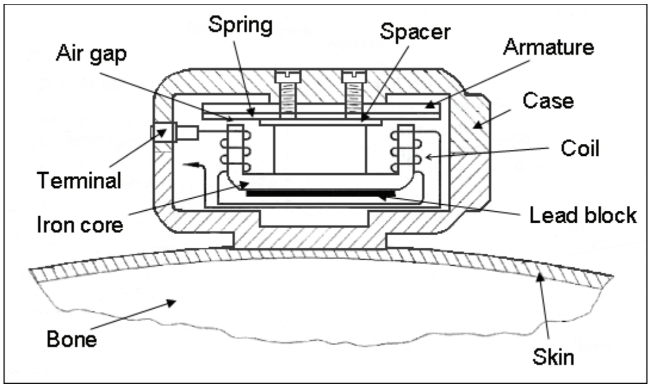


Figure 5-37. Schematic diagram of the RadioEar B-71 vibrator (Courtesy of RadioEar, Inc.).

Electrostatic transducers

A basic electrostatic (condenser) transducer consists of a thin diaphragm suspended at the center of two perforated flat metal plates. The plates and diaphragm form a capacitor. A high voltage bias is applied to the diaphragm polarizing it against both stationary electrodes. Because the suspended diaphragm is located at the center of the gap between the outer plates, the resulting attractive force to the outer plates is cancelled holding the diaphragm in a fixed position when no audio signal is applied to the electrodes. When an audio signal is applied to the outside plates, it creates an alternating electrostatic field between the plates pushing or pulling the suspended diaphragm. The movement of the diaphragm pushes air through the holes in the metal plates generating auditory signals. In some cases only one metal plate is used and the audio signal is delivered between the metal plate and the diaphragm. A schematic view of the electrostatic transducers is shown in Figure 5-38.

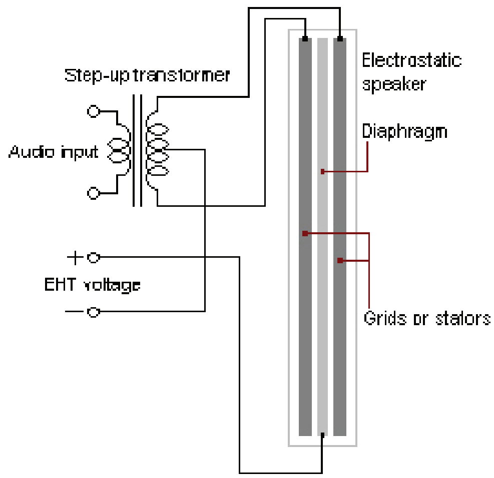


Figure 5-38. Schematic diagram of electrostatic transmitter (http://en.wikipedia.org).

Electrostatic transmitters are usually large, heavy, and expensive and require a high bias voltage. A step-up transformer is needed for the audio signal and is usually built into an adaptor box powered by commercial or generator power. Therefore they are seldom used in audio HMD systems. The primary advantages of electrostatic transducers are fast response, low distortion, and high fidelity sound reproduction. High quality electrostatic earphones include the Koss ESP950, Sennheiser HE60, Stax SR-1, and Stax 4070.

An *electret* transducer is an electrostatic device with the suspended dielectric diaphragm permanently polarized or with dielectric material filling the gap between the metal plate and the diaphragm. These transducers do nor require an external bias voltage and, thus, are much smaller, less expensive, and more rugged. However, they are very inefficient and thus only used in microphone assemblies.

Piezoelectric transducers

Piezoelectric transducers utilize the ability of crystals and some ceramic materials to generate a voltage in response to applied mechanical stress. When a voltage is applied across a piezoelectric material, the material is deformed. Conversely, if mechanical pressure is applied to the material, a potential difference is created on the opposite sides of the crystal. This unique property has many applications in electronic devices, especially in the audio industry. Because of this two-way effect, piezoelectric materials can be used as transmitters (earphone, vibrator) or receivers (microphone). An example of a piezoelectric earphone is shown in Figure 5-39.

Construction of piezoelectric transmitters is simple; they require fewer parts than magnetic and electrostatic transducers, and are very efficient. However, the frequency range of a piezoelectric transducer is limited. Therefore, in general, piezoelectric transducers are not suitable for applications where a wide, flat frequency response is required. Typical sizes of piezoelectric transducers used as microphones or buzzers are shown in Figure 5-40.

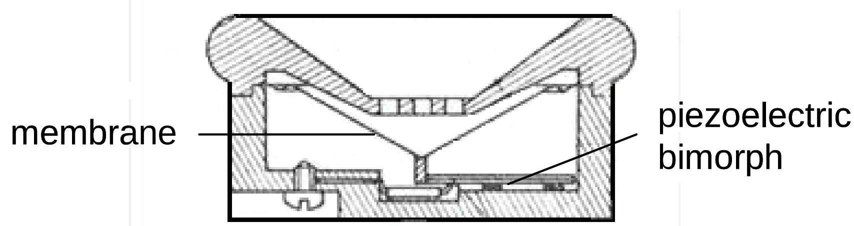


Figure 5-39. A simplifier view of a piezoelectric earphone (adapted from Kacprowski, 1956).

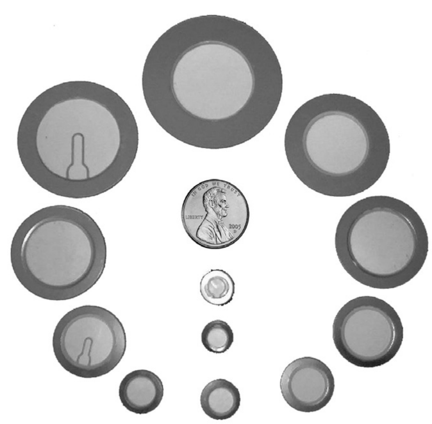


Figure 5-40. Piezoelectric transducers (Courtesy of Piezo Solutions).

Multiple-transducer designs

Audio HMD systems usually have one audio transmitter delivering a signal to the ear. However, some systems consist of two or more transducers delivering signals to the ear, built like large multi-loudspeaker systems. Such systems were mentioned above during the discussion of electromagnetic transducers.

In some high quality audio earphones there can be as many as three transducers; one for each frequency range - low-, mid-, and high-frequency. In some others, multiple transducers are used to create spatial displays. Recall that the spatial audio displays projected by proximal transmitters are normally created by means of HRTFs (discussed in Section 5-4). However, another method to create a spatial display is to use multiple transducers spatially separated slightly in an earcup and delivering multi-channel signal. For example, Konig (1996; 1997) described 4-transducer and 6-transducer earphones producing spatial sound without using HRTFs. The optimal arrangements of dynamic transducers inside the earcup for 4- and 6-channel headphones are shown in Figure 5-41. Figure 5-42 shows an actual arrangement inside the ear cup of a commercial surround sound headphone.

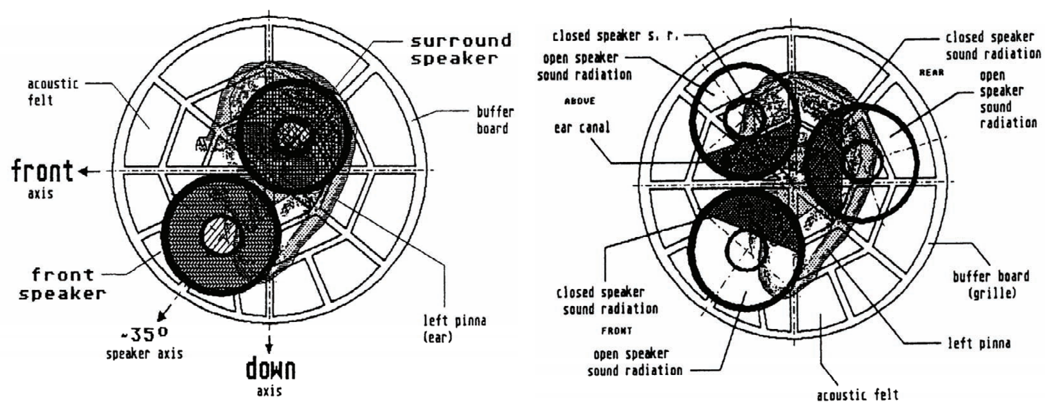


Figure 5-41. Transmitter arrangements inside left ear cup for 4-channel and 6-channel headphone (Konig, 1997).

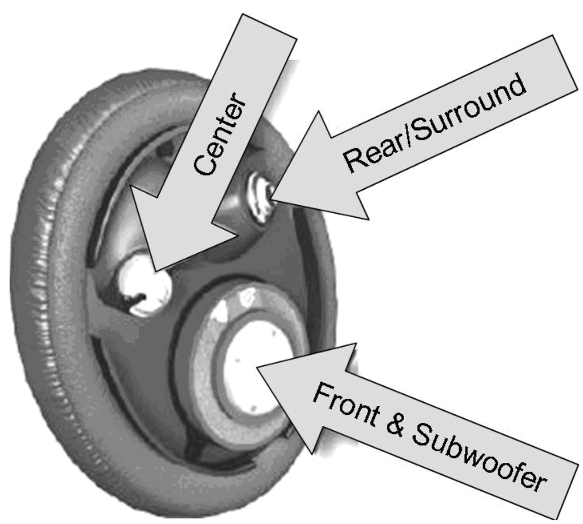


Figure 5-42. View of the earcup of LTB Magnum 5 surround audio display LTB-MG51-USB) (Courtesy of LTB Audio Systems, Inc.).

Few research papers describing this auditory spatial display technique have been published (e.g., Makowski and Letowski [1975] and Letowski and Makowski ([980]). The advantage of this technique is that there is little signal processing involved. However, today the use of HRTF in spatial displays is a more common practice due to increased microprocessor speed, reduced power consumption, and the low cost of hardware.

Audio transmitter calibration

The frequency response of an audio HMD system is typically measured as a pressure response using a standardized acoustic coupler. An acoustic coupler is an interface device that represents a standardized load to the acoustic transmitter used in the display. In the case of earphone-type audio systems it is a small chamber of specific shape and volume with an opening for coupling the audio transmitter (an earphone) to the chamber and with a measuring microphone terminating the chamber. Calibration procedure requires a specific static force pressing the transmitter against the coupler and specific environmental conditions to operate properly. In the case of bone conduction audio systems an accelerometer (motion transducer) is attached to a mechanical device providing a standardized load for the audio transmitter (a vibrator). The role of acoustic and mechanical couplers is to provide standardized and repeatable load conditions similar to the load conditions of the ear or the skull bones.

Standardized couplers provide repeatable data but such data are not necessarily a good representation of the signal delivered to the human listener. To know the actual frequency response of the audio transmitter seen by the ear or the bones of the head, the acoustic coupler used for such measurements must exactly represent the actual load provided by the human ear or the human head. The acoustic couplers that intend to represent the exact load provided by an average human ear are referred to as artificial ears or ear simulators (e.g., B&K 4153 and Larson-Davis AE100). The couplers that intend to simulate the load provided by a mastoid bone of the human head are called artificial mastoids (e.g., B&K 5090 and Larson-Davis AMC93). They have a specified range of frequencies within which such simulation can be assumed. Outside this range, the devices should be treated just as regular acoustic and mechanical couplers which do not necessarily match human characteristics.

Another method of measuring frequency response of earphone-based systems is to mount such displays on a manikin with artificial ears built in the head of the manikin. Such manikins are called artificial heads, binaural heads, or dummy heads by their developers. Examples of such heads include the B&K head and torso simulator (HATS), Aachen Head (HEAD Acoustics), and Knowles Electronic Manikin for Auditory Research (KEMAR). An artificial head provides a more natural coupling between the transmitter and the measuring system and simultaneous assessment of two transmitters (left and right) in their natural positions on the head. However the artificial ear terminates the ear canal rather then being mounted flush with the head and the collected data must be compensated for the additional travel of the acoustic wave along the canal.

Still another possibility is to put the display on a real person with miniature microphones mounted at the entrance to the ear canal. Such a real (human) load is not standardized and repeatable but provides the users with information regarding the effects of their own head on the auditory stimulus emitted by the transmitter.

The frequency response of a typical audio HMD is not flat and has several resonance and anti-resonance regions. In many applications not attempting to simulate the recording space or specific virtual environment such a frequency response is fully acceptable. However, any faithful reproduction of the recording environment requires a flat frequency response. This flat response is especially important if the signals are convoluted with a specific HRTF to create realistic immersion in the virtual environment.

There are three basic methods of earphone calibration/equalization: pressure equalization, free-field equalization, and diffuse field equalization. Each of these methods attempts to flatten the frequency response of an earphone with respect to a specific reference point. Pressure equalization flattens the frequency response in reference to the sound pressure measured using an artificial ear (acoustic coupler) or artificial head methods. Free-field equalization intends to recreate the conditions of sound field listening in which the listener is in front of a sound source in a non-reflective environment. Diffuse-field equalization assumes the sound source is not necessarily in front of the listener and the sound arrives at the ear with the same intensity and the same probability from any direction (Killion, Berger, and Neuss, 1987; Thiele, 1986; Larcher, Jot, and Vandernoot, 1998). Diffuse-field equalization is the most appropriate for simulating listening to distant sound sources in an enclosed space or in a free field environment when the sound sources surround the listener. Diffuse-field equalized earphones provide better spatial impression of the sound and make it easier to differentiate between sounds coming from the front and back of the listener. The compensated frequency response for diffuse-field listening is called a diffuse-field frequency response. Diffuse-field equalization is commonly built into high-quality earphones intended for music listening or virtual reality listening. Such products are provided by AKG (e.g., K 240D), Etymotic (e.g., ER 4P), and Sennheiser (e.g., HD 250, HD 580, HD 600, HD 650), Stax (e.g., Lambda Pro), and other manufacturers. An example of the relation between the pressure response and the diffuse-field response for the Telephonics TDH-39 earphones is shown in Figure 5-43.

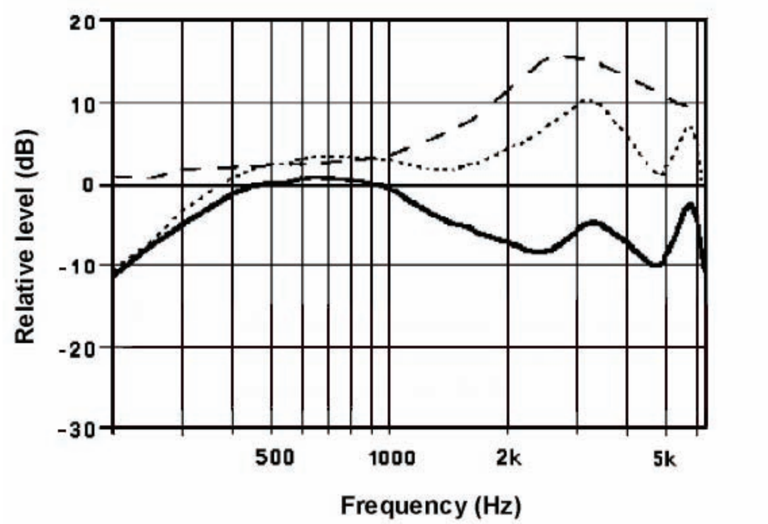


Figure 5-43. Relation between the pressure response (dotted line) and diffuse-field response (solid line) of the TDH-39 earphones. The dashed line shows the transformation function between pressure and diffuse-field environments (Cox, 1986; Killion, Berger and Nuss, 1986).

Audio receivers

Audio HMD systems are designed to provide information to the user via the auditory path, but the technologies used to generate acoustic stimuli may also function as collectors of mechanical (acoustic) energy and can be used to convert this energy into electric signals. The transducers that convert acoustic signals into electric signals are called audio receivers or microphones. Microphones used in conjunction with audio HMD systems are typically dynamic (moving coil microphones or condenser (electret) microphones. They can be used as air conduction microphones, bone conduction microphones, or throat microphones, which are mounted on the neck and receive signals directly from vibration of the vocal folds. Air conduction microphones are typically the noise-canceling type designed in such a way that the unwanted ambient noise is presented to two out-of-phase microphone elements while the desired speech communications is presented to only one microphone element. Using this technique, the unwanted noise may be reduced through phase cancellation. Bone conduction or throat microphones are much less susceptible to air conducted noise energy than air microphones and provide a good signal-to-noise ratio without additional signal processing.

Audio receivers should be designed and selected for specific applications. Frequency response, sensitivity, impedance, etc. of a microphone must be matched to the equipment to which it is attached and to the input conditions in which it operates. In addition, environmental conditions in which the microphone operates must be considered and they include dust, shock, vibration, rain, salt spray, temperature, and humidity extremes.

System interfaces and connections

The input to audio HMD systems comes from microphones, computers, intercom systems and/or radio communication systems through wired or wireless connections. To operate properly within the required communication regime, the audio system wiring and input circuitry should be compatible with technical specifications of the whole communication system. These requirements include type of connectors and pin assignments, signal level and impedance matching, and common ground requirements. Switching from the send to receive mode of operation and vice versa should be accomplished by a voice operating switch or by a push-to-talk (PTT) switch which is easily accessible and sufficiently large to be operated without removing gloves. All cables and fixed connections shall withstand 20 lbs tension to operate securely and reliably (PM Soldier Warrior, 2007).

For wired connections, audio HMD systems are normally connected to sound sources by using plugs and jacks to facilitate easy detachability. Different sizes of plugs are available to mate with different form factors of jacks. A mini-plug known as the 1/8 inch (3.5 mm) plug is the most common for portable devices; a smaller plug (2.5 mm) is common for cellular phones; and the full size 1/4 inch phone plug is often used in professional audio or laboratory applications. The universal serial bus (USB) connector is another new type of audio connector used to interface digital audio signals to/from personal computers or game consoles. As discussed in the initial part of this chapter, there are many methods to display audio signals to the listener, so these plugs can be mono or stereo plugs (2 or 3 electric contacts), or a connector with multiple pins to accommodate combinations of different signals as the application may require.

Cables connecting the audio source to the audio HMD system also contribute to the quality of the reproduced sounds. Unshielded cables and connectors are susceptible to electric interference from other sources. High conductivity cable provides improved signal transmission and results in less signal distortion. In some high quality sound systems, optical fiber is used for optimal signal transmission. When fiber optic cable is used the signal must be converted from a light signal to a mechanical signal using electronic converters located within the HMD system. For military applications, the standard electric interface to radios and intercommunications systems is the U-329/U connector shown in Figure 5-44. This connector provides single-channel audio to the headset (handset). Since military radios provide monaural audio, the military has not yet adopted a standard multi-channel connector configuration. Typically the connector is wired as follows:

- Pin A   Common Ground
- Pin B   Transmitter
- Pin C   Push-to-Talk Switch
- Pin D   Receiver (microphone)
- Pin E   DC Power (not standardized)

Figure 5-44. U-329/U Audio accessory connector used on U.S. military radios. This 6 pin connector is common to many military radios including the AN/VRC-111 and SINCGARS (Courtesy of Tactical Engineering).

Sound signals can be also delivered to audio HMD systems through wireless networks. Although cordless telephones have been available for many years, wireless audio HMD systems are only recently becoming common. The developments of digital radio frequency communication and of low cost transceiver microprocessors make wireless systems more attractive and affordable although they produce an electro-magnetic signature that is not desirable in some cases.

The most popular communication protocol used with wireless audio HMD systems is Bluetooth, also known as the Institute of Electrical and Electronic Engineers (IEEE) standard 802.15.1. It allows two devices to communicate with each other via unlicensed short range radio frequency (RF) signals. Bluetooth, developed in 1994 at Ericsson Radio Systems, Netherlands, was designed for low power short range communication (Institute of Electrical and Electronic Engineers, 2005; McDermott-Well, 2005). With Bluetooth technology, the audio HMD can receive signals at a maximum range of 1 m, 10m or 100m, depending upon the RF power of the system. A picture of an audio display utilizing Bluetooth technology is shown in Figure 5-45. For greater ranges, Wireless Fidelity -WiFi - a spread-spectrum system operating on several channels in the 2.4 GHz band (also known as IEEE 802.11, ANSI/IEEE Standard 802.11, 1999 edition (R2003)) is used. Other wireless signal transmission methods used with audio HMD systems are analog radio frequencies (very high frequency [VHF] or ultra high frequency [UHF]) or infrared light (Moy, 2001a).

Figure 5-45. Bluetooth mobile phone headset (http://en.wikipedia.org/wiki/Bluetooth).

Although wireless networking provides great convenience (no tangled wire, no tether), each wireless technology has its own advantages and limitations. In general, the wireless network is susceptible to interference, introduces noise, and drops connections occasionally. Infrared technology uses infrared light to transmit audio signals, thus requiring line of sight to the base system. Radio frequency signals can transmit through walls and often interfere with other surrounding radio frequency systems.

Mounting and hearing protection considerations

Mounting of the audio HMD on the head is very important consideration because it affects both effectiveness of the interface and user’s comfort. In considering a mounting solution for an audio HMD three basic factors need to be taken into account: technical quality of coupling the audio HMD to the user, user’s comfort, and system durability.

In order to provide proper interface to the ears or the head of the user an audio HMD can be built in the headgear, worn with a headband or other supporting structure, or inserted in the ear canal. There are several types of headbands used with different audio HMD systems that fit over the head, behind the head, behind the neck, or under the chin. The over-the-head mounting technique supports the weight of heavy ear cups and provides stability of the headphones on the head. The headband can be a hard bow conforming to the shape of the head or a soft harness around the head with optional top head support. The former type of headband is commonly used with circumaural and supra-aural audio systems whereas the latter one is used with bone conduction systems and lighter earphone-type audio HMD systems such as ear buds, and insert earphones.

In addition to soft harness design another headband style that is especially designed to be used with headgear is the behind-the-neck headband. The behind-the-neck headband is curved under the hair line along the neck and up around the ears. The headband hangs relatively loose behind the neck providing spring-like action holding audio transducers in place. This design interferes less with hair or helmet than over-the head or behind the head designs and is usually used with lighter audio display systems as supra-aural or ear bud systems.

Due to the variety of head sizes, the length of a headband needs to be adjustable for comfort and fit. The width of a headband also contributes to its pressure on the head. Padding with cushions provides better fit and makes long-term wear more comfortable. Some headbands are also foldable for better storage and portability. The over-the-head headband is probably the most reliable of all the on-the-head mounting system; however, it can interfere with other headgear.

Another mounting method is the clip-on-the-ear design, which includes a clip attached to the miniature audio transmitter by a hinge in such a way that it can be easily opened and closed to remove or keep it in place. This design is usually seen in light and inexpensive audio display systems using ear buds, supra-aural earphones, or insert earphones. The limitation of the clip-on-the-ear design is lack of comfort during the prolonged use due to the pressure of the clip on the ear.

Some small and lightweight audio display systems such as ear buds or insert earphones can be worn without any additional support by fitting snugly on the conchae or inside the ear canals. However, with this mounting technique, the earphones have a tendency to be pulled out of the ears by the weight of the connecting cables, especially during physical activities that require a variety of movements. Examples of mounting techniques used with helmet independent (add-on) audio display systems are shown in Figure 5-46.

Figure 5-46. Examples of various light-weight headsets with over-the-head (left), behind-the-neck (center), and clip-on-the-ear (right) mounting (http://www.amazon.com; http://www.thanko.jp; http://www.boscovs.com).

For military and firefighter applications it is desirable to integrate an audio HMD into the helmet. Transducers can be integrated into the impact resistant shell or embedded into the padding of the helmet support system. This mounting method gives users the convenience of fewer cables and only one piece of equipment to care for. For example, the Bose CVC helmet has its audio display and communication system integrated into the helmet shell. This helmet is shown in Figure 5-47 (left panel). Similarly, the Gentex aviation helmet HGU-56/P is equipped with audio HMD in a form of a pair of earphones mounted under the helmet. The CEP and CEPS in-the-ear systems can also be successfully used with the HGU-56/P and SPH-4B aviation helmets. These implementations are shown in Figure 5-47 (middle and right panels).

Audio displays integrated in the helmet are state-of-the art solutions for the aviators, tankers, and other users that need protection from noise and reliable communication within a moving platform. However, when the audio HMD system used for communication purposes is integrated into the helmet, the Warfighter does not have radio communication capability without donning the helmet. This creates the need for modular audio HMD systems for dismounted Warfighters, firefighters, security personnel, and others who may or may not wear helmets. Such systems employ C&HPS. The C&HPS may be worn using one of the mounting techniques described above or can be embedded into a fabric cup or harness worn under the helmet.

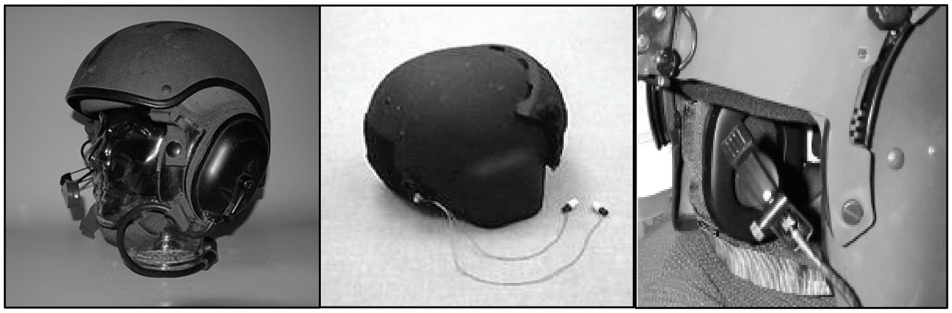


Figure 5-47. CVC helmet (left panel), HGU-56/P with CEP (middle panel), and HGU-56/P with CEPS (right panel) (Courtesy of USAARL).

In the case of the bone conduction audio HMD systems the selection of an appropriate mounting technique is challenging due to contradictory requirements of a minimum static force on the contact area needed for comfortable use of the display and the need for some minimum static pressure to provide good contact to the head and efficient sound transmission. These requirements favor large low-profile curved transmitters or a distributed network of miniature transmitters. In commercial bone conduction headsets, vibrators are secured with over-the-head, around-the-forehead, behind-the-head or behind-the-neck headbands (e.g., Sensory Devices and Vonia systems), or are incorporated into a web cap as in the case of the Temco HG-17 headset. Temco also produces an integrated bone conduction audio display and communication system intended to be mounted on a gas mask (Temco FM-1) or attached by an adhesive to the skin over the temporal bone behind the ear (Temco SK-1). For military Special Forces, security personnel, police, or intelligence agents the bone conduction transmitters can be secured on the head under the hair or mounted in inconspicuous head covers such as a baseball cap or hat.

Similar to air conduction transducers, over-the-head headbands are found in most commercial bone conduction headsets with the vibrators pressed securely to the face bones. Typically, the headband of a bone conduction audio HMD is stiff and flexible enough to maintain adequate static force on the vibrators. However, when worn with a helmet, the pressure of the helmet on the stiff headband can cause the vibrators to lose contact with the skull. Therefore such modular use requires soft harness mounting rather than hard headband mounting of the transmitters. Figure 5-48 shows typical mounting techniques used in commercial bone conduction audio HMD systems.

Figure 5-48. Examples of mounting techniques used with Temco bone conduction audio systems. The pictures show the over-the-head (left panel), behind-the-neck (center panel), and on-the-gas-mask (right panel) systems (Courtesy of Temco Communications, Inc.).

User's comfort is the most critical element of mounting considerations for audio HMD systems. Sound quality is usually considered the most important factor of audio HMD systems with comfort usually considered as a secondary requirement when the systems are used for short periods of time, that is, only when needed. However, comfort may actually be equally important as sound quality when an audio display system must be worn for long periods of time. Long-term discomfort results not only from an uncomfortable fit of the audio HMD but also from the degree of psychological isolation caused by the headgear and fatigue caused by system unbalance, weight, and a large number of controls that need to be operated when the system is used. A user may mildly complain about an audio HMD system with less than optimum sound quality but that same user will typically refuse to wear uncomfortable equipment for long periods of time – tens of hours, not minutes. Many factors (weight, compatibility with other equipment, mounting technique, fit) contribute to quality and comfort. If an audio HMD system is uncomfortable, it will not be used regardless of how well it performs and protects.

Earmuff-type HMD systems are typically built in the CVC and aviator's helmets providing hearing protection and housing audio communication transmitters. They perform a significant role in providing stability of the helmet on the user's head and overall comfort of the helmet (Mozo, 2001). They also isolate the ears from potential contact with the helmet liner, which increases the overall comfort of the helmet. Their main drawback in CVC and aviator's helmet applications is that they do not provide good ballistic and lateral impact protection (Shanahan, 1985). However, there are some design considerations (e.g., lower weight, modified structural strength) that may increase the lateral impact protection of earmuff-type HMD systems (Mozo, 2001).

The amount of hearing protection needed for the audio HMD system is a function of frequency and depends on the type of application and specific use of the system (with or without the helmet). Typical earmuff and earplug protectors provide some limited protection at low frequencies and the amount of protection increases with frequency. The minimum noise attenuation values by hearing protection tactical headsets recommended in a recent draft of the U.S. Army document are listed in Table 5-7.

The values listed in Table 5-7 reflect attenuation curves of typical HPDs. However, this curve is just the opposite to what may be required in most continuous vehicle, industrial, and environmental noises that typically have energy density distributions inversely proportional to frequency (1/f). This means that they are predominantly low frequency noises. In addition, good speech recognition requires good audibility of speech energy from 1000 Hz to 4000 Hz, which is usually significantly attenuated by most hearing protectors. Thus, in the applications that require live speech communication it makes sense to use audio HMD systems that offer noise attenuation that do not increase much with frequency. This philosophy is reflected, for example in the US. Air Force document MIL-PRF-89819/4 (DAF, 1997b) that specifies minimum attenuation for the in-flight headset-microphone to be approximately 20 dB for any frequency above 800 Hz and gradually lower attenuation with decreasing frequency due to the technical reasons.

It must be stressed that comfort and weight of an HMD (audio or otherwise) system are inversely related. In general, heavy HMD systems place more strain on the user's neck during movement or prolonged activity in any position. Thus, from a purely weight consideration, any lightweight in-the-ear or bone conduction devices may be favored over heavier and more bulky earmuff-type systems assuming they provide the same amount of hearing protection, speech intelligibility and situation awareness. However, the discomfort of having a device inserted in the ear canal or pressing against the head may be greater than the discomfort caused by the earmuff-type solutions. This stresses the need for considering the comfort of in-the-ear and bone conduction devices as a priority issue in designing such systems. Long-term comfort must be a primary consideration when designing any audio HMD, whether in-the-ear, bone conduction, or conventional systems. This calls for comfort evaluation of the devices by the users and response instruments (questionnaires, scales) that can provide a thorough feedback to the designers. An example of a scale that is used for comfort rating is the Wong-Baker FACES pain scale (Wong and Baker, 1988). This is a six step scale (from 0 to 5) illustrated by six faces expressing gradually increasing degree of pain from happy (no pain) to very unhappy (a lot of pain). This scale adapted for comfort rating is shown in Figure 5-49.

Table 5-7.
Noise attenuation criteria for hearing protection tactical headset.
(Modified from PM Soldier Warrior [2007]).

Octave band attenuation requirements [1] (dB)	Frequency (Hz)						
	125	250	500	1000	2000	4000	8000
Moderate noise level exposure [2]	15	16	21	27	28	35	35
High noise level exposure [3]	28	34	34	34	34	40	40

[1] *Levels should be determined following ANSI S12.42-2002 standard (ANSI, 2002) measurement procedure and subtracting 1 standard deviation*

[2] *Infantry Soldiers in wheeled tactical vehicles*

[3] *Infantry Soldiers in tracked tactical vehicles*

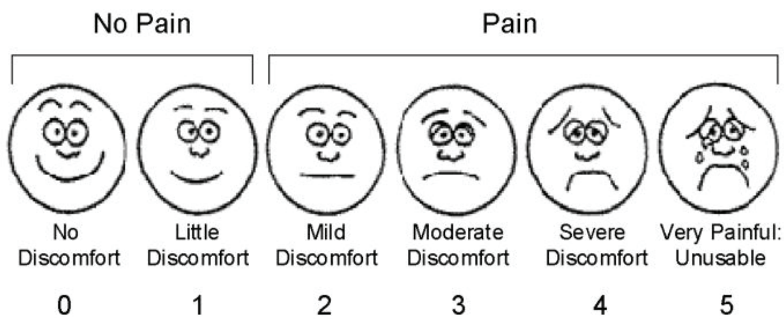


Figure 5-49. The Wong-Baker FACES pain scale adopted for comfort rating (Modified from: Hockenberry M, Wilson D, and Winkelstein ML: Wong's Essentials of Pediatric Nursing, ed. 7, St. Louis, 2005, p. 1259 (Copyright, Mosby) (used with permission).

The third important factor that needs to be considered in designing or selection of audio HMD systems is the durability of the system. Most COTS audio systems are not suitable for the harsh environments of military or firefighter operations. Audio HMD systems for the military must be sustainable in high impact, high temperature, and dusty environments. In some cases waterproof devices are required. For military applications under combat conditions, equipment should meet the requirements of MIL-STD-810F (DOD, 2001). This standard requires materiel to meet certain environmental design criteria and specifies tests and methods which replicate field conditions to verify compliance. This standard addresses and specifies minimum performance requirements for the following categories of environmental conditions: low pressure, high temperature, low temperature, temperature shock, contamination by fluids, solar radiation (sunshine), rain , humidity, fungus, salt, fog, sand dust, explosive atmosphere, immersion, acceleration, vibration, acoustic noise, shock, pyroshock, acidic atmosphere, gunfire vibration, temperature, humidity, vibration, and altitude , icing/freezing rain, ballistic shock, and vibro-acoustic and temperature conditions (DOD, 2001).

Speech intelligibility

Audio HMD systems are required to provide audio signals that result in auditory stimuli that are heard, recognized, and localized by the listener. The primary stimuli to consider are speech stimuli and its intelligibility. Speech intelligibility is defined as the percentage of speech units that can be correctly identified by an ideal listener over a given communication system in a given acoustic environment. If the properties of the listener, such as hearing loss or divided attention, are taken into consideration, it is more appropriate to refer to speech recognition rather than speech intelligibility.

Poor speech intelligibility increases task difficulty, compromises human performance, and may lead to loss of life (Peters and Garinther, 1990). The criteria for minimum required speech intelligibility in voice communication systems are stated in MIL-STD-1472F (DOD, 1999; Table VI) and are listed in Table 5-8.

The Modified Rhyme Test (MRT) criterion scores listed in Table 5-8 are the adjusted for guessing word recognition scores for the six-alternative MRT (House et al., 1965). The MRT is one of the three speech tests recommended for testing speech intelligibility in communication systems (ANSI, 1989).

The values listed in Table 5-8 are desirable goals and criteria for fielding live voice communication equipment and for natural person-to-person communication. However, the referenced standard does not specify the test conditions leaving some room for interpretation. More specific test conditions are included in the Communication Clarity Criteria being developed by the Program Manager Soldier Warrior office (PM Soldier Warrior, 2007) and shown in Table 5-9. Specified criteria that need to be met require performing the MRT as described in ANSI S3.2-1989 standard (ANSI, 1989), the talker to be in ambient noise environment of 75 dB SPL or more, and the listener in a pink noise environment with the overall sound pressure level as specified in Table 5-9. These criteria are based on the U.S. Air Force criteria for headset-microphone (DAF, 1997a). The difference between these two documents is that the Air Force document requires 85 % speech intelligibility at 105 dB SPL and 80% intelligibility at 115 dB SPL.

Table 5-8.
Intelligibility criteria for voice communication systems.

Communication Requirement	MRT* Score
Exceptionally high intelligibility; separate syllables understood	97 %
Normal acceptable intelligibility; approximately 98% of sentences correctly heard; single digits understood	91%
Minimally acceptable intelligibility; limited standardized phrases understood; approximately 90% sentences correctly heard (not acceptable for operational equipment)	75%

* Modified Rhyme Test

Table 5-9.
Communication clarity criteria for hearing protection tactical headsets.
(Program Manager - PM Soldier Warrior, 2007)

Sound pressure level of pink noise (dB SPL)	75	95	105
Minimum score percent correct (adjusted for guessing) in %	90	85	80

Relations between speech intelligibility scores (in %) and speech level (in dB A) for the communication system of the SPH-4B aviation helmet operating without and with the addition of CEP or Bose ANR system are shown in Figure 5-50. The data were obtained in UH-60 helicopter cabin noise of approximately 110 dB produced during flight at a forward speed of 120 knots (Mozo, 2001; Mozo and Murphy, 1997).

In practical situations, when worn equipment is being used in adverse listening conditions, the actual speech intelligibility is much worse than was expected. Therefore, it is critical to test speech intelligibility under the worst expected operation conditions (the worst case scenario) as well as under normal operational conditions. In addition, worn equipment does not perform as well as new equipment and should be tested periodically.

One important consideration in selecting an audio HMD system is the bandwidth of the radio communications channel that will be used to provide the data. Typical telecommunications and military radio systems are frequency band-limited to a pass-band from approximately 300 Hz to 3.4 kHz. With the introduction of digital telephony, based on the International Telecommunication Union standard G.711 (the standard for encoding telephone audio on a 64 kbps channel), the upper frequency limit of the telephone network is now commonly accepted to be approximately 3.3 kHz at best. The last Bell public switched telephone network tests in 1984 showed significant high-frequency roll-off at 3.2 kHz for short and medium distance connections, dropping to 2.7 kHz in long distance connections. The telephone network carries frequencies no lower than 220Hz, and most commonly the lower limit is 280 or 300 Hz. (Rodman, 2003). For this application an audio HMD with frequency response tailored to human voice data transmitted over a bandwidth limited channel will outperform an audio HMD with a frequency response which covers the full range of human hearing due to both noise and power constraints. Conversely, for applications where sound source localization cues are required and sufficient bandwidth is available to present needed high-frequency sound energy, the auditory performance of the listener would be hampered by a frequency-limited HMD (i.e., an HMD lacking a flat frequency response over the entire range of frequencies perceived by a human).

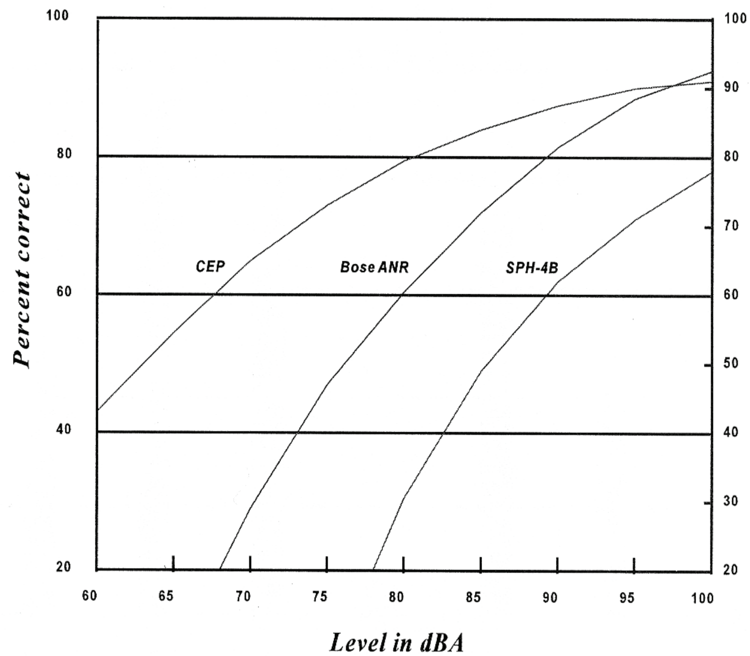


Figure 5-50. Radio communication speech intelligibility scores for SPH-4B aviation helmet without (SPH-4B) and with CEP (CAP) or Bose ANR (Bose ANR) systems as a function of speech level in UH-60 helicopter noise (Mozo, 2001) (Courtesy of USAARL).

Other primary factors affecting speech intelligibility are poor speech articulation by the talker and loss of signal intensity during speech transmission. The sound attenuation provided by the aircrew and tanker helmets and by other ear-encapsulating headgear greatly affects intelligibility of natural live speech. For example, Garinther and Hodge (1987) observed that the presence of the M25 respiratory mask and the NBC (nuclear-biological-chemical) protective hood restricted effective speech communication range to less than 12 meters. Conversely, the typical infantry helmets provide only minimal speech attenuation in the frequency range below 4.0 kHz (Randall and Holland, 1972), that is within the range that is responsible for providing more than 80% of speech intelligibility (ANSI, 1997b). The potential detrimental effect of an infantry helmet on speech communication is in providing false cues regarding the direction of incoming speech.

Audio and radio communication systems that provide good speech intelligibility have been reported to improve combat performance and decrease Warfighter's fatigue. Garinther and colleagues (Garinther, Whitaker and Peters 1994; Whitaker, Peters and Garinther 1989) reported that a specified percent of improvement in speech intelligibility provides an almost equal improvement in crew performance. The functional relationship between mission success and speech intelligibility is shown graphically in Figure 5-51.

Audio HMD Systems: Closing Remarks

The audio HMD system is a sub-system of a larger multi-functional display system of the helmet assembly. It is a challenging sub-system since it must provide hearing protection and auditory situation awareness in addition to an audio display. In addition, to operate properly as a part of the large system, the audio HMD must be physically and electrically compatible with the remainder of the HMD system and not interfere with the functioning of the non-audio system components of the HMD. This requires appropriate design considerations so as to provide an engineered solution suitable for the desired application. Issues such as comfort, power requirements, size, weight, location, desired sound pressure level, fidelity (bandwidth and dynamic range), number of audio channels, wiring, and connectivity must be considered together and from the perspective of optimizing auditory performance of the user awhile considering total system requirements and functionality. For example, high power and wide bandwidth audio signals may require larger and heavier transmitters that may not be feasible to be incorporated in the overall design.

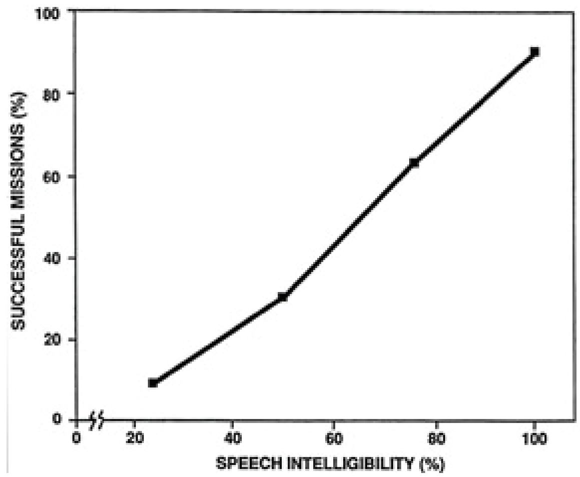


Figure 5-51. Percent of mission success as a function of speech intelligibility (Garinther, Whitaker and Peters, 1994).

Human hearing range extends from approximately 0 dB SPL to 120 dB SPL; therefore high quality output of an audio HMD may theoretically require a 120 dB dynamic range. This would be, for example, the ideal transmission range for high fidelity symphonic orchestra listening in ideal listening conditions. However, in most applications this wide intensity range is not necessary and may be dangerous. Prolonged listening to sounds (signal and/or noise) with intensity exceeding 85 dB SPL can be a source of hearing loss. For listening in quiet to normal verbal messages, the dynamic range of speech communications can be drastically limited since the effective dynamic range of speech is only approximately 50 dB. In the audio HMD systems operating in varying environmental conditions or used for environmental listening this range needs to be extended to accommodate various voice intensities from whisper to shouting and must allow hearing faint environmental sounds. Note that limited dynamic and frequency range of the transmitting channel also removes some contextual and environmental information and adversely affects transmission of emotions and physical state of the talker. However, for environmental listening it is also necessary to have an intensity limiter built into the system to protect the listener from dangerous high intensity sounds.

To protect the user from the harmful effects of high level environmental and military noises, the amount of hearing protection provided to the user by the audio HMD system must be carefully considered and integrated into the overall design from the beginning; it cannot be added as an afterthought. Natural speech communication and auditory awareness of the environment must be considered in parallel with the hearing protection system. Overprotection is actually worse than under protection since the user most-likely will defeat the protection or fail to use the system as it was intended. As discussed previously in this chapter, hearing protection can be provided in two primary forms, in-the-ear, or over-the-ear (circumaural earmuff). Adding in-the-ear hearing protection will reduce the efficiency of any external earphone-based audio systems and may make certain types of audio HMD systems unusable. Conversely, in-the-ear protection systems work well with bone conduction audio HMD systems. Circumaural hearing protection may be acceptable when used with both in-the-ear and bone conduction systems, but maximum efficiency is achieved when the audio HMD system and hearing protection are implemented as one fully integrated C&HPS. Circumaural audio HPDs are difficult to integrate into the overall helmet system because there is limited available space under the helmet in the vicinity of the ears. The decision as to which approach to take when designing audio HMDs - earmuff- or insert-type earphone, linear or non-linear, active or passive or active hearing protection, must be dictated by the mission which must be accomplished. There is no one-size fits all solution. In summary, audio HMD systems selected for a specific platform and operations must be tailored to their intended use, both operationally and environmentally. In addition, regardless of operational requirements of the system, it has to provide a long-term comfort for the user. An uncomfortable system will never be worn properly and used all the time when needed, thereby affecting users' mission effectiveness and safety.

References

American National Standards Institute. (1986). Methods for measurement of impulse noise. ANSI S12.7-1986 (R2006). New York, NY: American National Standards Institute.

American National Standards Institute. (1989). Method for measuring the intelligibility of speech over communication systems. ANSI S3.2-1989. New York, NY: American National Standards Institute.

American National Standards Institute. (1994). Acoustical terminology. ANSI S1.1-1994. New York, NY: American National Standards Institute.

American National Standards Institute. (1995a). Microphone-in-the-real-ear and acoustic test fixture methods for the measurement of insertion loss of circumaural hearing protection devices. ANSI S12.42-1995 (R2004). New York, NY: American National Standards Institute.

American National Standards Institute. (1995b). Bio-acoustical terminology. ANSI S3.20-1995 (R2003). New York, NY: American National Standards Institute.

American National Standards Institute. (1997a). Method for the measurement of real ear attenuation of hearing protectors. ANSI S12.6-1997 (R2002). New York: American National Standards Institute.

American National Standards Institute. (1997b). Methods for calculation of the Speech Intelligibility Index. ANSI S3.5-1997. New York, NY: American National Standards Institute.

Békésy, G., von. (1939). Über die piezoelektrische Messung der absoluten Hörschwelle bei Knochenleitung, *Akustiche Zeitschrifts*, 4, 113-125.

Békésy, G. von. (1960). *Experiments in Hearing*. New York: McGraw-Hill.

Bolia, R.S., D'Angelo, W.R., and McKinley, R.L. (1999). Aurally aided visual search in three-dimensional space, *Human Factors*, 41(4), 664-669.

Chapin, W., Jost, A., Cook, B., Suruku, F., Foster, S., Zurek, P., Desloge, J., Beaudoin, R., Bolas, M., McDowall, I., Lorinmer, E., Shinn-Cunningham, B., and Durlach, N. (2003). Concept and technology exploration for transparent hearing. Project Final Report. Scorpion Audio Team. Natick, MA: U.S. Army Natick Soldier Research, Development and Engineering Center.

Christian, E. (2000). The detection of warning signals while wearing active noise reduction and passive hearing protection devices. Master's Thesis. Virginia Polytechnic Institute (Industrial and Systems Engineering).

Cox, R.M. (1986). NBS-9A Coupler-to-eardrum transformations: TDH-39 and TDH-49 earphones, *Journal of the Acoustical Society of America*, 79(1), 120-123.

Dancer, A., and Hamery, P. (1998). Nonlinear hearig protection devices, *Proceedings of the 27th Annual National Hearing Conservation Association (NHCA) Conference*. Albuquerque, NM: National Hearing Conservation Association.

Department of the Air Force. (1997a). Headset-Microphone, (Hearing protective type, high and moderate ambient noise Levels); General specification for Revision: A (dated: 30 January 1997). MIL-PRF-87819A (USAF). Washington, DC: Department of the Air Force.

Department of the Air Force. (1997b). Headset-Microphone (Hearing protective type, high and moderate ambient noise levels), up to 105 dB, in-flight aircrew headset-microphone); General specification for Revision: A (dated: 30 January 1997). MIL-PRF-87819/4-01 (USAF). Washington, DC: Department of the Air Force.

Department of the Army (DA). (2000). *Aeromedical Training for Flight Personnel*. Chapter 7: Noise and Vibration in Army Aircraft. Field Manual FM 3-04.301. Washington, DC: Department of the Army.

Department of the Army (DA). (1998). *Hearing Conservation.* DA Pamphlet 40-501 (DA PAM). Washington, DC: Department of the Army.

Department of Defense. (1991). Hearing conservation program. Instruction 6055.12. Washington, DC: Department of Defense.

Department of Defense. (1997). Noise limits. MIL-STD 1474D. Washington, DC: Department of Defense.

Department of Defense. (1999). Human engineering. MIL-STD-1472F. Washington, DC: Department of Defense.

Department of Defense. (2001). Test method standard for environmental engineering considerations and laboratory tests. MIL-STD 810F. Washington, DC: Department of Defense.

Garinther, G.R., and Hodge, D.C. (1987). Effect of the M25 protective mask and hood on speech intelligibility and voice level. HEL Technical Memorandum HEL-TM-19-87. APG, MD: Human Engineering Laboratory.

Garinther, G.R., Whitaker, L.A., and Peters, L.J. (1994). The effects of speech intelligibility on military performance, *Proceedings of Symposium on Speech Intelligibility*, 1-12. Washington, DC: National Academy of Science.

Goodhill, V., and Holcomb, A. L. (1955). Cochlear potentials in the evaluation of bone conduction, *Annals of Otology, Rhinology, and Laryngology*, 64(4), 1213-1233.

Gowers, D.W., and Casali, JG (1994). Speech intelligibility and protective effectiveness of selected noise reduction and conventional communication headsets, *Human Factors*, 36 (2), 359-369.

Harris, J., Haines, H., and Myers, C. (1953). A helmet-hold bone conduction vibrator, *Laryngoscope*, 63, 998-1007.

Hartman, W.M., and Wittenberg, A. (1996). On the externalization of sound images, *Journal of the Acoustical Society of America*, 99, 3678.

Henry, P., and Letowski, T. (2007). Bone conduction: Anatomy, physiology and communication, ARL Technical Report ARL-TR-4138. APG, MD: U.S. Army Research Laboratory.

Henry, P., and Mermagen, T. (2004). Bone conduction communication in a military vehicle, *Proceedings of the NATO Vehicle Habitability Conference*. Prague, Czech Republic: NATO, 14-1 to 14-10.

Hockenberry, M., Wilson, D., and Winkelstein, M.L. (2005). *Wong's Essentials of Pediatric Nursing*, 7th Ed, 1259. St. Louis, MO: Mosby.

House, A.S., Williams, C.E., Hecker, M.H., and Kryter, K.D. (1965). Articulation-testing methods: Consonantal differentiation with a closed-response set, *Journal of the Acoustical Society of America*, 37, 158-166.

Institute of Electrical and Electronic Engineers (IEEE). (2005). Information technology-Telecommunications and information exchange between systems- Local and metropolitan area networks- Specific requirements. IEEE Standard 802.15.1-2005. New York: IEEE.

Isaacs, A. (1996). *A Dictionary of Physics*. New York: Oxford University Press.

Kacprowski, J. (1956). *Zarys Elektroakustyki*. Warszawa, Poland: Wydawnictwa Telekomunikacyjne.

Khanna, S., Tonndorf, J., and Queller, J. (1976). Mechanical parameters of hearing by bone conduction, *Journal of the Acoustical Society of America*, 60, 139-154.

Killion, M.C., Berger, E.H., and Nuss, R.A. (1987). Diffuse field response of the ear, *Journal of the Acoustical Society of America*, 81(1 Suppl), S75.

Knudsen, V.O., and Jones, I.H. (1931). Bone conduction, *Archives of Otolaryngology*, 31, 489-505.

Konig, F.M. (1997). A new supra-aural dynamic headphone system for in-front localization and surround reproduction of sound. *The 102nd Convention of the Auditory Engineering Society. Preprint 4495*. Munich, Germany: AES.

Lancaster, J.A., and Casali, J.G. (2004). REAT and MIRE insertion loss comparison for eight headphones of various passive and ANR designs, inclusive of NRR and spectral attenuation. Audio Lab Report 9/07/04-2-HP. Blacksburg, VA: Virginia Polytechnic Institute and State University.

Larcher, V., Jot, J-M., and Vandernoot, G. (1998). Equalization methods in binaural technology. *The 105th Audio Engineering Conference. Preprint 4858*. San Francisco, CA: AES.

Letowski, T., and Makowski, W. (1980). Psychoacoustic investigation of listening with quadraphonic earphones, *Archives of Acoustics*, 5, 59-72.

Letowski, T., Karsh, R., Vause, N., Shilling, R., Ballas, J., Brungart, D., and McKinley, R. (2001). Human Factors Military Lexicon: Auditory Displays. ARL Technical Report ARL-TR-2526. APG, MD: U.S. Army Research Laboratory.

Letowski, T., Mermagen, T., Vause, N., and Henry, P. (2004). Bone conduction communication in noise: A preliminary study, *Proceedings of the XXI International Congress on Sound and Vibrations*. St. Petersburg, Russia: ICSV, 3037-3044.

Letowski, T., Henry, P., and Mermagen, T. (2005). Use of bone conduction transmission for communication with mounted and unmounted soldiers, *Proceedings of the ASNE Human System Integration Symposium on "Enhancing Combat Effectiveness Through War Fighter Performance,"* Arlington, VA: ASNE.

MacDonald, J., Henry, P., and Letowski, T. (2006). Spatial audio through a bone conduction interface, *International Journal of Audiology*, 45, 595-599.

MacDonald, J., and Tran, P. (2008). The effect of HRTF measurement methodology on localization performance in spatial audio interfaces, *Human Factors*, 50(2), 256-263.

McBride, M., Letowski, T., and Tran, P. (2005). Bone conduction head sensitivity mapping: Bone vibrator. ARL Technical Report ARL-TR-3556. APG, MD: U.S. Army Research Laboratory.

McDermott-Well, P. (2005). What is Bluetooth, *Potential*, 23, 33-35.

Makowski, W., and Letowski, T. (1975). An analysis of quadraphonic headphone construction, *Przeglad Telekomunikacyjny*, 6, 195-198.

Moy, C. (2001a). A quick guide to headphones. Headwize Technical Paper. Retrieved June 21, 2007 from http://headwize.com/articles/hguide_art.htm

Moy, C. (2001b). Active noise reduction headphone systems. Headwize Technical Paper. Retrieved June 21, 2007 from http://www.headwize.com/tech/anr_tech.htm

Mozo, B.T. (2001). Acoustical performance. In: Rash, C.E. (Ed.), *Helmet Mounted Displays: Design Issues for Rotary-Wing Aircraft*. Bellingham, WA: SPIE Press.

Mozo, B.T. (2004). Communication enhanced and protection system. White paper. Retrieved on March 25, 2007 from http://www.cep-usa.com/CEPS%20Product%20Description.pdf

Mozo, B.T., and Murphy, B.A. (1997). The assessment of sound attenuation and speech intelligibility of selected active reduction devices and the communication earplug when used with the HGU-56/P aviator helmet. USAARL Report 97-8. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Musicant, A.D., and Butler, R.A. (1984). The influence of pinnae-based spectral cues on sound localization, *Journal of the Acoustical Society of America*, 75, 1195-1200.

Moulton, D. (2004). Lecture 28. Rocking armature transducer: Part 1. White paper. Retrieved December 30, 2006 from http://www.moultonworld.pwp.blueyonder.co.uk/ Lecture28_page.htm

Neitzel, R, Somers, S., and Seixas, N. (2006). Variability of real-world hearing protector attenuation measurements, *Annals of Occupational Hygiene*, 50, 679-691.

Nelson, W.T., Hettinger, L.J., Cunningham, J.A., Brickman, B.J., Hass, M.W., and McKinley, R.L. (1998). Effects of localized auditory information on visual target detection performance using a helmet-mounted display, *Human Factors*, 40(3), 452-460.

Nilo, E.R. (1968). The relation of vibrator surface area and static application force to the vibrator-to-head coupling, *Journal of Speech and Hearing Research*, 11, 805-10.

Nixon, C., McKinley, R., and Steuver, J. (1992). Performance of active noise reduction headsets. In: Dancer, A., Henderson, D., Salvi, A., and Hamernik, R. (Eds.), *Noise-Induced Hearing Loss*. Baltimore, MD: Mosby Year Book.

Pinedo, C., Young, L., and Esken, R. (2006). Effects of combined 3-D auditory/visual cueing system on visual target detection using a helmet-mounted display. Air Force Research Laboratory, Dayton, OH. Report AFRL:-HE-WP-TR-2006-0128.

PM Soldier Warrior (2007). Performance requirements for hearing protection tactical headsets. Fort Belvoir, VA: PM Soldier Warrior. Draft June 14, 2007.

Price, G.R. (2005). Executive summary of the development and validation of AHAAH. White paper. Retrieved June 21, 2007 from http://www.arl.army.mil/ahaah/html_docs/ executive_sum.htm

Price, G.R. (2007). Validation of the auditory hazard assessment algorithm for the human with impulse noise data. *Journal of the Acoustical Society of America*, 122, 2786–2802.

Randall, R., and Holland, H. (1972). The effect of helmet form on hearing: Speech intelligibility and sound localization. HEL Technical Note HEL-TN-10-72. APG, MD: Human Engineering Laboratory.

Rash, C.E. (2006). Head-mounted displays, *Special Operations Technology*, 4, 16-20.

Rodman, J. (2003). The effects of bandwidth on speech intelligibility. Pleasanton, CA: Polycom Corp. White paper. Retrieved June 21, 2007 from: http://www.polycom.com/common/documents/whitepapers/ effect_of_bandwidth_on_speech_intelligibility_2.pdf

Scharine, A., Henry, P., and Binseel, M. (2005). An evaluation of selected communications assemblies and hearing protection systems: A field study. ARL Technical Report ARL-TR-3475. APG, MD: U.S. Army Research Laboratory.

Shanahan, D.F. (1985). Basilar skull facture in U.S. Army aircraft accidents. USAARL Report No. 85-11. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Shaw, E.A.G. (1982). Hearing protector design concepts and performance limitations. In: Alberti, P. (Ed.), *Personal Hearing Protection in Industry.* New York, NY: Raven Press.

Stenfelt, S., Håkansson, B., and Tjellström, A. (2000). Vibration characteristics of bone conducted sound in vitro, *Journal of the Acoustical Society of America*, 107, 422-431.

Thiele, G. (1986). On the standardization of the frequency response of high-quality headphones, *Journal of the Audio Engineering Society*, 34, 956-969.

USACHPPM. (2006). Noise levels of common Army equipment. Retrieved from http://chppm-www.apgea.army.mil/hcp/NoiseLevels.aspx. APG, MD: U.S. Army Center for Health Promotion and Preventive Medicine.

Vaudrey, M.A., and Sachindar, S. (2003). A real-time audio tele-presence device for remote acoustic monitoring. ARL Technical Report CR-0502. APG, MD: U.S. Army Research Laboratory.

Wagstaff, A.S., Tvete, O., and Ludvigsen, B. (1996). The effect of a headset leakage on speech intelligibility in helicopter noise, *Aviation, Space, and Environmental Medicine*, 67, 1034-1038.

Wenzel, E.M., Arruda, M., Kistler, D.J., and Wightman, F.L. (1993) Localization using nonindividualized head-related transfer functions, *Journal of the Acoustical Society of America*, 94, 111-123.

Whitaker, L., Peters, L., and Garinther, G. (1989). Tank crew performance: Effects of speech intelligibility on target acquisition and subjective workload assessmen, *Proceedings of the 33rd Human Factors Society Annual Meeting.* Santa Monica, CA: Human Factors Society.

White, G.D. (1987). *The Audio Dictionary*. Seattle, WA: University of Washington Press.

Whittle, L.S. (1965). A determination of the normal threshold of hearing by bone conduction, *Journal of Sound and Vibration*, 2, 227-248.

Wightman, F.L., and Kistler, D.J. (1989a). Headphone simulation of free-field listening. I: Stimulus synthesis, *Journal of the Acoustical Society of America*, 85, 858-867.

Wightman, F.L., and Kistler, D.J. (1989b). Headphone simulation of free-field listening. II: Psychophysical validation, *Journal of the Acoustical Society of America*, 85, 868-878.

Wikipedia (2007). Noise-cancelling headphone. White paper. Retrieved June 15, 2007 from: http://en.wikipedia.org/wiki/Noise-cancelling_headphone

Watson, N. (1938). Limits of audition for bone conduction, *Journal of the Acoustical Society of America,* 9, 294-300.

Wong, D., and Baker, C. (1988). Pain in children: comparison of assessment scales, *Pediatric Nursing,* 14(1), 9-17.

Zahorik, P. (2000). Limitations in use of Golay codes for head-related transfer function measurements, *Journal of the Acoustical Society of America*, 107, 1793-1796.

Zotkin, D.N., Duraiswami, R., Grassi, E., and Gumerov, N.A. (2006). Fast head-related transfer function measurement via reciprocity, *Journal of the Acoustical Society of America*, 120, 2202-2215.

Part Three

The Human Visual and Auditory Sensory Systems

In addition to understanding the visual and auditory displays that are part of the HMD system, it is critical to understand the properties of the human sense organs – the eyes and the ears – and their associated perceptual systems – vision and audition. Understanding the anatomy, structure, physiology and functions of these systems is necessary to design an effective human-machine interface. It is instructive to follow the energy – light and sound – generated by objects in the environment as it is captured by the sense organs and transformed into electrical signals that follow the sensory pathways to the human brain. The related perceptual experiences of vision and audition are the result of complex processes that are not yet completely understood.

6 BASIC ANATOMY AND PHYSIOLOGY OF THE HUMAN VISUAL SYSTEM

Corina van de Pol

The human eye is a complex structure designed to gather a significant amount of information about the environment around us. It is *the* sensor used by the Warfighter in the visually-rich battlespace. In designing a head/helmet-mounted display (HMD) system for the Warfighter, the human visual system (which begins with the eyes) could be considered as an integral component of the HMD and not as a separate and different system that subsequently is mated with the HMD. It is therefore important that HMD designers have an understanding of both anatomy and function of the human eye itself.

In the following chapter (Chapter 7, *Visual Function*), the functional operations of the human eye, its pointing and tracking mechanisms and the integration into a binocular visual system will be described. In this chapter, the goal is to provide the HMD designer with a basic understanding of the anatomy and physiology of this critical element of the human visual system. This chapter provides a brief overview of the visual system (inclusive of the eye organ itself), beginning at the front surface of the eye and progressing to the primary visual cortex at the back of the brain. Topics include:

- The Protective Structures of the Eye
  - The Orbit
  - The Lids
  - The Sclera
- The Anterior Segment of the Eye
  - The Cornea
  - The Aqueous Humor
  - The Iris
  - The Crystalline Lens and Ciliary Muscle
- The Posterior Segment of the Eye
  - The Retina
  - The Vitreous Humor
- The Visual System Pathways to the Brain
  - The Optic Nerves and Optic Tracts
  - The Lateral Geniculate Nucleus
  - The Visual Cortex

For more detailed discussions of the human eye's anatomy and physiology, the reader should refer to the large volumes of texts available, e.g., *Adler's Physiology of the Eye* (Kaufman and Alm [Eds.], 2003).

The Protective Structures of the Eye

The two orbits, sometimes referred to as "sockets," that protect the human eyes are situated at the front of the skull, each with a wider opening to the front narrowing to a small opening at the rear where the optic nerve exits to connect through the visual pathways and the brain. The orbits are angled outward approximately 23° with respect to the midline of the skull. The human eye itself is approximately 24 millimeters (mm) (0.94 inches [in]) in diameter and occupies about 25% of the volume of the orbit, allowing for the extraocular muscles, blood vessels, nerves, orbital fat and connective tissue that surround and support the eye (Figure 6-1). The orbit surrounds and supports most of the human eye, while the cornea and part of the anterior globe extend somewhat beyond the orbital rims. These structures are protected by the eyelids.

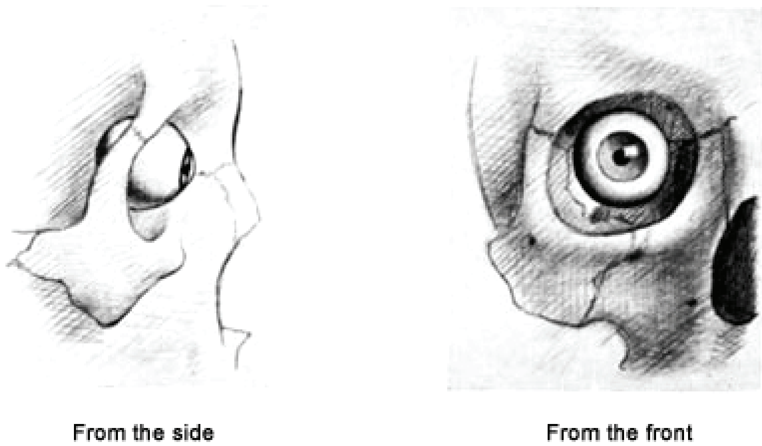


Figure 6-1. The position f the eye in its socket (Wolff, 1933).

The upper and lower eyelids form an aperture that is generally 30 mm (1.2 in) wide and 10 to 12 mm (0.4 to 0.5 in) high when the eye is “open.” The lids themselves have cartilage-like tarsal plates within their structure that provide shape to the lids and additional strength for protection of the eye. Each lid has a row of cilia or eyelashes that are very sensitive to touch or particles near the eye, which when stimulated bring on the blink reflex. The lids also contain the glands responsible for maintenance of the tear layer.

The globe itself is predominately formed of and protected by the sclera that extends from the edges of the clear cornea at the front of the eye (the “limbus”) to the optic nerve at the back of the eye. The sclera is a thick, opaque white tissue that covers 95% of the surface area of the eye. It is approximately 530 microns (μm) in thickness at the limbus, thinning to about 390 μm near the equator of the globe and then thickening to near 1 mm (0.04 in) at the optic nerve. At the posterior aspect of the eye, the sclera forms a netlike structure or “lamina cribrosa” through which the optic nerve passes. The sclera also serves as the anchor tissue for the extraocular muscles.

The Anterior Segment of the Eye

The portion of the eye visible to the observer without special instrumentation is considered the anterior (or “front”) segment of the eye. Most of the structures responsible for focusing images onto the retina of the eye are here. The cornea is the primary focusing structure, providing about 75% of the focusing power of the eye. The crystalline lens provides the remaining variable focusing power and serves to further refine the focus, allowing the eye to focus objects at different distances from the eye. The iris controls the aperture or pupil of the eye for different light levels. The iris is actually an extension of the ciliary body, a structure that has multiple functions in the anterior segment, from production of the fluid that fills the anterior segment (aqueous humor) to suspension and control of the shape of the crystalline lens of the eye. Figure 6-2 shows most of the major structures of the human eye, including the components of the anterior segment, the protective sclera and the posterior segment (described in the next section).

The cornea

The cornea is a unique biological tissue that is transparent to light and contains no blood vessels. This small transparent dome at the front of the eye is approximately 11 mm (0.43 in) in diameter and 500 μm thick in the center, thickening to around 700 μm at the periphery. At the very edge of the cornea, transparency is slowly lost over a 1-mm (0.04-in) range in an area known as the "limbus", which is where the cornea integrates into the opaque sclera. The cornea is more curved than the rest of the globe with an average radius of curvature of 7.7 mm (0.3 in), while the radius of curvature of the globe is approximately 12 mm (0.5 in).

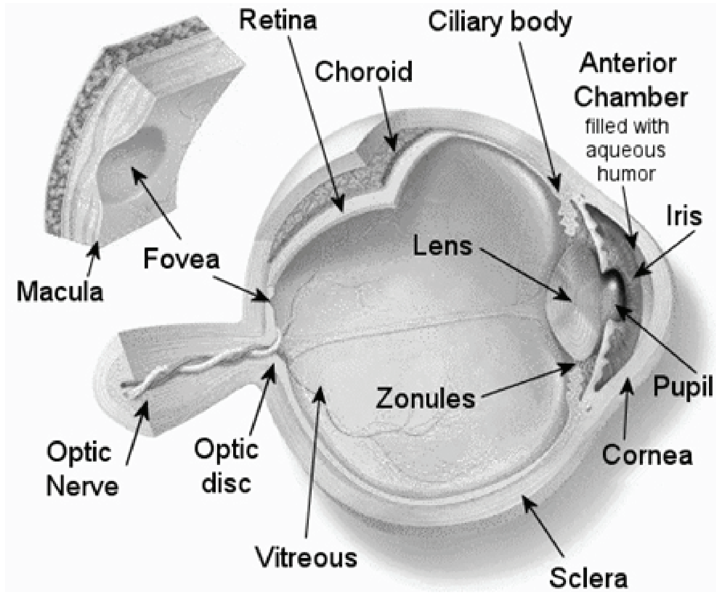
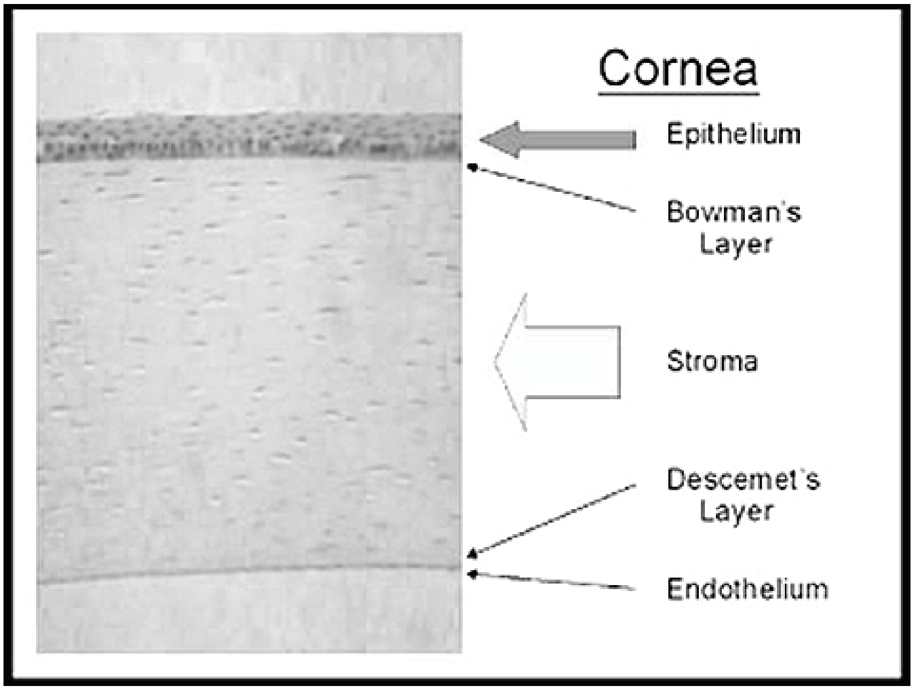


Figure 6-2. Cross-sectional view of the eye (*http://www.gimbeleyecentre.com/images/Cross_Section_Labelled.gif).*

With the primary function of transmitting and focusing light into the eye, all the structures of the cornea are very specifically arranged (Figure 6-3). About 90% of the cornea is made up of evenly spaced collagen fibrils arranged in sections that crisscross to cover the entire extent of the cornea. This layer is known as the "stroma" and it provides not only transparency, but strength. Four more layers make up the remaining 10% of the cornea, the epithelium and Bowman's layer at the front of the cornea and Descemet's membrane and the endothelium at the back of the cornea.

The epithelium of the cornea, much like the epithelium of the skin, serves as a barrier to bacteria or other pathogens. Additionally, the epithelium helps to maintain the stroma at a proper level of hydration by preventing fluid from entering the stroma through its tight cell junctions and the pumping of a small portion of fluid out of the stroma. Bowman's layer is a very thin (12 μm) membrane right beneath the epithelium and, in mammals, is only found in primates. Its purpose is not entirely known, although it may aid in protection of the stroma.

At the back or posterior aspect of the cornea is another very thin membrane called Descemet's membrane that is between 10 to 15 μm thick. It also is felt to have some protective function. The endothelium is a single layer of cells at the very posterior aspect of the cornea. The endothelium is in direct contact with the aqueous humor, the fluid that fills the anterior chamber of the eye. The endothelium pumps nutrients, such as glucose, from the aqueous humor into the cornea while actively pumping fluid out of the cornea. The hydration balance maintained by the endothelium and somewhat assisted by the epithelium is important to the transparency of the cornea, since excess fluid would disturb the regularity of the corneal fibrils and result in increased light scatter. In mild cases of edema, such as may occur when contact lenses are worn too long or under hypoxic conditions, the cornea may become slightly cloudy (Jones and Jones, 2001; Liesegang, 2002; Morris et al., 2007). Under more extreme conditions, such as anoxic conditions, or in cases of endothelial dystrophies, the swelling of the stroma could result in complete opacity of the cornea.



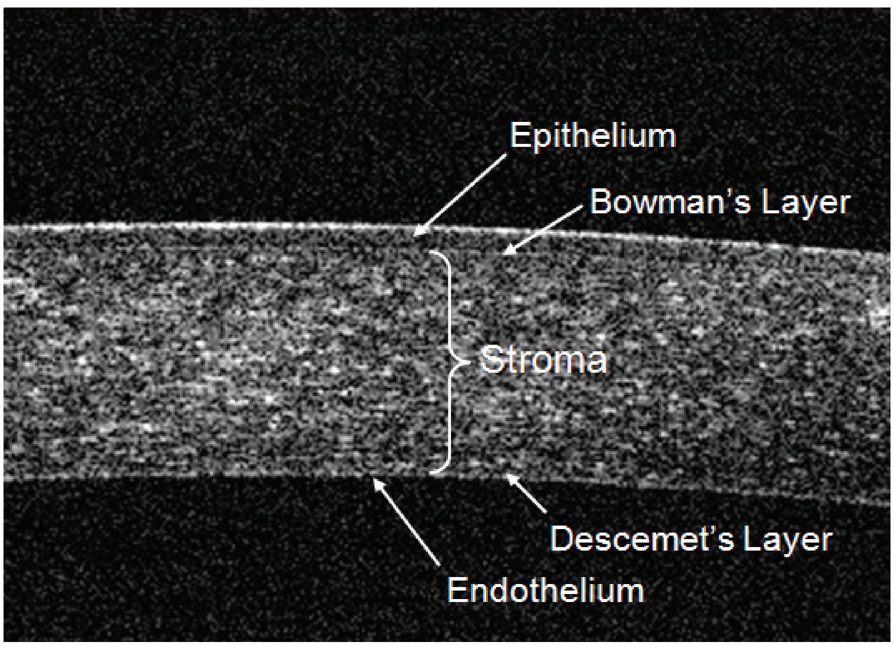


Figure 6-3. Cross-section of the cornea (top) (adapted from http://www.opt.pacificu.edu/ce/catalog/10603-AS/Cornea.jpg); actual Optical Coherence Tomography image of a cross-section of the author's cornea.

The aqueous humor

The fluid that fills the anterior chamber of the eye, that area between the cornea and the front surface of the crystalline lens, is called the aqueous humor. Aqueous is produced by the ciliary body that is just posterior to the root of the iris and extends backwards along the inner globe to the anterior aspect of the retina (Figure 6-2). Aqueous finds its way into the anterior chamber by flowing between the crystalline lens and the iris through the pupil.

Aqueous has two functions; it provides nutrients to the cornea and is part of the optical pathway of the eye. Aqueous humor is basically a fortified blood plasma that circulates in the anterior chamber, providing nutrients to the cornea and the crystalline lens. It is a transparent fluid with an index of refraction of 1.333, which is slightly less than the index of refraction of the cornea (1.376) and less than the index of refraction of the lens (gradient index of 1.406 to 1.386). As is discussed in another chapter, it is these differences in index of refraction between media coupled with the curvature of the various optical surface interfaces that result in the bending of light at each interface.

As nutrients are drawn from the aqueous into the cornea by the endothelium, the aqueous fluid is circulated out of the eye and replaced by newly produced aqueous produced. To move out of the anterior chamber, it flows out of the eye primarily through the trabecular meshwork, a "drainage" system that lies behind the limbus in the angle between the cornea and the anterior iris. There is some resistance to outflow of aqueous at the trabecular meshwork that serves to maintain a pressure within the eye of approximately 15 mmHg. If there were no resistance, the eye would lose its shape and therefore its optical integrity. If there is too much resistance (or too much production of aqueous), the pressure in the eye may exceed the eye's tolerance and damage to the optic nerve may occur, a condition known as "glaucoma."

Glaucoma generally results in a loss of mid-peripheral vision with sparing of central vision until the condition has progressed significantly. It is most commonly hereditary with a higher prevalence in certain ethnic groups (Friedman et al., 2004; Leske, 2007; Rivera, Bell, and Feldman, 2008; Wadhwa and Higginbotham, 2005); however, it may occur in individuals without a family history of glaucoma or may result secondarily to blunt trauma to the eye Cavallini et al., 2003; Kenney and Fanciullo, 2005; Sihota, Sood, and Agarwal, 1995). Glaucoma can be slowly progressive, as in the case of *primary open angle glaucoma* (POAG) or *low tension glaucoma* (LTG), and the loss of vision may be initially barely noticeable. A third type of glaucoma, a*ngle closure glaucoma* (ACG), is more acute and may or may not be accompanied by pain in and around the eye when it occurs (Ang and Ang, 2008; Congdon and Friedman, 2003). During routine eye exams, measurement of intraocular pressure and assessment of visual fields are essential for early detection of glaucoma.

The iris

The iris is visible through the cornea and is what gives the eye its "color." All irides have a dark pigmented posterior layer; it is the amount of pigment in the anterior or stromal layer that produces different colors. A "blue" eye results from the selective absorption of long wavelength light by the stroma of the iris and the reflection of short wavelength (blue) light by the posterior pigmented layer. In a "brown" eye almost all visible wavelengths are absorbed by the iris stroma and very little light is left to reflect out of the eye.

The main purpose of the iris, however, is to block excess light from entering the eye and to control the iris aperture or "pupil" for differing amounts of ambient light (Figure 6-2). There are two opposing muscles in the iris; the sphincter muscles that serve to constrict the pupil and the dilator muscles that serve to dilate the pupil. Parasympathetic nerves innervate the sphincter muscles and sympathetic nerves innervate the dilator muscles. It's because the sympathetic system is heightened relative to the parasympathetic system during "fight or flight" situations that pupils dilate when danger is sensed. Most pupil responses are controlled by a complex set of signals sent through the midbrain (specifically the Edinger-Westphal nucleus) in response to the amount of light striking the retina or as part of the accommodative triad (discussed in Chapter 7, *Visual Function*).

There are very few conditions that affect the iris directly; however, changes in the normal response of the pupil to light or accommodation can result from lesions in the neural pathways or direct trauma to the iris. If the iris does not constrict in response to light, likely the parasympathetic system has been affected by such conditions known as Adie's tonic pupil or *third nerve palsy*. This lack of constriction may also occur in response to anticholinergic drugs, such as found in scopolamine patches, or adrenergic drugs, such as found in some eye drops used for “red eye.” If the iris fails to dilate under low light conditions, likely the sympathetic system has been affected by a condition known as *Horner’s syndrome*.

The crystalline lens and ciliary muscle

Like the cornea, the crystalline lens is a transparent structure. Unlike the cornea, it has the ability to change its shape in order to increase or decrease the amount of refracting power applied to light coming into the eye. Transparency is maintained by the regularity of elongated fiber cells within the lens. These cells originate at the equator of the lens and lay down across the surface of other fiber cells while growing toward the anterior portion of the lens and the posterior portion of the lens until they meet at the central sutures. During elongation they pick up crystallins, hence the name “crystalline lens.” It is these crystallins that give the lens a higher index of refraction than the aqueous and vitreous humors. The gradient index of refraction of the lens ranges from about 1.406 through the center to about 1.386 through the more peripheral portions of the lens (Hecht, 2002). This is due to the fiber cells near the surface having a lower index of refraction than deeper cells, which results in a decrease in spherical aberrations and therefore a more refined quality of focus.

The lens is surrounded by an elastic extracellular matrix known as the “capsule.” The capsule not only provides a smooth optical surface, but it provides an anchor for the suspension of the lens within the eye. A meshwork of nonelastic microfibrils or “zonules” anchor into the capsule near the equator of the lens and, much like a suspension system around a trampoline, connect into the ciliary muscle (Figure 6-2). When the ciliary muscle is relaxed, the tension on the zonules is highest and the lens is “pulled” to its flattest curvature. This generally results in focus for a distant object when the eye is emmetropic (e.g. does not have any refractive errors, such as myopia or hyperopia). When the ciliary muscle contracts, it moves slightly forward, but mostly inward towards the center line of the eye. This releases the tension on the zonules and allows the lens to take up its preferred shape, which is more rounded and thereby more powerful. This increases the focal power of the eye to focus on nearby objects.

Since the lens continues to lay down fiber cells throughout life, it becomes denser and less flexible resulting in a loss of the ability to change focus for near objects with age. This process called *presbyopia* will be covered in a later chapter. A *cataract* is a condition in which the crystalline lens starts to develop opacities or lose its transparency. Cataracts can be associated with environmental factors such as smoking, health conditions such as diabetes, or the use of certain medications such as corticosteroids (Delcourt et al., 2000; Rowe et al., 2000). The effect of cataracts on vision is generally a reduction in contrast sensitivity, an increase in glare and halos at night and some shift in color sensitivity due to the “yellowing” of the lens.

The Posterior Segment of the Eye

The retina lines the interior of the posterior portion of the globe and is where images are formed. Initial processing of the image occurs at this highly specialized sensory tissue. Vitreous is the clear gel that fills the posterior segment and serves to provide for light transmission through the eye and to protect the retina.

The retina

The retina is a mostly transparent thin tissue designed to capture photons of light and initiate processing of the image by the brain. The average thickness of the retina is 250 μm and it consists of 10 layers (Figure 6-4). From the surface of the retina to the back of the eye the layers are the inner limiting membrane, the nerve fiber layer (axons of the ganglion cells), the ganglion cell layer, the inner plexiform layer (synapses between ganglion and bipolar or amacrine cells), the inner nuclear layer (horizontal, bipolar amacrine and interplexiform cells, along with the retina spanning glial cells), the outer plexiform layer (synapses between bipolar, horizontal and photoreceptor cells), the outer nuclear layer (photoreceptor cells), the outer limiting membrane, the receptor layer (outer and inner segments of the photoreceptor cells) to the retinal pigment epithelium (RPE). The RPE is the outmost layer of the retina and serves as the primary metabolic support for the outer segment of the receptor cells and also acts as the final light sink for incoming photons that reduces intraocular glare. Its light absorbing pigmentation is why the pupil appears black.

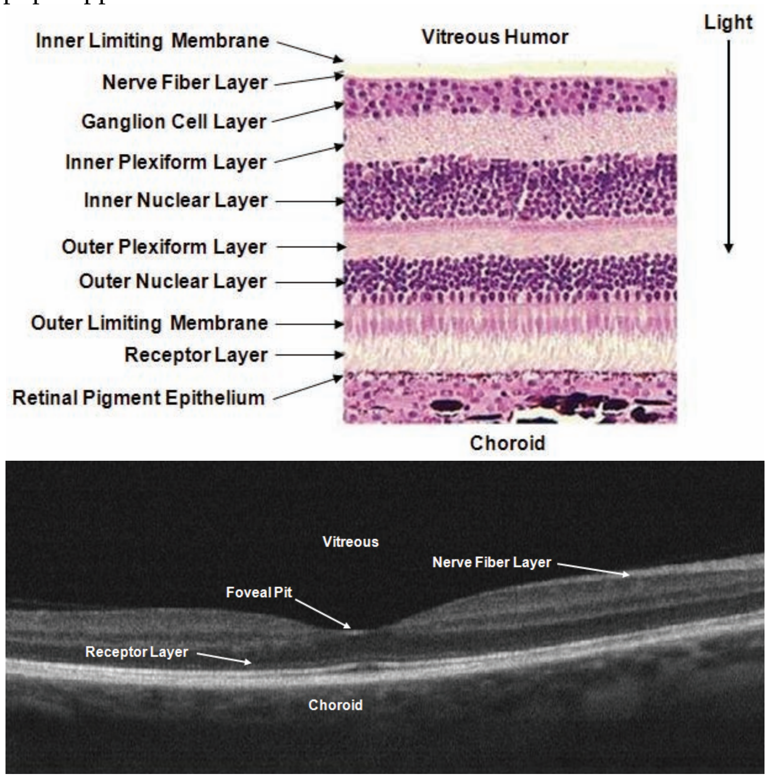


Figure 6-4. Cross-section of the retina (top) (adapted from http://www.opt.pacificu.edu/ce/catalog/12059-PS/Fig1N.jpg); actual Optical Coherence Tomography image of a cross-section of the author's retina.

The fact that the receptor layer is deep within the retina means that photons of light actually must pass through most layers of the retina before reaching the receptors. The receptors absorb and convert photons to neural signals, which are than processed through the network of bipolar, horizontal, amacrine and ganglion cells. The output axons of the ganglion cells form the nerve fiber layer that collects at the optic nerve to exit the eye. It's the intricate interconnections of the various neural cells in the retina that complete the first processing of the visual information being sent to the brain.

There are two types of receptors in the receptor layer, rods and cones, essentially named for their shape. The outer segment of the receptor cells contain the light sensitive visual pigment molecules called "opsins" in stacked disks (rods) or invaginations (cones). There are approximately 5 million cones and 92 million rods in the normal adult retina. Cones provide the ability to discern color and the ability to see fine detail and are more concentrated in the central retina. Rods are mainly responsible for peripheral vision, vision under low light conditions and are more prevalent in the mid-peripheral and peripheral retina.

At the most posterior aspect of the retina, where most of the light that the eye receives is focused, is a region called the macula lutea. The macula is an area approximately 5 to 6 mm in diameter which has a greater density of pigments (lutein and zeaxanthine). These pigments help to protect the retinal neural cells against oxidative stress. Within the macular area is the fovea centralis, the small region at the center of the retina where vision is most acute. In this small 1.5 mm (0.06 in) diameter area there are no rods, only cones and the overlying neural layers are effectively swept away so that there is a depression in the retina. The average thickness of the retina drops to around 185 ìm in this “foveal pit.” The area immediately outside the fovea is called the parafoveal region and is where there is a transition from cone-dominated to rod-dominated retina.

The retina receives its nourishment from two sources, the retinal vasculature serves the inner layers of the retina and the choroidal vasculature, which lies between the RPE and the sclera, serves the metabolically active RPE and outer layers of the retina. In order to maximize photon capture in the central retina, the retinal capillary system does not extend in to the fovea centralis, an area known as the foveal avascular zone. This area depends on the blood supply provided by the choriocapillaris.

One of the most common conditions that can affect the retina is age-related macular degeneration (ARMD), in which there is a loss of vision in the center of the visual field (Klein et al., 2004; Nicolas et al., 2003; van Leeuwen et al., 2003). In ARMD, the ability of the retinal pigment epithelium to remove the waste produced by the photoreceptor cells after processing light coming into the eye is reduced. As a result, waste builds up in the form of “drusen.” These drusen further disrupt the metabolic process and eventually the retina starts to deteriorate. If blood vessels from the choriocapillaris break through (“wet ARMD”) the condition can become significantly worse. ARMD is generally hereditary and early signs are detectable through routine eye exams.

The vitreous humor

The vitreous body is a gel-like structure that fills the posterior portion of the globe. Vitreous humor is comprised of collagen fibrils in a network of hyaluronic acid and is a clear gel (Kaufman and Alm [eds.], 2003). The vitreous body is loosely attached to the retina around the optic nerve head and the macula and more firmly attached to the retina at the ora serrata just posterior to the ciliary body. The connections at the anterior portion of the vitreous body help to keep the anterior and posterior chamber fluids separated. The connections around the optic nerve and macula help to hold the vitreous body against the retina.

With aging, the vitreous starts to liquefy and shrink. When this happens, aqueous from the anterior chamber can get into the posterior chamber of the eye. Additionally, there can be increased tugging at the attachment points on the retina causing a release of cells that the individual sees as “floaters.” If there is significant traction at the attachment points, the retina can be pulled away from the inner globe and a retinal tear or detachment can result.

The Visual System Pathways to the Brain

The neural signals initially processed by the retina travel via the axons of the ganglion cells through the optic nerves, dividing and partially crossing over into the optic chiasm and then travelling via the optic tracts to the lateral geniculate nucleus (LGN). From the LGN, the signals continue to the primary visual cortex, where further visual processing takes place (Figure 6-5).

The optic nerves and optic tracts

The optic nerve of each eye consists of a bundle of approximately 1 million retinal ganglion cell axons. The nerve connects to the posterior aspect of the eye in a position that is about 15° nasal to the macula. The connection is referred to as the optic nerve head and is visible when looking into the eye using an ophthalmoscope. The optic nerve head is approximately 1.8 mm (0.07 in) in diameter. Since there are no photoreceptors (rods or cones) overlying the optic nerve head, there is a small blind spot or “scotoma” of approximately 5° in size about 15° temporal to fixation in the visual field of each eye. When both eyes are open, the blind spot of each eye is “filled in” by the visual field of the other eye.

The optic nerves of each eye continue posteriorly and then meet at the optic chiasm. It is here that axons of neurons from the nasal retina (temporal visual field) cross to the opposite or “contralateral” optic tract (e.g. axons from the right eye temporal visual field cross to the optic tract on the left side of the brain). Axons of neurons from the temporal retina (nasal visual field) continue along the same side or "ipsilateral" optic tract (same side of the brain). This means that visual signals from the right side of the visual field are traveling to the brain via the left optic tract and signals from the left visual field are traveling via the right optic tract. Each optic tract terminates at its LGN.

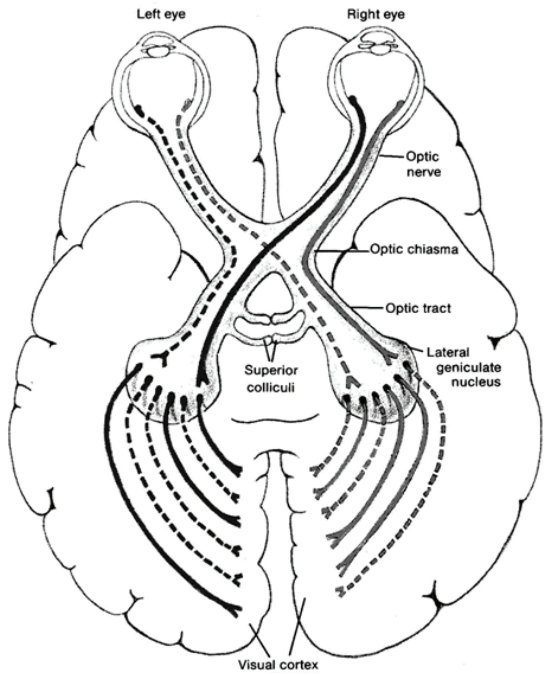


Figure 6.5. Visual system pathway (http://www.skidmore.edu/~hfoley/images/Brain.top.jpg).

If a stroke, aneurism or tumor causes damage along the visual pathway, it is often possible to diagnose the exact location of the insult by measuring the visual field. For instance, a pituitary tumor would appear near the optic chiasm and the impact on the visual field would be on the fibers that are crossing to the other side of the brain. Since these fibers are from the nasal retina of each eye, the loss of vision would be in both temporal visual fields or a bitemporal visual field defect (Figure 6.5). Whereas an insult to one of the optic tracts would result in a loss of vision to the opposite or contralateral side of the visual field. For instance, a defect to the right optic tract would cause a loss of the left visual field of both eyes (the temporal visual field of the left eye and the nasal visual field of the right eye).

The lateral geniculate nucleus (LGN)

The LGN is a paired structure located at the dorsal thalamus. It is here that visual information to the brain, specifically the visual cortex, appears to be regulated and the first stage of coordinating vision from both eyes begins. Each LGN has six layers, three receiving input from the right eye and three receiving input from the left eye. Because of the way the retinal ganglion cell axons are distributed through the chiasm and on to the optic tracts, the information processed in any one layer of the LGN represents specific areas of the visual field for one eye.

Four of the layers are composed of the Parvocellular (small) ganglion cells from the retina that are primarily from the fovea. These cells are most sensitive to color and fine detail. Two of the layers are composed of the

Magnocellular (large) ganglion cells from the retina. These cells are mostly from the perifoveal and more peripheral retina and are largely responsible for the processing of motion.

The LGN then sends forward neurons via the *optic radiations* to the p*rimary visual cortex*.

The Visual Cortex

The visual cortex in the occipital lobe of the brain is where the final processing of the neural signals from the retina takes place and "vision" occurs. The occipital lobe is at the most posterior portion of the brain. There are a total of six separate areas in the visual cortex, known as the V1, V2, V3, V3a, V4 and V5.

The p*rimary visual cortex* or V1 is the first structure in the visual cortex where the neurons from the LGN synapse. In V1, the neural signals are interpreted in terms of visual space, including the form, color and orientation of objects. V1 dedicates most of its area to the interpretation of information from the fovea. This mapping is known as "cortical magnification" and is typical in primates and animals that rely on information from the fovea for survival. The signals then pass through to V2 where color perception occurs and form is further interpreted.

As the neural signals continue further into other areas of the visual cortex, more associative processes take place. In the portions of the visual cortex that make up the p*arietal visual cortical areas*, motion of objects, motion of self through the world and spatial reasoning occur. In the *temporal visual cortical areas*, including the *middle temporal* (V5) area, recognition of objects through interpretation of complex forms and patterns occurs. The final psychological and perceptual experience of vision also includes aspects of memory, expectation/prediction and interpolation subserved by other apparently non-visual areas of the brain.

References

Ang, L.P., and Ang, L.P. (2008). Current understanding of the treatment and outcome of acute primary angle-closure glaucoma: An Asian perspective. *The Annals, Academy of Medicine, Singapore*, 37(3), 210-215.

Cavallini, G.M., Lugli, N., Campi, L., Pagliani, L., and Saccarola, P. (2003). Bottle-cork injury to the eye: a review of 13 cases. *European Journal of Epidemiology*, 13(3), 287-291.

Congdon, N.G., and Friedman, D.S. (2003). Angle-closure glaucoma: impact, etiology, diagnosis, and treatment. *Current Opinion Ophthalmology*, 14(2), 70-73.

Delcourt, C., Cristol, J.P., Tessier, F., Leger, C.L., Michel, F., and Papoz, L. (2000). Risk factors for cortical, nuclear, and posterior subcapsular cataracts: the POLA study. Pathologies Oculaires Liees a l'Age. *American Journal of Epidemiology*, 151(5), 497-504.

Friedman, D.S., Wilson, M.R., Liebmann, J.M., Fechtner, R.D., and Weinreb, R.N. (2004). An evidence-based assessment of risk factors for the progression of ocular hypertension and glaucoma. *American Journal of Epidemiology*, 138(3 Suppl), S19-31.

Hecht, E. (2002). *Optics*. Reading, MA: Addison-Wesley.

Jones, L.W., and Jones, D.A. (2001). Non-inflammatory corneal complications of contact lens wear. *Contact Lens and Anterior Eye*, 24(2), 73-79.

Kaufman, P.L., and Alm, A. (Eds.). (2003). *Adler's Physiology of the Eye* (10th ed.). St. Louis: Mosby.

Kenney, K.S., and Fanciullo, L.M. (2005). Automobile air bags: friend or foe? A case of air bag-associated ocular trauma and a related literature review. *Optometry*, 76(7), 382-386.

Klein, R., Peto, T., Bird, A., and Vannewkirk, M.R. (2004). The epidemiology of age-related macular degeneration. *American Journal of Epidemiology*, 137(3), 486-495.

Leske, M.C. (2007). Open-angle glaucoma – an epidemiologic overview. *Ophthalmic Epidemiology*, 14(4), 166-172.

Liesegang, T.J. (2002). Physiologic changes of the cornea with contact lens wear. *Contact Lens Association of Ophthalmologists Journal*, 28(1), 12-27.

Morris, D.S., Somner, J.E., Scott, K.M., McCormick, I.J., Aspinall, P., and Dhillon, B. (2007). Corneal thickness at high altitude. *Cornea,* 26(3), 308-311.

Nicolas, C.M., Robman, L.D., Tikellis, G., Dimitrov, P.N., Dowrick, A., and Guymer, R.H. (2003). Iris colour, ethnic origin and progression of age-related macular degeneration. *Clinical and Experimental Ophthalmology,* 31(6), 465-469.

Rivera, J.L., Bell, N.P., and Feldman, R.M. (2008). Risk factors for primary open angle glaucoma progression: what we know and what we need to know. *Current Opinion Ophthalmology*, 19(2), 102-106.

Rowe, N.G., Mitchell, P.G., Cumming, R.G., and Wans, J.J. (2000). Diabetes, fasting blood glucose and age-related cataract: the Blue Mountains Eye Study. *Ophthalmic Epidemiology*, 7(2), 103-114.

Sihota, R., Sood, N.N., and Agarwal, H.C. (1995). Traumatic glaucoma. *Acta Ophthalmol Scand,* 73(3), 252-254.

van Leeuwen, R., Klaver, C.C., Vingerling, J.R., Hofman, A., and de Jong, P.T. (2003). Epidemiology of age-related maculopathy: a review. *European Journal of Epidemiology,* 18(9), 845-854.

Wadhwa, S.D., and Higginbotham, E.J. (2005). Ethnic differences in glaucoma: prevalence, management, and outcome. *Current Opinion Ophthalmology,* 16(2), 101-106.

Wolff, E. (1933). *Anatomy for Artists*. London: Lewis, H.K.

7 VISUAL FUNCTION

Thomas O. Salmon
Corina van de Pol
Clarence E. Rash

In order to design a helmet-mounted display (HMD) that most effectively couples with the eye and optimizes visual performance, the designer should have a basic understanding of the capabilities and performance of the visual system. This includes an understanding of the following:

- The physical nature of light
- How the eye forms an image
- Refractive errors and their correction
- Spatial vision, including visual acuity and contrast sensitivity
- Peripheral vision
- Adaptation to high and low illumination
- Color vision
- Accommodation
- The eye's temporal responsiveness
- Eye movements
- Binocular vision

The Physical Nature of Light

While vision is predominately a physiological process, it is all made possible by that part of the electromagnetic (EM) spectrum we call light. We see the world around us and the objects in it because of light energy that is either emitted by or reflected off of these objects. An elementary understanding of light and its role in vision can be both instructive and useful.

The universe is filled with energy. The total span of this energy is represented by the EM spectrum (Figure 7-1). At any given place along the spectrum, the energy is characterized by a specific frequency (or wavelength). Frequency (f) is inversely proportional to wavelength (λ), as shown in Equation 7-1, where c is the speed of light.

$$f = \frac{c}{\lambda} \qquad \text{Equation 7-1}$$

While continuous in nature, it is convenient to divide the spectrum into subdivisions. At one end of the spectrum is the highest frequency (shortest wavelength) subdivision known as the gamma rays (Figure 7-2). Gamma rays have frequencies to the order of 10^{20} Hertz (Hz) and higher, and wavelengths of 10^{-11} meters. These rays have the more energy than any other part of the EM spectrum. They are produced by atoms undergoing radioactive decay and by nuclear explosions. Gamma rays have practical applications in medicine and in industry. In medicine they are used to kill cancerous cells and sterilize medical equipment. In the food industry, they are used to kill bacteria and insects and to maintain freshness (Tauxe, 2003).

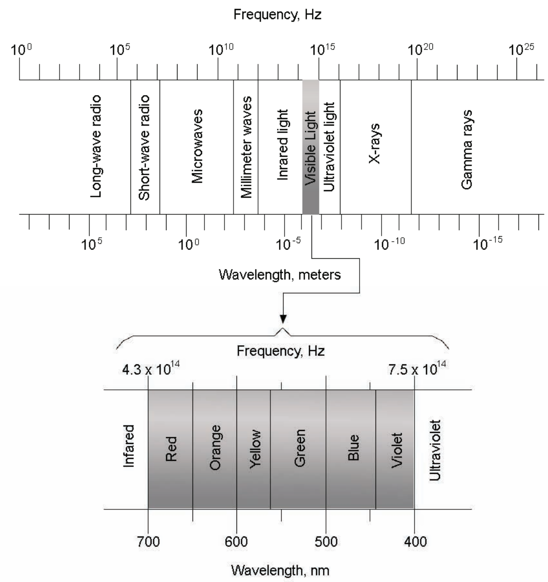


Figure 7-1. The electromagnetic spectrum.

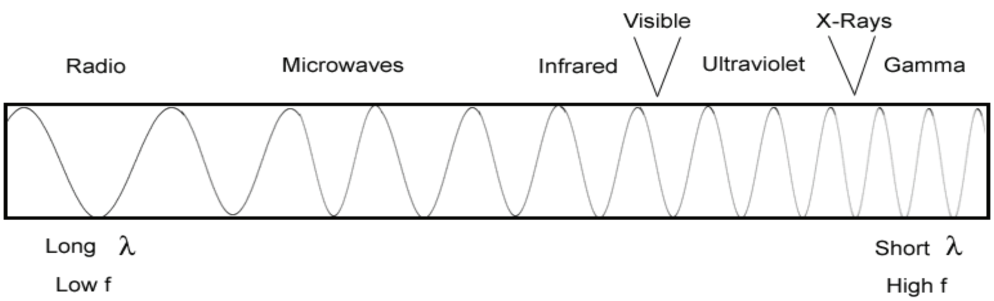


Figure 7-2. The electromagnetic spectrum as a range of frequencies and wavelengths.

At the other end of the spectrum are radio waves, having the lowest frequencies ($<10^6$ Hz) and longest wavelengths ($>10^2$ meters). The radio wave part of the spectrum is often further divided into short- and long-waves. This part of the spectrum is the least energetic. Uses of radio waves include AM and FM radio, television and cell phones.

For our purpose the most important part of the EM spectrum is visible light, i.e., that part of the spectrum that the human eye can detect or “see.” It is a very small part of the complete spectrum. When studying vision, it is customary to refer to the wavelength of a specific part of the visible light spectrum. There are no exact bounds to the visible spectrum. A typical human eye will respond to wavelengths from 400 to 700 nm, but this can vary from person to person. As will be explained later, during daylight the eye typically has its maximum sensitivity at around 555 nm, and in low illumination, the eye is optimized at approximately 510 nm.

Different wavelengths within the visible spectrum are associated with certain colors. That is, when we see light of a particular wavelength, we perceive a particular color. Sir Isaac Newton is credited with first showing that light shining through a prism will be separated into its different wavelengths and will thus show the various colors of visible light. This separation of visible light into its different colors is known as dispersion.

Newton divided the visible spectrum into seven named colors: Red, orange, yellow, green, blue, indigo, and violet, which are represented by the mnemonic “ROYGBIV.” For accuracy, “indigo” is not actually observed in the spectrum but is traditionally added to the list so that there is a vowel in Roy's last name. The red is associated with the longer wavelengths and violet with the shorter wavelengths. Between red and violet, there is a continuous range of wavelengths and, hence, colors.

The last important principle of light (and the entire EM spectrum), for the purpose of this discussion, is known as particle-wave duality. It is generally accepted that light is composed of packets of energy called photons, which display some of the properties of waves and some of the properties of particles. The energy of an individual photon is proportional to its frequency; the higher the frequency (or shorter the wavelength), the greater the energy. The photon represents the smallest amount of light energy that can be produced. The human eye is remarkable in that under ideal conditions, a rod receptor in the retina at the back of the eye can respond to the energy of a single photon.

The particle nature of light explains the reflection of light rays and the photoelectric effect; the wave nature of light explains refraction, interference and polarization. Simply stated, light exhibits properties both of particles and of waves. For human vision and the following discussion of how the eye forms an image, the dual nature of light will be used in the sense that the path of light entering the eye will be treated as light rays associated with waves that obey the laws of reflection and refraction.

How the Eye Forms an Image

In a simplistic representation, vision can be separated into two mechanisms: one that encompasses the collection and focusing of light on the photoreceptors in the retina at the back of the eye and the one that consists of the physiological and cognitive processes that follow.

Consider the diagram in Figure 7-3. A simple object of interest, here represented as a tree, is depicted. As a luminous object, the tree can be seen only when light from a source such as the sun or moon falls upon the tree and is reflected. Light will be reflected from, and can be considered as originating from, every point on the tree. It is convenient to treat light originating from each point as rays that travel in straight lines.

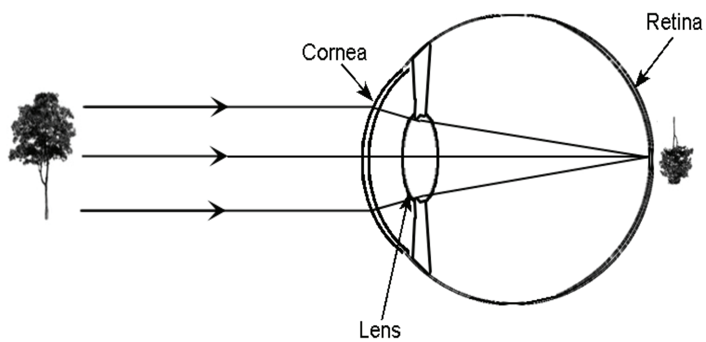


Figure 7-3. Formation of an image.

We need only to concern ourselves with those rays that enter the eye. Due to the nature of optics, we also need only consider a representative number of these rays in order to investigate the image formation process. In Figure 7-3, three rays have been depicted, one each from the many that originate from the top, middle and bottom of the tree.

In our basic model, the eye uses a simple lens system (cornea plus the lens) to form an image of the tree on the retina. In an often-used analogy, the eye is compared to an old-fashion analog camera (Figure 7-4). In this analogy, the retina acts as the film, the lenses of the eye acts as the lenses of the camera, and the iris acts as a diaphragm controlling the amount of light entering the eye-camera. Except for those entering along the optical axis of the eye, the light rays are refracted (bent) by the lenses and focused onto the retina. As rays from all points of the tree are considered, a two-dimensional image of the tree, although inverted, is formed on the retina (Figure 7-3). The brain later turns this image "right way up" in the stages leading up to conscious perception.

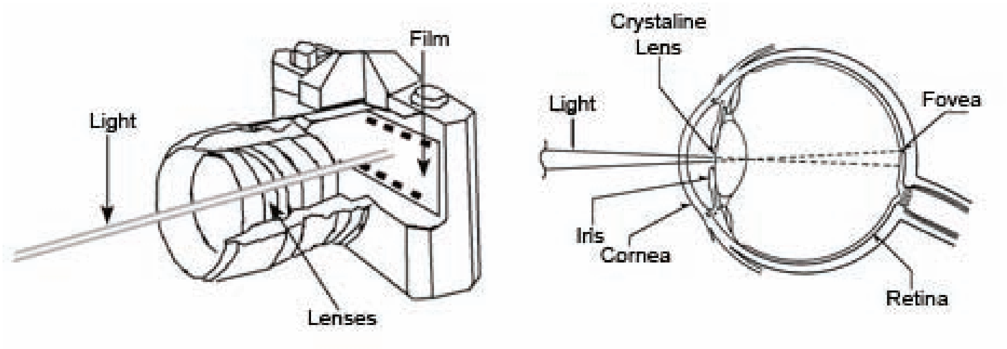


Figure 7-4. The camera-eye analogy.

Note that the eye's optical system includes two lenses, the *cornea*, at the front of the eye, and the lens, inside the eye. The cornea acts as a fixed-focus converging lens, providing about 65% of the focusing power of the eye; the internal lens acts as a variable focus lens. It is controlled by a set of muscles (the ciliary muscles) that relax and contract, thereby changing the lens' curvature and power. This mechanism provides the fine-focusing that allows the cornea-lens system to form a sharp image on the retina over a range of object distances (Atchison and Smith, 2000; Benjamin, 2006; Bennett and Rabbetts, 1991; Goss and West, 2002).

The cornea

The cornea is a thin, transparent tissue at the front of the eye consisting mostly of a collagen-based stromal layer that is about 0.5 mm thick. A thin tear film coats the anterior corneal surface, making it into a smooth high-quality optical surface. Wind, low humidity, high altitude, certain diseases, drugs or refractive surgery can affect the tear layer and lead to corneal surface drying, which causes irritation and transient blurred vision. The cornea is a living tissue and requires oxygen, absorbed directly from the atmosphere, in order to maintain normal metabolism and transparency. Hypoxia due to the environment or contact lens wear can lead to corneal swelling, optical distortion, and loss of transparency. Surface drying and hypoxia can be especially troublesome for pilots or aircrew who wear contact lenses, or who have had refractive surgery. The cornea's refractive power is determined largely by its anterior and posterior surface curvatures. One way to correct refractive errors of the eye is to alter the curvature or shape of the anterior corneal surface, as is done in refractive surgery (Bron, Tripathi and Tripath, 1997; Kaufman and Alm, 2003).

The lens

The internal lens, also called the crystalline lens or simply called the lens, has less than half the refractive power of the cornea, but it fulfills an important unique function. By adjusting its shape it allows the eye to accommodate, that is, focus for different viewing distances. Accommodation declines with age. By about age 45, most people have difficulty focusing at normal reading distances, and need help from bifocals or reading glasses. Opacities of the lens, known as cataracts, may be caused by trauma, disease, toxicity, exposure to radiation, or as a normal process of aging. Depending on the severity and distribution, they can degrade vision. If the visual impact is severe enough, the cataract can be surgically removed and replaced with an artificial intraocular lens (Benjamin, 2006; Goss and West, 2002).

The pupil

The pupil, the aperture at the center of the iris, controls the amount of light entering the eye by changing size in response to light. The pupil's diameter is usually close to 4 mm, but in dim illumination it can dilate to about 7 mm, and in bright illumination it constricts to about 2 mm. Retinal illumination, in trolands (E'), may be computed by the following formula, where object luminance (L) is expressed in candelas/m^2 and pupil area (A) is given in square mm (Schwartz, 2004).

$$E' = LA \qquad \text{Equation 7-2}$$

Since pupil area changes with the square of its radius, retinal illumination also changes with the square of pupil radius. Pupil size varies somewhat from person to person, and with age, race, distance of the object being viewed, emotional state, fatigue and in response to certain drugs. Pupil size also affects retinal image quality. A small pupil increases the eye's depth of focus and minimizes the affect of small optical errors. For example, following LASIK (laser in-situ keratomileusis) refractive surgery, patients with small residual optical aberrations may see well during the day when illumination is high and the pupils are small. At night, however as the pupils naturally dilate, residual aberrations may degrade vision noticeably. Another example is an aviator aged 40-50, who may be able to read without bifocals in high illumination when the pupil is small, but who may have difficulty reading in low light.

The retina

The retina is an intricate tissue layer that contains 10 distinct sub-layers, over 100 million photoreceptor cells and complex neural networks that process the image. It is about 0.5 mm thick and lines the back half of the eyeball's interior, and so receives the extended image formed by the cornea and lens. If you examine the retina using an ophthalmoscope, it will appear as a red surface, due to its rich blood supply, with a prominent pale oval, on the nasal side, which is the optic nerve head (Figure 7-5). Seen emerging from the optic nerve are the retinal arteries and veins. On the temporal side of the optic nerve is a slightly darker region, known as the macula, and at the center of the macula is a tiny, but critically important area called the fovea. The fovea corresponds to the central 2° of the visual field, and because of its extremely high photoreceptor cell density, it supports the best visual acuity in the retina. The fovea is the most important part of the retina since it provides high-definition vision and is the focus of our visual attention. While damage to other areas of the retina may go unnoticed, damage to the fovea causes a debilitating loss of vision in that eye (Bron et al., 1997; Kaufman and Alm, 2003; Schwartz, 2004).

Refractive Errors and Their Correction

The most common cause of poor vision is an uncorrected refractive error. Ideally, the cornea and lens focus the optical image precisely onto the retina, but when refractive errors are present, the lens-to-retina focal distance is incorrect and the image is blurred.

Lower-order aberrations (defocus and astigmatism)

The largest refractive aberrations in the normal eye are defocus and astigmatism. These are sometimes referred to as the lower-order aberrations. Defocus includes the common refractive errors of *myopia* (near sightedness) and *hyperopia* (far sightedness). An eye that has no lower-order aberrations and therefore no refractive error is considered emmetropic (Figure 7-6a). In the case of myopia, the image comes to focus in front of the retina (Figure 7-6b). Distant objects are blurred for patients with myopia. In hyperopia the focal plane is behind the retina (Figure 7-6c). Depending on the degree of hyperopia, hyperopes usually have more difficulty focusing on near objects. Astigmatism is a condition in which some of the eye's optical surfaces are curved like the side of an American football with greater curvature in one meridian (vertical) and a lesser curvature 90° away (horizontal). As a result, light in the eye forms a linear focus at one distance, a perpendicular linear focus at some greater distance and a blurred interval in between (Figure 7-6d), causing blurred vision for both far and near objects. The simplest refractive errors, such as myopia or hyperopia can be compensated in optical instruments such as binoculars or night vision goggles (NVGs) by adjusting the instrument's focusing ring. Astigmatism however, is more complex and requires customized correction with spectacles or contact lenses (Benjamin, 2006; Goss and West, 2002).

Figure 7-5. Photograph of a normal retina. This is what you would see if you looked into someone's right eye. The nose is to the right of the picture, the temple to the left (Copyright NSU Oklahoma College of Optometry; used with permission).

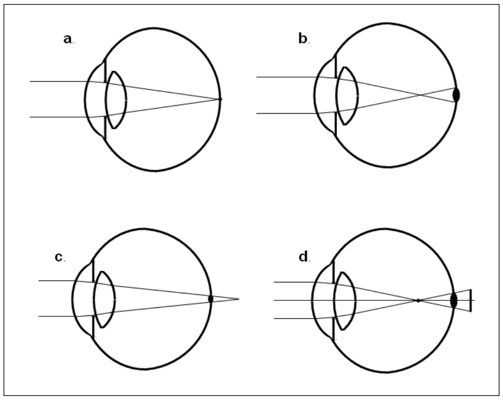


Figure 7-6. Refractive errors. In emmetropia (a) light focuses onto the retina. In myopia (b) light over-converges and forms a blur circle on the retina. In hyperopia (c) light under-converges and forms a blur circle on the retina. In astigmatism (d) some light over-converges, while some under-converges, resulting in a blur circle on the retina.

Higher-order aberrations

Higher-order aberrations are refractive errors that are more complex than myopia, hyperopia or astigmatism, and cannot be well corrected with conventional spectacles or contact lenses (Atchison and Scott, 2000; Campbell et al., 2004; Salmon and van de Pol, 2005). Fortunately, in most normal eyes, they are small and have little noticeable effect on vision (Thibos et al., 2002). The most common aberrations in the normal eye are coma, trefoil and spherical aberration (Salmon and van de Pol, 2006). When the lower-order aberrations are fully corrected, the presence of higher-order aberrations, along with light scatter in the eye cause symptoms of halos, glare, and reduced contrast sensitivity in some eyes (Chalita and Krueger, 2004; Mrochen and Semchishen, 2003).

Chromatic aberration

Refraction, which is the bending of light used by lenses to focus light, varies according to wavelength and is proportional to the wavelength. Therefore, any optical system using white light will be in focus for only one wavelength, while other wavelengths will be out of focus. This wavelength-dependent focusing discrepancy is referred to as chromatic aberration. In the case of the human eye, the focus difference between the shortest and longest wavelengths (longitudinal chromatic aberration) amounts to about 2 diopters (Thibos et al., 1990; Thibos, Bradley and Zhang, 1991). If the eye's optics form an in-focus retinal image for 555-nm light, slightly blurred, out-of-focus images from the other wavelengths will be superimposed. The net result will be a slightly more blurred image in white light, than in monochromatic (single wavelength) light. Fortunately, the eye's sensitivity to different wavelengths is biased toward middle wavelengths (see the CIE [Commission Internationale de l'Eclairage or International Commission on Illumination] V̩[λ]) function section below), and this significantly diminishes the adverse blur caused by chromatic aberration (Bradley, 1992). In addition, the lateral magnification of extended objects, or the location of peripheral objects, imaged on the retina, will vary with wavelength, and this can contribute to blur, especially in the peripheral retina. Chromatic aberration is not an issue for monochromatic displays or optical systems, but should be considered in any system that uses multiple wavelengths (colors) or white light. Chromatic aberration can also arise when optical instruments are not correctly centered relative to the eyes.

Spectacles and contact lenses

The most common way to correct refractive error is through the use of spectacles or contact lenses. In order to correct myopia, the correcting lens has to increase the divergence of light entering the eye, which effectively pushes the focus of the system back towards the retina. Myopia-correcting lenses, which diverge light, have a negative focal power and are referred to as *minus* lenses. For hyperopia the correcting lens must increase convergence of light such that the focus of the system is pulled forward towards the retina. Hyperopia-correcting lenses increase convergence and have positive focal power. They are therefore known as a *plus* lens. To correct astigmatism, a *cylinder* lens is used that has max power in one meridian and minimum power in the perpendicular meridian. This lens must be correctly oriented with the axis, which is the orientation of the eye's astigmatic refractive error (Benjamin, 2006).

Most eyes have a combination of defocus and astigmatism, so spectacle and contact lens corrections may contain correction for both kinds of refractive errors. Spectacles place this correction about 10 to 15 mm in front of the eye, whereas contact lenses are placed directly onto the cornea. There are advantages and disadvantages to each of these corrections, however in terms of compatibility with most head-mounted displays, contact lenses have the distinct advantage of providing a more unencumbered visual correction. That is not to say that contact lenses are the perfect solution since there are increased risks of eye infections and ocular discomfort with contact lenses, especially under austere or harsh environmental conditions (Benjamin, 2006; Bennett and Weissman, 2005).

Except for some recent developments in spectacle lens design, most spectacle lenses correct only lower-order aberrations. Besides correcting the lower-order aberrations of defocus and astigmatism, some aspheric contact lenses may correct spherical aberration (a higher-order aberration) for some patients. Currently the only contact lens type that can correct most higher-order aberrations of the cornea is a rigid contact lens. It covers and in effect, replaces the cornea as the anterior refractive surface of the eye. However, if higher-order aberrations are present in the intraocular components (posterior cornea and internal lens), these would not be corrected by rigid contact lenses. Efforts are under way to develop spectacles and contact lenses that are customized for each person's specific lower and higher-order aberrations.

Refractive surgery

Refractive surgery directly modifies the eye's optics in order to correct refractive errors. This can be accomplished through reshaping of the cornea (keratorefractive), implanting a lens in addition to the eye's natural lens (corneal inlay or phakic intraocular lens) or replacing the eye's natural lens (clear lens extraction). This often frees the patient from the need to wear spectacles. The most common refractive procedures are corneal, such as photorefractive keratectomy (PRK) or laser in-situ keratomileusis (LASIK). These techniques use a laser to reshape the cornea to either increase its power (to correct hyperopia) or decrease its power (to correct myopia). Keratorefractive surgery can also correct for astigmatism. Early forms of keratorefractive surgery often inadvertently increased higher-order aberrations and left patients with poor vision that was uncorrectable with standard spectacles or contact lenses. These aberrations were particularly problematic with large pupils, so they were most noticeable in low light (Bailey et al., 2004; Fan-Paul et al., 2002; Hammond, Puri and Ambati, 2004; Schallhorn et al., 2003; Yamane et al., 2004). Recent improvements in refractive surgery have decreased the risk of residual higher-order aberrations through the use of wavefront-guided customized corrections (Kaiserman et al., 2004; Krueger, Applegate and MacRae, 2004).

Testing the Visual System

A large number of tests are available to evaluate the visual system. They may be divided into: 1) tests of optical performance and 2) tests of visual performance. Clinical tests of optical performance usually measure refractive errors and enable doctors to prescribe the appropriate optical correction to restore a clear focus on the retina. Autorefractors (Figure 7-7) are tabletop instruments that objectively measure myopia, hyperopia and astigmatism, while newer instruments, known as aberrometers (Figure 7-8) measure all of these as well as higher-order aberrations. Optometrists and ophthalmologists have developed methods to determine spectacle prescriptions based on subjective responses from the patient, and subjective techniques are considered more accurate than auto-refraction or aberrometry for measuring myopia, hyperopia or astigmatism. However, aberrometers provide the only practical way to measure higher-order aberrations in a clinical setting (Salmon and van de Pol, 2005).

These tests measure only the optical portion of the visual system, while visual tests measure the performance of the entire system; that is the end result of both optics and neural processing. The most familiar visual tests measure visual acuity, contrast sensitivity, the visual field and color vision.

Spatial Vision

Spatial visual performance is defined here as how well we see static monochromatic images, while motion (temporal vision), color and depth perception will be considered separately.

Figure 7-7. Example of a clinical autorefractor. Figure 7-8. Example of a clinical aberrometer

Resolution visual acuity (visual acuity)

The most familiar spatial vision test is visual acuity, which uses the Snellen letter chart found in most doctors' offices. In order to correctly read a letter, such as an E, the patient must be able to resolve the separation between the strokes of the letter. This kind of visual task is therefore referred to as resolution visual acuity and the smallest gap that a person can resolve between the strokes of a letter is referred to as the minimum angle of resolution (MAR). A person with normal vision should be able to resolve a letter E with a 1.0-arc minute MAR. A standard Snellen acuity letter with a 1.0-arc minute MAR and height of 5.0 arc minutes (Figure 7-9) is 8.7 mm tall if the viewing distance is 6 meters (approximately 20 feet). If a person can read letters of this size, he is said to have a visual acuity of 20/20 in the United States, 6/6 in the United Kingdom, or 1.0 in many other countries. If the patient has worse-than-normal visual acuity, he will require larger letters.

If for example, the smallest letter he can read has an MAR of 10.0 arc minutes, which is ten times as large as a 20/20 letter, his visual acuity would be recorded as 20/200, 6/60 or 0.1. (Kaufman and Alm, 2003; Schwartz, 2004)

Figure 7-9. MAR and angular dimensions of a Snellen 20/20 letter E.

Table 7-1 lists different ways of recording equivalent visual acuities. Most visual clinical visual acuity charts use black letters on a white background (high contrast). In some cases subtle changes in vision may be detected more easily if the chart uses low contrast gray letters since low contrast is more difficult to see. Other special purpose charts may use only a limited set of letters, symbols or shapes to test visual acuity.

Contrast sensitivity

Contrast sensitivity provides a more comprehensive test of spatial vision than visual acuity. In a contrast sensitivity test, the patient views test patterns such as letters or stripes that vary not only in size and in contrast as well. Figure 7-10 shows one example of a contrast sensitivity chart with vertical stripes arranged in rows. On this chart contrast decreases from left to right, while stripe size decreases from top to bottom. Although letters, stripes, or any other pattern could be used to test spatial vision, there are theoretical advantages to using gradient stripe patterns with transverse brightness profiles that change sinusoidally. These sine-wave grating patterns (Figure 7-10) are frequently used in vision research (Nadler, 1990; Schwartz, 2004).

Table 7-1.
Different ways to record equivalent visual acuities.

MAR	0.75	1.0	2.0	10.0
log(MAR)	-0.125	0	0.30	1.00
US Snellen	20/15	20/20	20/40	20/100
UK Snellen	6/4.5	6/6	6/12	6/60
Decimal	1.33	1.0	0.5	0.1
Note: The left column shows the best visual acuity scores. The shaded column indicates the standard for normal well-corrected vision, and the two right columns indicate worse-than normal vision.				

.

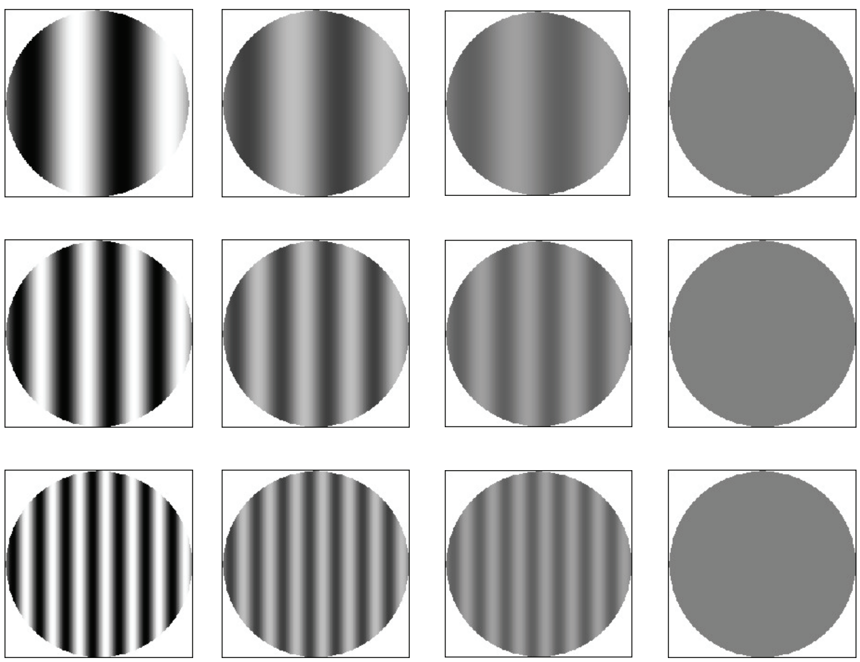


Figure 7-10. Clinical contrast sensitivity charts use targets such as these to test vision. In this simple chart, spatial frequency varies from top to bottom with 2, 4 and 7 cycles in Rows 1, 2 and 3 respectively. Contrast decreases from left to right, with approximate values of 1.0, 0.5, 0.25 and 0 in Columns 1, 2, 3 and 4, respectively. Actual clinical test charts usually include more spatial frequency and more contrast levels. They are designed to measure the minimum contrast a person can see for each spatial frequency.

The two key parameters of a sine wave grating that affect its visibility are stripe width and contrast. Stripe width is specified by the number of repeating light/dark cycles per degree of visual angle, as seen from the eye. Broad stripes have fewer cycles per degree, and are therefore said to have a low spatial frequency. Narrow stripes have more cycles per degree, or a higher spatial frequency. Low spatial frequency gratings (broad stripes) test how well we see large objects, while high spatial frequency gratings (narrow stripes) test how well we see small objects. A contrast sensitivity chart includes gratings with a range of spatial frequencies representing the range of sizes visible to a normal human eye. A high-contrast 30-cycles-per-degree grating corresponds in size to the 20/20 letter on a Snellen eye chart, and should be readable for a person with normal vision.

Contrast is the other key parameter that affects visibility. Low contrast is always more difficult to see than high contrast. Vision scientists define contrast according to the Michelson formula (Equation 7-3), where variables L_{max} and L_{min} refer, respectively, to the luminances of the brightest and darkest portions of the test pattern. Michelson contrast has a maximum value of 1.0, which is the contrast of a black object against a pure white background in the case of a typical visual acuity chart, or the contrast of bright green symbology on a black background as in an aircraft heads-up display. The minimum contrast value is 0, which is the contrast of a gray letter against an equal-luminance gray background – that is, a uniform gray field.

$$C = (L_{max}-L_{min})/(L_{max}+L_{min}) \qquad \text{Equation 7-3}$$

In contrast sensitivity testing, we determine the minimum contrast (contrast threshold) a person can see across a range of spatial frequencies (sizes). A person with good vision is capable of seeing low contrast, and would have a low contrast threshold. A high threshold indicates poor vision. Contrast sensitivity is computed as the inverse of the contrast threshold. High contrast sensitivity (low threshold) indicates good vision, and low contrast sensitivity indicates poor vision. Figure 7-11 presents a typical contrast sensitivity function. It peaks at about 4 cycles per degree, and drops offs on either side. On the high frequency side, the curve steadily declines to zero. The spatial frequency at that point is referred to as the cut-off frequency, and represents the resolution limit of that visual system. A person with excellent vision would be able to resolve 40-60 cycles per degree.

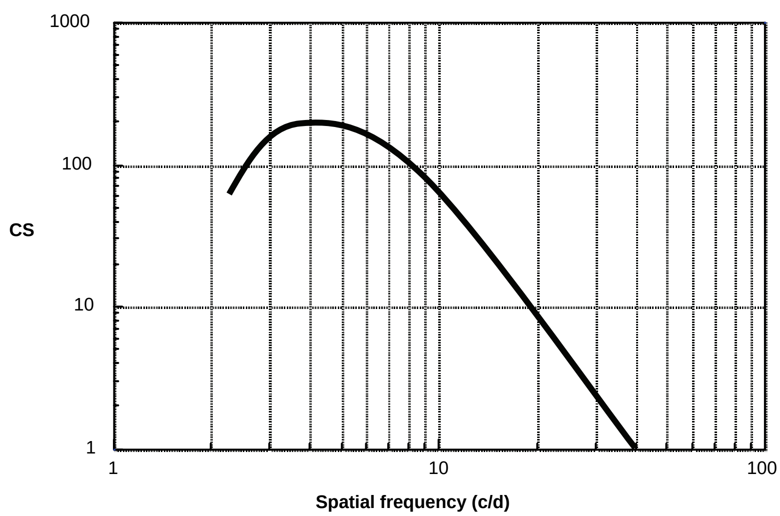


Figure 7-11. A typical contrast sensitivity function.

New technologies are making it possible to correct optical errors more perfectly than ever before. This raises an interesting question: “How well could a person see if he had perfect optics?” Theoretically, with a large pupil, the eye’s optics should be capable of imaging approximately 200-cycle-per-degree patterns onto the retina, (Atchison and Smith, 2000) but because of the size of retinal photoreceptors cells, the retina cannot resolve spatial frequencies greater than about 75 cycles per degree (Applegate, Thibos and Hilmantel, 2001). This corresponds to a Snellen visual acuity of 20/8 (American notation), which would be four rows better than 20/20 on a standard chart.

Although most visual scenes contain a complex mix of spatial frequencies and contrasts, the contrast sensitivity function (CSF) characterizes the basic spatial vision capabilities of the visual system by testing at discrete spatial frequencies. In some respects it resembles the modulation transfer function (MTF) used in optical engineering, however it differs from an optical MTF because the CSF also takes into account neural processing by the brain. Various optical or pathological problems can affect vision, and they can affect different aspects of the CSF to different degrees. For example, small refractive errors mainly reduce the CSF in the high spatial frequencies only. This makes small objects more difficult to see, but large objects are unaffected. A cataract or even a dirty helmet visor can cause poor vision across a broader range of spatial frequencies. Cockpit instruments, especially those used at night provide high contrast, and are therefore easy to see. However, other important visual information, for example, maps in the cockpit, or outside terrain features, personnel or equipment, may have low contrast, which makes them difficult to see.

Since high spatial frequencies and lower contrasts are harder to see, optical devices can improve vision by decreasing the spatial frequencies of images, increasing contrast or both. Magnification is one way to decrease the spatial frequency of objects. Another simple way to decrease spatial frequency is to move closer to the object. Visibility of computer monitors or cockpit displays can be improved by increasing contrast. Spectacles or contact lenses correct optical blur, which improves contrast at high spatial frequencies thereby making small objects easier to see.

Vernier acuity

Another important spatial visual task is the ability to detect a difference in the relative position of two objects. For example, what is the smallest angular offset of one line, relative to the other that the visual system can detect (Figure 7-12)? This kind of task is referred to as Vernier acuity, and under ideal conditions we are capable of detecting angular offsets as small as 10 arc seconds. This is equivalent to detecting a 1-mm offset at a distance of 20 meters. Because Vernier acuity is so good, it is sometimes called hyperacuity. High precision measuring devices sometimes use tick marks or images that an observer must align, in order to take advantage of Vernier acuity. Vernier acuity is also used by naval aviators to verify the correct glide path during aircraft carrier landings (Figure 7-13). Vernier acuity is also important for aiming weapons since the shooter must aligning the back sight with the front sight and target.

Figure 7-12. Vernier acuity example. Which line is higher?

Peripheral Vision

Unless otherwise specified, we assume that visual acuity, contrast sensitivity and most other vision tests are measuring central or straight-ahead vision. In this case, the object of interest is imaged onto the part of the retina known as the fovea. Although the fovea accounts for only the central 2° of the visual field, it provides the majority of the visual information that occupies our visual. High photoreceptor density in the fovea optimizes resolution and makes 20/20-or-better visual acuity possible. Photoreceptors are more sparely placed in the peripheral retina, where visual acuity is worse, usually in the range of 20/200, which is ten times worse that foveal vision.

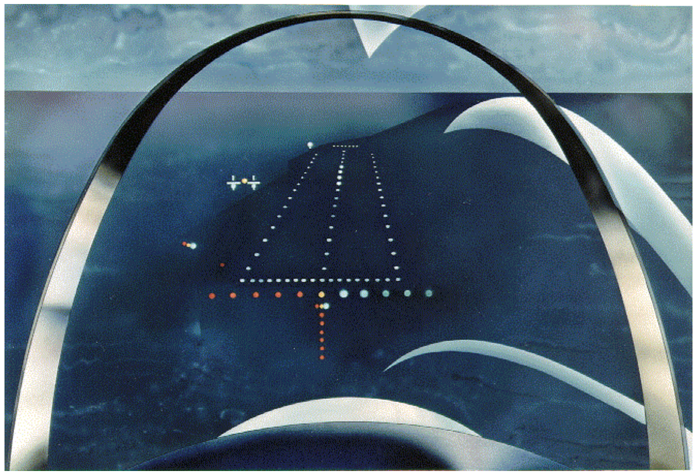


Figure 7-13. Navy pilots judge the alignment of landing lights to verify their glide path during a carrier landing, a Vernier alignment visual task. (Used with permission from NAVAIR Lakehurst; http://www.lakehurst. navy.mil/nlweb/icols.gif)

Although visual acuity declines peripherally, peripheral vision is important, especially for detecting large objects or moving objects. With static straight-ahead gaze, the monocular visual field extends as far out as 50, 60, 70 and 90° in the superior, nasal, inferior and temporal directions, respectively. As shown in Figure 7-14, with both eyes fixating straight ahead, the full extend of the binocular visual field is about 180°. Since each eye's visual field extends 60° nasally beyond straight-ahead gaze, the two monocular visual fields have considerable overlap. This means that objects located within the central 120° are seen by both eyes, thereby giving us the advantages of binocular vision. Objects located in the far periphery to the right and left are seen only by one eye. (Anderson and Patella, 1999)

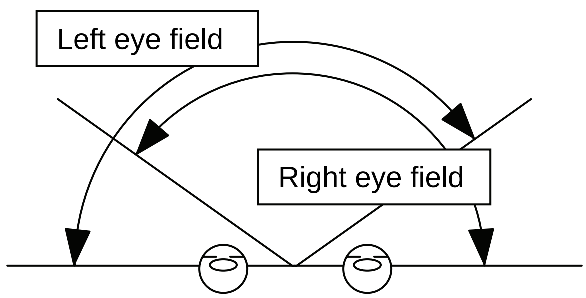


Figure 7-14. Top view of each eye's visual field. The right and left eye visual fields extend from about 90° temporal, across the midline and about 60° into the opposite field. The horizontal extent of each eye's field is about 150°. The central 120° of the two fields overlap. The far temporal periphery of each eye's field is seen by one eye.

Although unobstructed eyes may have the monocular or binocular visual fields limits described above, optical devices, such as NVGs usually restrict vision to a narrower field. However, by scanning with the eyes and moving the head or body, the relatively narrow 40-degree field-of-view of NVGs can cover nearly 360° of visual space. External obstructions such as window frames, seats, or shoulder harnesses that restrict movement are also factors that limit the effective field-of-view.

Although visual resolution is worse in the periphery, some aspects of peripheral vision are better than central vision. We can usually detect motion better in the periphery, and the mid-peripheral retina, about 20° outside the fovea, is actually better than the fovea for seeing faint objects in the dark. With this in mind, personnel who must detect faint objects at night should be taught to look slightly to the side of rather than directly at the objects they are trying to see.

Visual Adaptation to High and Low Illumination

Among the neurons in the retina are two classes of photoreceptor cells, the *rods* and *cones*, which start the visual process by responding to light in the image created by the eye's optics. Using a complex biochemical process known as phototransduction, the photoreceptors convert optical energy into electrophysiological signals that are relayed to the brain. The rods and cone photoreceptor cells differ in terms of their intracellular structure, and range of sensitivity. The presence of these two photoreceptors systems enable the visual system to operate over a wide range of light levels (Kaufman and Alm, 2003; Schwartz, 2004).

Scotopic vision

Rods support vision in low light (*scotopic* vision) and are designed to maximize photon capture. They are absent in the fovea but present in the rest of the retina. Since there are no rods in the fovea, under scotopic conditions, such as at night, the central 2° of the visual field becomes a tiny blind spot. The rods are designed to capture light when photons are sparse, and scientists have found that a rod cell is capable of responding to a single photon of light. Perceptual awareness requires simultaneous absorption of at least 10 photons (Cornsweet, 1970). The scotopic system operates from nearly complete darkness up to luminance values of about 1 candela/m^2. Rods photoreceptors do not contribute to color perception. Because of the distribution of rods and their supporting neurons, scotopic visual acuity is poorer than cone-mediated acuity. On the other hand, the rod system is better at integrating light from a larger area of the retina, and it therefore provides vision in low light, below the threshold for cones. As illumination increases and approaches the upper limit for the rods, the cones begin to work. For intermediate light levels both rods and cones are working. Vision in this range of illumination is known as *mesopic* vision. As illumination increases above the mesopic level, the rods become saturated and cease to function. We then transition to cone-mediated, that is, *photopic* vision. The output of most NVGs is sufficiently bright that the eye's response is in fact cone-mediated, which has implications for vision in unaided areas of the visual field at night (see below).

Photopic vision

Cones are present throughout the retina, but are most highly concentrated in the fovea and more sparsely distributed in the periphery. At the center of the fovea, cone density is about 120,000 cells/mm^2 (Bron et al., 1997), and provides distinct input to the visual center in the brain. This allows for high-resolution vision of at least 20/20. In the peripheral retina, cone density decreases, and input from cones is pooled. This limits visual resolution and visual acuity to about 20/200. In addition to providing high-resolution central vision, the cone system supports color vision. Because of their structure, cones capture light most effectively if rays enter straight on, that is, perpendicularly to the retinal surface. Light rays striking at wider angles stimulate the cones less efficiently, a phenomenon known as the Stiles-Crawford effect. Because of this, light rays entering the peripheral pupil appear less bright than rays entering centrally. A benefit of the Stiles-Crawford effect is that light scattered within the eye has little effect on cone vision. Eventually illumination becomes so high that the cones become saturated and vision fails.

It is interesting to note that personnel using NVGs may be using mesopic and photopic vision for the central 40° of their visual field, while depending on scotopic vision to scan for objects in the far periphery.

The CIE V(λ) function

Both rods and cones are sensitive to a wide range of wavelengths, but sensitivity varies for different wavelengths (Figure 7-15). Both rod and cone sensitivities peak near the middle of their respective ranges. In terms of absolute sensitivity, the rods are more sensitive than the cones. Figure 7-15 also shows that the scotopic (rod) sensitivity spectrum peaks at shorter wavelengths than the photopic (cone) sensitivity spectrum. Because of this, as illumination decreases and we transition from photopic to scotopic vision, we may perceive that shorter wavelength hues become relatively brighter while longer wavelength hue become less bright, a phenomenon known as the Purkinje shift. The photopic curve in Figure 7-15 is referred to as the CIE luminous efficiency curve, or V(λ) function (Stockman et al., 2006). It defines the how efficiently the each wavelength stimulates the vision of a normal human observer, and is therefore foundational for the field of photometry. Photometry defines standard units for illumination and luminance, which simply put, specifies perceived light intensity. Illumination quantifies the amount of light falling on an area. A basic metric unit for illumination is the lux (lumens/m^2); the English unit is the foot-candle (lumens/ft^2). Luminance quantifies the amount of light emitted from an extended source. A basic metric unit for luminance is the nit (candelas/m^2); an English unit is the foot-Lambert ($1/\pi$ candelas/ft^2) (Schwartz, 2004).

Dark adaptation

Within their working ranges, rods and cones must adapt to changes in light level. When illumination increases (light adaptation), the photoreceptors become less sensitive to light. When illumination decreases (dark adaptation), they become more sensitive. Cones dark adapt more quickly than rods and reach their maximum sensitivity after about 15 minutes in the dark. Rods, on the other hand require about 40 minutes to fully dark adapt and reach maximum sensitivity. After complete dark adaptation, exposure to any light begins the process of light adaptation and visual sensitivity declines. If a dark-adapted person needs to use a light, yet hopes to preserve dark adaptation, the loss of sensitivity can be minimized by using a long-wavelength red light. Long wavelengths are relatively poorly absorbed by rods, so red light has minimal impact on rod dark adaptation. Meanwhile, cones are about equally sensitive to rods at long wavelengths, so they can contribute to vision in low light (Kaufman and Alm, 2003; Schwartz, 2004).

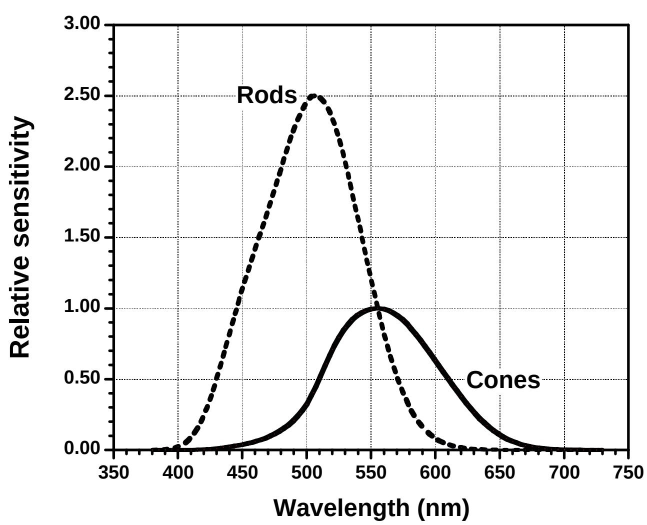


Figure 7-15. Relative sensitivity of the rod (scotopic) and cone (photopic) systems as a function of wavelength.

Color Vision

Color is a critical feature of vision since it helps us better discriminate objects, such as different teams, medical test results or cockpit data. Color describes the sensation created by the visual system primarily based on the wavelength absorbed by the cones. Someone once said, “Color is not a property that inheres in external objects but is rather an internal construct of the individual, dependent on the wavelength composition of light entering the eye and on the structure of the eye and nervous system” (Swanson and Cohen, 2003). Therefore, although different wavelengths of light exist in the physical world, color exists only in the mind of the beholder.

Hue, saturation, brightness

As was mentioned earlier in this chapter, the human visual system is sensitive to electromagnetic radiation with wavelengths between approximately 700 and 400 nm, which corresponds to hues ranging from red, orange, yellow, green, blue and violet, which is near 400 nm. The three basic characteristics of a color are its *hue*, *saturation*, and *brightness*. Hue refers to the aspect of color that most obviously distinguishes one wavelength from another, and is often used synonymously with the word “color.” For example, “red,” “green” and “blue” refer to different hues. However any hue can vary in appearance because of differences in color saturation, which describes how pure or vivid a particular hue is. For example, a highly saturated version of red looks deep red while a desaturated version looks pink. Two colors with the same hue and saturation can also look different due to differences in brightness.

L, M and S cones

Color perception is based on the ability to discriminate wavelengths, and this is possible because the retina contains three types of cone photoreceptors, each of which responds to a different range of wavelengths. Figure 7-16 shows an example of how the three cones types, known as S, M and L cones, are distributed in a portion of one person’s retina. (Roorda et al., 2001; Roorda and Williams, 1999) The three classes are designated S, M and L cones because their peak sensitivities are located, respectively, in the short, middle and long wavelength ranges. Although they have peak sensitivities at different wavelengths, each absorbs a broad band of wavelengths across the visible light spectrum, as shown in Figure 7-17. The overlapping across different absorption spectra makes it possible to uniquely encode any wavelength by the ratio of three cone responses. The three cones send their signals into a complex network of neurons within the retina that further process the wavelength and brightness information and send it to the brain, which completes our perception of color (Schwartz, 2004).

The CIE (Commission Internationale de l’Eclairage or International Commission on Illumination) color specification system

Just as our visual system can sense any color based on the response of three cone types, it is possible to create a wide range of colors by mixing three primary colors. For example, TV and computer monitors simulate different colors by mixing red, green and blue colored lights. This is an example of additive color mixing. Subtractive color mixing occurs when pigments rather than lights are mixed to create other colors. Three primaries that are often used for subtractive color mixing, as is done in printing, are yellow, cyan, and magenta.

Different color specification systems have been developed, but the most popular is 1931 CIE color specification system. It matches color by the additive mixture of three primaries known as X, Y and Z. The system has been designed so the relative proportion of each primary in a mix is represented by chromaticity coordinates x, y and z, the sum of which always add up to a value of 1.0 for each wavelength hue. If any color’s x and y coordinates are known, the z value can directly be computed, so the x and y chromaticity coordinates are sufficient to specify any color. By plotting the x and y coordinates for any wavelength hue, we can generate an arc of points that represents every color of the spectrum, as shown in Figure 7-18. This is the CIE chromaticity diagram, and is frequently used in science and engineering to analyze and perform calculations with color. The points along the arc represent all colors of the spectrum, and points within the arc represent all other colors that can be created by any mixture of the spectral hues. The straight border along the bottom of the CIE chromaticity field represents various shades of purple that can be created by mixtures of violet and red. Pure white has chromaticity coordinates x=0.33, y=0.33, z=0.33 since it is made up of equal amounts of the three primary lights (Schwartz, 2004; Stockman et al., 2006). Another widely used color specification system is the CIE Lab system. It may be derived from the CIE standard color designation by transforming the original x, y and z coordinates into three new reference values known as L, a and b. This transformation creates a color coordinate system that better expresses color differences as perceived by a normal eye (Agoston, 1979; McLaren, 1976).

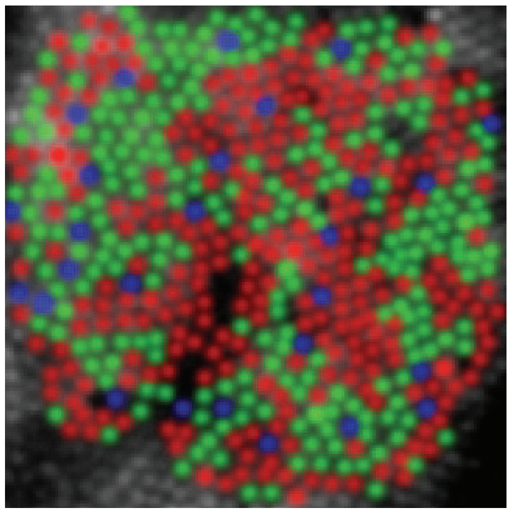


Figure 7-16. Distribution of S, M and L cones in the paracentral retina of one subject. Blue, green and red dots respectively show the locations of S, M and L cones in a 136 x 136 µm (0.5 x 0.5 degree) region on the retina. Calculated from data provided by Austin Roorda, and available in a downloadable Excel spreadsheet at: http://vision.berkeley.edu/ roordalab/

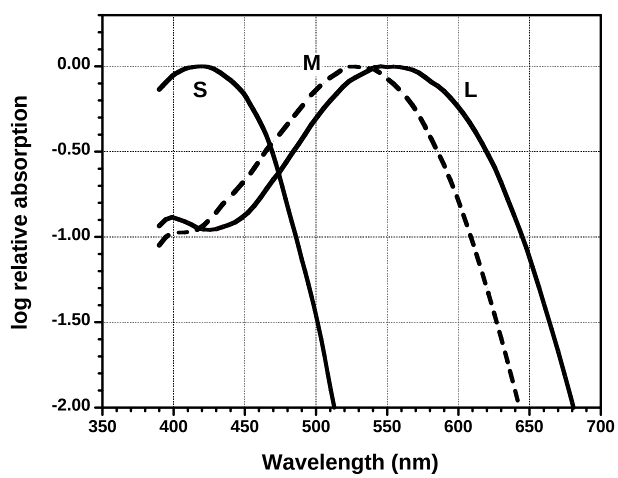


Figure 7-17. Relative absorption spectra of the S, M and L cones based on the Stockman and Sharpe 2000 data set, downloadable from: http://www.cvrl.org/

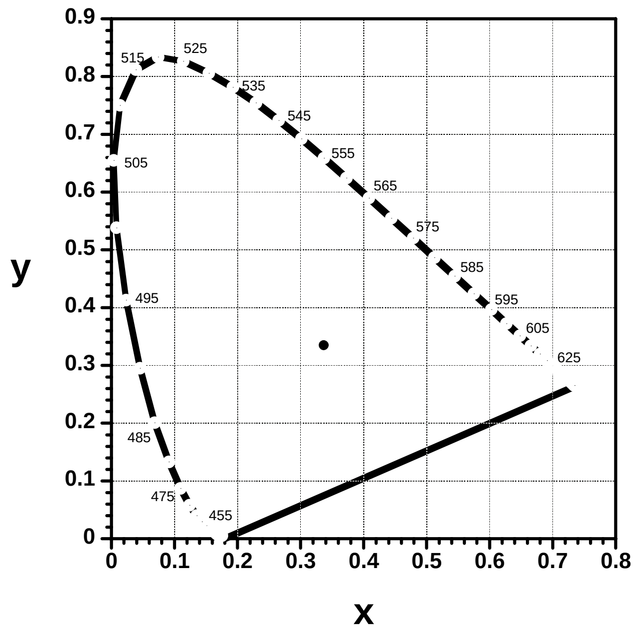


Figure 7-18. The CIE chromaticity diagram. Labels on the curve indicate wavelength in nanometers (nm). The dot indicates the coordinates for pure white.

Color blindness and abnormal color vision

Cataracts, other diseases or toxicity can cause anomalous color vision, but most color vision defects are hereditary. Hereditary color anomalies are classified into three categories based on the cone type that is affected. Patients with problems in the L, M or S cones are respectively referred to as *protans*, *deutans* and *tritans*. Absolute protans have no L cones, while patients with mild protanomalous vision experience a less severe color anomaly due to defective L cones. Similarly M-cone defects are subdivided into deuteranopia, where the M-cones are absent, or deuteranomaly, where they are present but anomalous, and S-cone defects are likewise divided into tritanopia and tritanomaly. Hereditary color vision anomalies affect about 8% of males but only about 0.4% of females. Among them, deuteranomaly is the most common, affecting 5% of males. Both protans and deutans have difficulty discriminating long and middle wavelength colors. This range of wavelengths includes the hues red and green, so both protans and deutans are sometimes referred to as having red-green color anomalies. Interestingly, although color blindness and color anomalies predominantly affect males, men with the defective gene inherit it from their mothers. Since red-green color anomalies are not rare, affecting about 8% of the male populations, engineers who plan to incorporate color into displays or signals should be careful to avoid colors that can be confused by red-green color anomalous patients. These patients have difficulty discriminating hues such as red, orange, yellow and greenish-yellow, but would be able to discriminate red from blue. To ensure that they can distinguish different items, they should use appropriate colors or other visual cues, such as brightness, flicker or different-shapes to avoid confusion.

Many color vision tests diagnose anomalies by presenting colors that confuse certain categories of color-anomalous patients. Since red-green anomalies are the most common, many color vision tests only diagnose protans and deutans from normals, but do not distinguish between them. The Ishihara color vision plates (Birch, 1997) (Figure 7-19) display a colored number embedded in a background made up of another color. The colors are selected from among those that are confused by color anomalous persons. Depending on which colors are used, it's possible to differentially diagnose protans, deutans and tritans. One of the most well-designed and easy to use color vision tests is the HRR test (named after the designers, Hardy, Rand, Rittler,). Like the Ishihara test, it consists of a book of plates with colored figures embedded in a gray background. It can diagnose all three classes of color anomalies as well as grade their severity (Bailey et al., 2004). Some tests, such as the widely used D-15 test, require patients to arrange color samples in a particular order. Another test, sometimes used to screen aviators, the Farnsworth Lantern test (Figure 7-20), presents patients with red, green or white lights, which they must identify. There is no effective cure or treatment for hereditary color vision defects, but patients with color vision anomalies often learn to compensate and may have little problems identifying colored objects in natural environments. They are more likely to make mistakes, however, when viewing man-made signal lights or symbology (Benjamin, 2006; Schwartz, 2004).

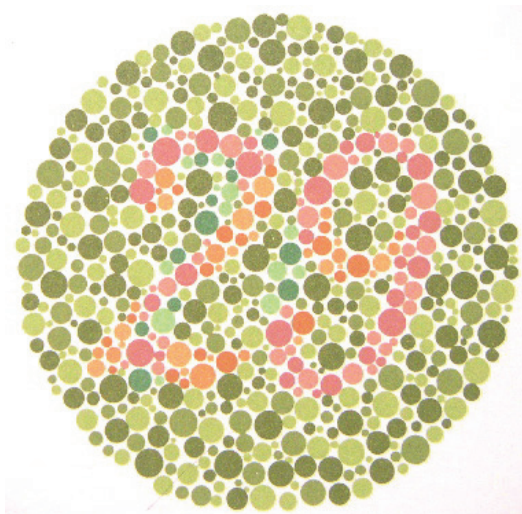


Figure 7-19. One page from the well-known Ishihara color vision test book. People with certain color anomalies have difficulty reading the number (29).

Figure 7-20. In the Farnsworth Lantern test, the patient must identify the color to two lights, each of which may be red, green or white.

Accommodation

Accommodation is the auto focusing mechanism of the eye that allows us to clearly see objects at different distances. By increasing or decreasing its curvature, the eye's internal lens changes the eye's refractive power thereby enabling it to refocus. To see near objects, the eye requires more refractive power, and less power is needed to focus on distant objects. The internal lens is suspended by a system of fibrils that originate from the ciliary body, an annular muscle lining the inside of the eye just behind the iris. In the non-accommodative state, the emmetropic eye is focused for an infinitely distant object. The ciliary muscle is relaxed and relatively flat, pulling outward on the zonular network and lens. This outward radial tension pulls on the lens periphery and flattens the anterior surface, causing less refractive power (Figure 7-21). During accommodation, when the eye focuses at near, the ciliary muscle, a sphincter, contracts, bulges and shortens its internal radius. This releases tension on the zonules, allowing the lens to "bulge" or increase its anterior surface curvature and refractive power (Benjamin, 2006; Goss and West, 2002).

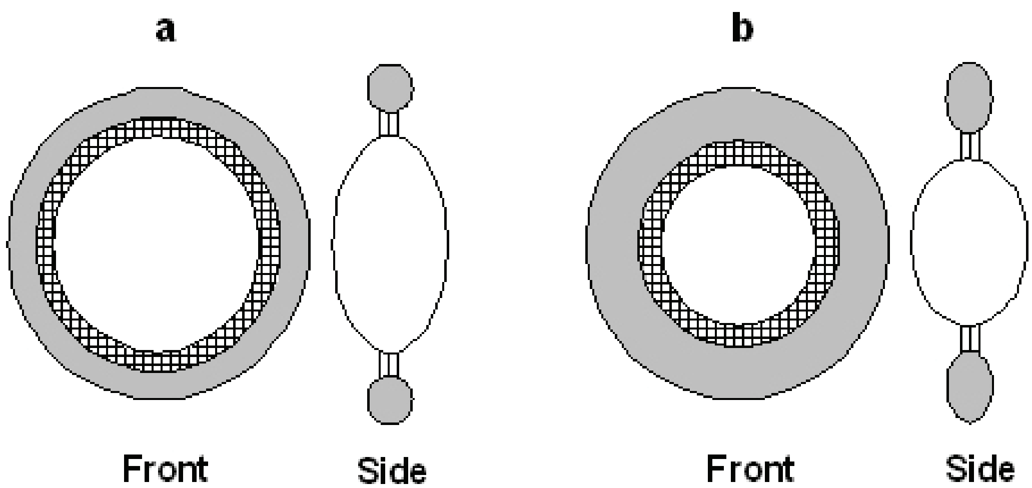


Figure 7-21. Schematic diagram of Helmhotz' theory of accommodation. In the unaccommodated state (a), the ciliary muscle (gray annulus) is relaxed and has a large inner diameter. This pulls the crystalline lens (white) flat, via the connecting zonular fibrils (lines). During accommodation (b), the ciliary muscle, which is a sphincter, flexes and decreases its inner diameter. This allows the crystalline lens to bulge, thereby increasing its focal power.

The accommodative triad

When focusing on near objects, three actions occur simultaneously: the eyes rotate inward (converge), the pupils constrict and the lenses of both eyes increase in power (Benjamin, 2006; Goss, 1995; Goss and West, 2002). Convergence orients both eyes toward the near object of interest so that its image is focused on each eye's fovea. Pupil constriction increases the eye's depth of focus, which improves clarity of near objects. Finally, the lenses of each eye change shape to accommodate; the amount of accommodation is generally symmetrical between the two eyes. These three components of accommodation are known as the accommodative triad or near reflex and are neurologically coupled at the Edinger-Westfal (EW) nucleus in the brain. It is therefore possible to drive lens accommodation and pupil constriction just by converging the eyes. Alternatively a stimulus to accommodation can cause the eyes to converge. This near reflex can work against clear, comfortable vision in binocular or bi-ocular head-mounted displays when there are imbalances in convergence or divergence demands on the eyes due to misalignment of oculars or other optical components. Difficulties can also arise when there is an optical accommodative demand in the system due to excessive minus power in one or both oculars; this can drive the eyes to converge and lead to fatigue and/or double vision.

Stimulus to accommodation

Blur is the primary stimulus for accommodation. Retinal blur stimulates the EW nucleus, which then stimulates accommodation. In the absence of any other information, the eye will increase accommodation to make the image on the retina clear. If increasing accommodation further blurs the image, a feedback loop changes the accommodative response from positive to negative, decreasing accommodation. If this were the only mechanism to determine the direction of accommodation, the eye would constantly search for a focus when regarding objects at different distances (or levels of blur). As noted above, convergence of the eyes also stimulates accommodation, thereby assisting the optically driven accommodative mechanism. Longitudinal chromatic aberration of the eye also contributes to accommodation. Since short wavelengths focus closer to the lens than long wavelengths, the spectral composition of the blur provides directional information for the accommodative response. In head-mounted displays where some of the additional cues to accommodation, such as color and convergence demand, are not present, the accommodative system may be less precise and the visual system may become more fatigued.

Night myopia

If no objects are visible, such as in the dark, or when viewing featureless haze outside a cockpit, the eyes will have no optical input to stimulate accommodation. In these situations, the accommodative system does not relax, but rather tends to accommodate slightly and focus for an intermediate distance of about one meters. This dark focus of accommodation causes temporary myopia (night myopia) and will cause distant objects, should they appear, to be slightly blurred. Some pilots who frequently fly at night may need a slightly more myopic eyeglass prescription for night flying to compensate for night myopia.

Testing accommodation

The amplitude of accommodation can be determined in a number of ways. The simplest technique involves presenting a small target at a comfortable distance from the eye and moving it closer until a clear focus cannot be maintained. The near point of accommodation may be recorded in terms of cm from the eye or converted to optical power in diopters by computing the inverse of the distance in meters. Accommodation can also be measured using divergent (minus) lenses of increasing power placed in front of the eye while a distance target is observed. Since the eye accommodates (adds plus focal power) to counter the minus lens, the amount of accommodation is equivalent to the highest power lens that can be cleared. Similarly, some instruments measure accommodation by determining the near point through translation of a target or through changes in the instrument’s optical power.

Presbyopia

The ability to accommodate gradually decreases with age (Figure 7-22) (Atchison, 1995; Koretz et al., 1989). Young children may have 15 to 20 diopters of accommodation, giving them the ability to see objects clearly as close as 5 cm from the eye. As we age, the lens continues to grow, becoming denser and less flexible. The result is a decreasing ability to focus on near objects so that by age 20, accommodative ability will have declined to approximately 12 diopters so the near point will have receded to about 8 cm from the eye. By age 40, accommodation may have decreased to 4 or 5 diopters causing further regression of the near point to about 20 to 25 cm from the eye. Around age 45, the loss of accommodation reaches the point where reading and other close work becomes difficult without a reading correction for most people. By age 60 the lens becomes inflexible and unable to accommodate. This is complete presbyopia and requires an optical correction such as reading glasses or bifocals with a power in diopters equal to the inverse of the reading distance (e.g. +2.50 diopters to focus on objects 40 cm from the eye) (Benjamin, 2006; Bennett and Rabbetts, 1991; Goss and West, 2002).

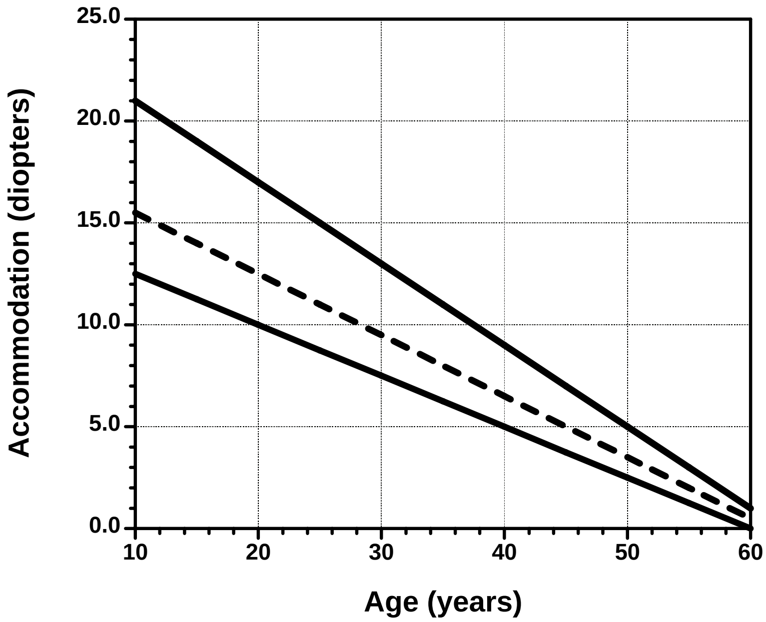


Figure 7-22. Maximum, average and minimum amplitude of accommodation expected with age, according to Hofstetter's formulas.

Presbyopic corrections

Since presbyopia is a condition where distance vision is maintained while the ability to see near is impaired, the goal in correcting presbyopia is to provide additional power for near only. For presbyopic emmetropes (no distance refractive error) or presbyopes wearing contact lenses to correct distance vision, the simplest solution is a pair of reading glasses. The drawbacks to this solution are few, but could be significant depending on the circumstances. Reading glasses provide focus at near or intermediate distances only, therefore distance objects will appear blurry through these lenses. Since the glasses have to be removed to see distant objects, they can be easily misplaced and may not be readily available when needed. Bifocals or multifocals can be a solution for myopes, hyperopes or astigmats who need a distance and near correction and for emmetropes who do not want their distance vision blurred by reading glasses. The top portion of the multifocal lens is designed for distance vision and the lower portion of the lens has either a segment of additional power for near, called an *ADD*, or a gradual increase in power from top to bottom, called a *progressive* multifocal (Benjamin, 2006).

Contact lens corrections for presbyopia include monovision, where one eye is corrected for distance and the other eye is corrected for near, or multifocal, where different zones in the contact lens provide distance or near focus. In monovision, the brain must adapt to one eye being dominant for distance and the other eye being dominant for near. While both eyes contribute to vision at both distances, there is a reduction in visual quality in the eye not focused at the particular distance. In general, about 70% of patients who are fit with monovision successfully adapt to it (Westin, Wick and Harrist, 2000). Multifocal contact lenses are most often designed to provide simultaneous vision, either through the use of discrete distance/near zones or through an aspheric design that gives a continuous, distance-to-near power change over the lens area. When the eye is looking at distant objects, the portion of the lens devoted to distance vision focuses these rays onto the retina; at the same time, near objects will be focused onto the retina by the near optics. Most simultaneous design lenses compromise image contrast due to the “splitting” of focal power for both far and near. A translating multifocal contact lens is another solution. It has discrete far- and near-focus zones that shift as the wearer changes gaze from far (up gaze) to near (down gaze). The lens rests on the lower lid and slides upward when the eyes look downward, thereby centering the near zone over the pupil. Translating lenses work best as rigid gas permeable (hard) lenses, rather than as soft contact lenses and are therefore less popular (Bennett and Weissman, 2005).

Surgical procedures to correct presbyopia are on the increase. PRK or LASIK can be used to create a monovision correction by leaving one eye slightly near-sighted and correcting the other eye for distance. These procedures can also be used to create an aspheric corneal surface that works much like an aspheric contact lens that provides simultaneous distance and near vision. Intra-corneal inlays are in development and include small lenses that provide a central “reading” power as well as small aperture devices that increase the eye’s depth of focus. Intraocular procedures include small accommodating lenses that are surgically implanted inside the eye. Two designs include intraocular lenses that shift forward to increase the near focus of the eye, and multifocal lenses that provide simultaneous vision. The intraocular lenses require removal of the eye’s internal lens and are usually reserved for patients undergoing cataract surgery (Krueger et al., 2004).

The Eye’s Temporal Responsiveness

Although it is often studied as a separate topic, the temporal response of the visual system is strongly influenced with the luminance, spatial, color and surrounding aspects of a stimulus and whether it is located in the central or peripheral visual field. Temporal considerations are important even for a task as simple as light detection, because detection requires that sufficient light be collected over time, a phenomenon known as temporal summation. Temporal resolution, that is, the ability to resolve two visual stimuli separated by time as two, is another fundamental aspect of temporal vision. Flickering lights, which are simply a series of repeat presentations, are also used to study temporal vision. If the rate of flicker is high enough, the visual system will no longer be able to resolve the individual flashes and will perceive a steady light. The rate at which the flicker fuses into a steady perception is known as the critical flicker fusion (CFF) frequency. Temporal contrast sensitivity is determined by changing the contrast level of targets and then determining the CFF for each level. When you combine temporal contrast sensitivity with spatial contrast sensitivity you obtain a fairly complete representation of the limits of the human visual system (Van Hateren, 1993). Motion processing is a special case of temporal vision where the spatial position of the object changes with time, either due to movement of the object across the field of regard or movement of the observer’s eyes or head with respect to the object, which causes the image to move on the retina (Kaufman and Alm, 2003; Schwartz, 2004).

Temporal summation and the critical duration

When attempting to detect a dim light, there is an inverse relationship between stimulus intensity (brightness per unit time) and its duration up to a critical duration. That is, a dimmer light must be left on longer in order to be seen, while brighter lights can be detection with shorter durations. This is referred to as time-intensity reciprocity and is described by *Bloch’s law*

$$Bt = K \qquad \text{Equation 7-4}$$

where B = luminance, t = duration, and K = a constant value.

If the time the stimulus is left on exceeds the critical duration, the intensity required for detection remains constant. This relationship is depicted schematically in Figure 7-23.

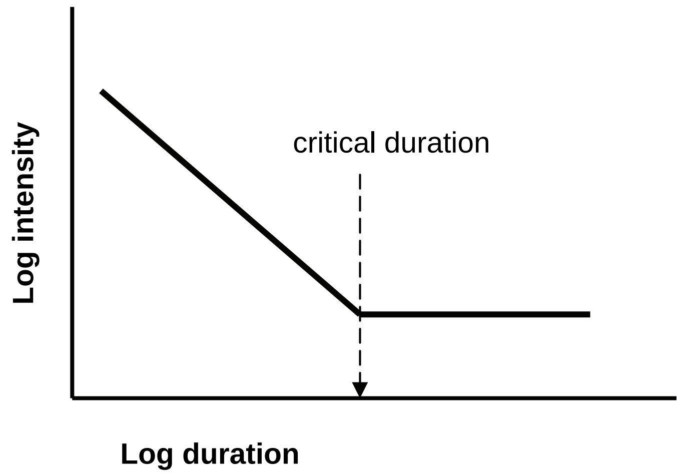


Figure 7-23. Schematic diagram of Bloch's Law, also known as time-intensity reciprocity.

The critical duration is generally between 40 to 100 milliseconds, but it depends on retinal adaptation level, location in the visual field, and color. For instance, the critical duration increases for under scotopic (dark) conditions, for small stimuli, or for single wavelength or narrow-band stimuli.

Critical flicker fusion frequency (CFF)

The CFF is particularly important to display technology, where the refresh rate must exceed the CFF in order to avoid the perception of flicker, which can be annoying. Refresh rate is not the only factor that influences CFF, however. The brightness, color, location in the visual field, and size of the stimulus, as well as variability of the individual, also play a role.

CFF increases as the stimuli becomes brighter; a brighter display would therefore, require a faster refresh rate to avoid the perception of flicker. Based on the same principle, someone who can detect and is annoyed by the flicker of a bright display can eliminate the flicker by reducing the brightness. This phenomenon is described by the Ferry-Porter law, which states that the CFF increases linearly with the log of the luminance. The Ferry-Porter law holds for stimuli of different wavelengths, visual field location and size; although the slopes vary, as described below.

Since the CFF depends on processing speed of the retinal photoreceptor (e.g. faster photoreceptors can perceive flicker at higher frequencies), the type of receptor will influence the CFF. For colored stimuli, the three receptors are the S, M and L cones, which have respective peak sensitivities in the blue, green and red wavelength ranges. When the CFF is measured for these three wavelengths, the slope of the Ferry-Porter function is steepest for the middle-wavelength stimuli, indicating that the M cones are the fastest processors and are more sensitive to flicker with increasing luminance. The function is shallowest for the short-wavelength stimuli, indicating the S cones are slowest and least sensitive to flicker.

The peripheral retina is more sensitive to flicker than the central retina. This can be observed by noting that the flicker of a fluorescent light is more noticeable when viewed with peripheral vision rather straight on. The mid-peripheral retina is the most sensitive to flicker and beyond about 60° from fixation, flicker sensitivity declines.

The Granit-Harper law states that the CFF increases linearly with the log of stimulus area. This applies for retinal eccentricities out to 10° and stimulus sizes up to 50°. Some of this relationship is driven not so much by stimulus size, but by the most temporally-sensitive portion of the retina within the stimulus. For instance, any stimulus that falls even partly on the more sensitive mid-peripheral retina will result in an increased CFF.

Temporal Contrast Sensitivity

Similar to spatial contrast sensitivity, the visual system is variably sensitive to flicker at different frequencies and amplitudes. At lower temporal frequencies (slower on/off), the visual system is fairly consistently sensitive to flicker across a wide range of retinal adaptation levels. As the flicker frequency increases, the visual system becomes more sensitive up to a peak frequency, which under normal photopic levels is around 15 to 20 cycles per second. Figure 7-24 shows how the sensitivity varies as a function of temporal frequency, that is, flicker rate. Beyond the peak frequency, sensitivity declines, and the point where the function intersects the x-axis is the high-contrast temporal resolution limit.

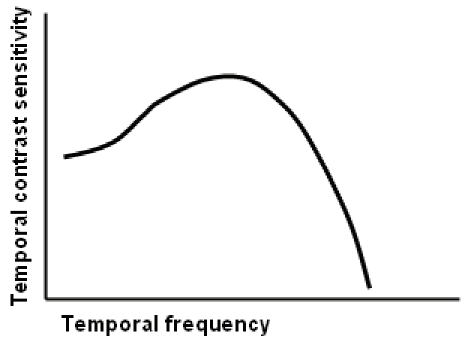


Figure 7-24. Schematic diagram showing how temporal contrast sensitivity varies as a function of temporal frequency.

Eye Movements

A complex system of nerves and muscles work together to coordinate binocular eye movements, with the overall goal to keep the fovea of each eye aimed at the object of regard. The six extraocular muscles (Figure 7-25) are controlled by three cranial nerves (III, IV and VI) and as three agonist/antagonist pairs they serve to move the eyes horizontally (lateral and medial rectus muscles), vertically (superior and inferior rectus muscles), and torsionally (superior and inferior oblique muscles). Two intraocular muscles (iris sphincter and ciliary muscle) are controlled by the coulometer nerve (III) and serve to manage pupil size and accommodative state. The iris dilator muscle is stimulated by sympathetic neurons in the long ciliary nerves (Netter, 1975).

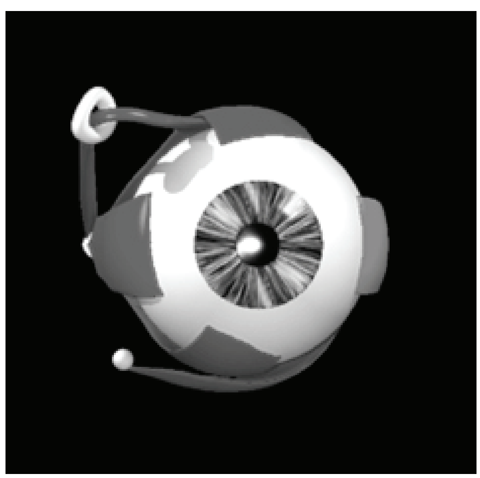


Figure 7-25. Front view of the left eye showing the six extraocular muscles that move the eyeball (Courtesy of Dr. Jason Ellen).

Conjugate eye movements

In conjugate eye movements or *versions*, the eyes move in the same direction. These movements are used for tracking an object that is moving across the visual field (pursuits) or to move quickly towards an object in another part of the visual field (*saccades*). Under the slower pursuit movements, the velocity of the eye is approximately 20° to 50° per second, whereas under faster saccades the velocity of the eye is between 300° to 700° per second. It is important to note that during saccades visual input is momentarily suppressed. This *saccadic suppression* minimizes visual distortion that would occur due to images that rapidly move across the retina. Saccades can either be voluntary, where individuals move their eyes to an object of interest, or involuntary, where the individuals move their eyes in response to an external stimulus, which could be visual, auditory or sensory (pain).

The slow conjugate eye movements include smooth pursuits and the vestibular-ocular reflex (VOR). As previously noted, smooth pursuits are designed to stabilize the image of a moving target on the retina. If the target speed exceeds 50° per second the eyes will start to combine saccades with small intervals of pursuits in order to maintain fixation. This can be demonstrated by moving an object from left to right across the visual field; when the object moves slowly, the eyes follow without saccades, however as speed increases, the eyes are less able to maintain fixation using only smooth pursuits. The VOR helps to keep the eyes on target when the head moves. As the head is rotated, the semi-circular canals in the vestibular system are stimulated, and the information is transmitted to the oculomotor system, which allows the eyes to maintain fixation on the object of regard.

Vergence eye movements

In vergence eye movements, the eyes move in opposite directions; both eyes move towards the midline during *convergence* and away from the midline during divergence. Just as in conjugate eye movements, the primary purpose of vergence eye movements is to keep the foveas of both eyes fixated on the object of regard. As objects come closer to the eyes, the visual axes of the eyes must converge. To accomplish this, the medial rectus muscles of both eyes are engaged. As objects move further away, the opposite occurs; the visual axes diverge due to the action of the lateral rectus muscles of both eyes.

Binocular Vision

Two eyes provide certain advantages over vision with just one eye. For most visual functions such as visual acuity, contrast sensitivity, and extent of the visual field, binocular vision enhances monocular vision. The most significant benefit of binocular vision is stereopsis, which is the powerful perception of depth that is based on the fact that the two eyes view objects from slightly different positions. There are some disadvantages to binocular vision compared to monocular vision. It requires more complex control and processing and thereby renders the person susceptible to problems when the system fails. For example, if the eyes do not align properly the patient may experience double vision and confusion, problems that normally cannot exist in monocular vision. Other binocular problems can lead to eyestrain, fatigue or headaches (Benjamin, 2006; Steinman, Steinman and Garzia, 2000; Tychsen, 1992).

Binocular fusion and alternatives to fusion

Each eye receives an image that is relayed to the brain. The brain combines the two images into one, a process known as binocular fusion. Binocular fusion may be divided into two stages: motor fusion and sensory fusion. Motor fusion refers to the action of the extraocular muscles that rotate the eyes to keep them fixated on the object of regard. Complex neurological mechanisms use control and feedback to coordinate the actions of the twelve extraocular muscles (six for each eye) and point each eye in the correct direction. Assuming good optics, if the eyes are looking at the same object, the two retinal images will be nearly identical. This is a prerequisite for sensory fusion, which is the neurological process by which the brain combines the two images into one. With defective motor fusion the eyes will not look at the same object and the brain will be faced with a sensory dilemma—how to fuse dissimilar images. If the brain attempts fusion despite the differences, the person will experience visual confusion and diplopia (double vision). Confusion occurs because two different objects will appear in the same location, and diplopia occurs because the object will appear in two non-overlapping positions. This causes visual discomfort and stress, so the brain usually responds automatically to resolve the visual crisis. One solution is to switch attention back and forth between the two images, a condition known as binocular rivalry, but if this condition continues for long, the brain will probably begin to give preference to one image and ignore the other (suppression) (Kimchi et al., 1993). This resolves the sensory dilemma but the person will no longer enjoy binocular vision and its unique benefits.

Stereopsis

Since the two eyes see the world from slightly different vantage points, the right and left retinal images are not exactly identical. In addition the location of objects seen by each eye is slightly different. The differences in visual directions will more pronounced the closer the objects are. If the differences between the right and left eye images are not too great, the visual system is capable of fusing the images. In fact, the visual system detects and analyzes small disparities between the two images to generate the sense of depth perception known stereopsis. Stereopsis allows amazingly fine depth perception, but primarily for objects closer than about 20 feet. For distant objects, differences between the right and left images become insignificant, and stereopsis makes little contribution to depth perception. Instead, we rely on monocular depth cues such as the relative sizes of objects, interposition, convergence of parallel lines, shadows and lighting to provide us with depth perception for faraway objects.

At a typical reading distance of 40 cm, using stereopsis, you should be able to detect that one object is nearer than another if they are separated by a mere 0.5 mm. However, when looking at an object 1000 m away, the minimum separation required for stereopsis is about 750 m. In fact, beyond about 300 meters stereopsis is essentially useless, and we depend primarily on monocular depth cues to judge distances. If the effective separation between the eyes can be increased, image disparities will increase, and there will be a greater stimulus for stereopsis. As an example, field binoculars or large ship-mounted binoculars increase the separation between the viewing positions of the eyes, thereby providing for hypersteropsis*;* that is enhanced stereoscopic depth perception. Some helmet-mounted visual systems create the same effect because each eye's telescope is mounted on the outside of the helmet.

Other differences in the images seen by the two eyes can affect binocular vision depending on the degree of the difference. For example, the quality of binocular vision is not significantly affected by small amounts of monocular blur, but large amounts of blur will degrade binocular visual acuity and stereopsis. It may also lead to eyestrain, rivalry and suppression of the blurred eye. Small image size differences between the two eyes can be tolerated, but they may lead to distorted space perception. Sensory fusion will be difficult for size difference greater than 10%, and will probably lead to diplopia, rivalry or suppression. Differences in the brightness of the retinal images can also affect binocularly perceived brightness. This can also cause an erroneous perception of depth for moving objects, an effect known as the Pulfrich phenomenon. This is sometimes demonstrated by having a person binocularly view a pendulum that is swinging in a plane parallel to the forehead. If a tinted lens is placed over the right eye, the pendulum will appear to swing inward, toward the observer when it moves from left to right, and away from the observer when it swings back.

Ocular dominance

Even if the brain receives equal input from the two eyes, one eye is usually preferred as the dominant eye. The dominant eye is sometimes defined as the eye that is used for sighting or aligning objects under binocular conditions, but there are other definitions for ocular dominance. For example, the eye with the more substantial seeming image, or the eye that is more sensitive to optical blur. The degree to which one eye is dominant over the other varies from person to person, and the dominant eye can switch depending on viewing distance or visual task. In many cases hand and eye dominance are on the same side, but occasionally they are opposite, a condition known as crossed dominance, which can lead to problems in weapons use.

In summary, binocular vision provides enhanced vision compared to monocular vision, and for short distances stereopsis provides extremely precise depth perception.

Conclusion

The eye can be thought of as an optical instrument, but vision depends on much more than optics. Complex physiological processes, including neural image processing are required to perceive an image. The ability to see is further complicated because the visual scene may extend across a wide angular field, objects can be located at different distances, they may be in motion and lighting conditions can vary drastically. Color perception and the addition of a second eye significantly complicate the visual process. By understanding the optics and physiology of vision, engineers and scientists can better design systems that do not interfere with, but rather enhance, vision.

References

Agoston, G. (1979). *Color Theory and Its Application in Art and Design*. (p. 137). Berlin: Springer-VERLAG.

Anderson, D., and Patella, V. (1999). Chapter 1. Introductory concepts. In: *Automated Static Perimetry* (p. 2-9). St. Louis: Mosby.

Applegate, R., Thibos, L., and Hilmantel, G. (2001). Optics of aberroscopy and super vision. *Journal of Cataract and Refractive Surgery*, 27(7), 1093-1107.

Atchison, D., and Scott, D. (2000). Mathematical treatment of ocular aberrations: A user's guide. In: V. Lakshminarayanan (Ed.). *Vision Science and Its Applications*, 35, 110-130. Washington, DC: Optical Society of America.

Atchison, D., and Smith, G. (2000). *Optics of the Human Eye*. Oxford: Butterworth-Heinemann.

Atchison, D.A. (1995). Accommodation and presbyopia. *Ophthalmic and Physiological Optics*, 15(4), 255-272.

Bailey, J., Neitz, M., Tait, D., and Neitz, J. (2004). Evaluation of an updated HRR color vision test. *Visual Neuroscience*, 21, 431-436.

Bailey, M., Olson, M., Bullimore, M., Jones, L., and Maloney, R. (2004). The effect of LASIK on best-corrected high and low-contrast visual acuity. *Optometry and Vision Science*, 81(5), 362-368.

Benjamin, W. (2006). *Borish's Clinical Refraction*. St. Louis: Butterworth-Heinemann, Elsevier.

Bennett, A., and Rabbetts, R. (1991). *Clinical Visual Optics*. Oxford: Butterworth-Heinemann, Ltd.

Bennett, E., and Weissman, B. (2005). *Clinical Contact Lens Practice*. Philadelphia: Lippincott Williams and Wilkins.

Birch, J. (1997). Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic and Physiological Optics*, 17(5), 403-408.

Bradley, A. (1992). Perceptual manifestations of imperfect optics in the human eye: Attempts to correct for ocular chromatic aberration. *Optometry and Vision Science*, 69(7), 515-521.

Bron, A., Tripathi, R., and Tripath, B. (1997). *Wolff's Anatomy of the Eye and Orbit*. London: Chapman and Hall.

Campbell, C., Loshin, D., Applegate, R., Horowitz, L., MacRea, S., Neal, D., Schwiegerling, J., Thibos, L., Webb, R., and Williamsand, D. (2004). ANSI Z80.28-2004 - Methods for reporting optical aberrations of eyes. *American National Standard for Ophthalmics*. Merrifield, VA: Optical Laboratories Association.

Chalita, M.R., and Krueger, R.R. (2004). Correlation of aberrations with visual acuity and symptoms. *Ophthalmology Clinics of North America*, 17(2), 135-142, v-vi.

Cornsweet, T. (1970). *Visual Perception*. New York: Academic Press.

Fan-Paul, N., Li, J., Miller, J., and Florakis, G. (2002). Night vision disturbances after corneal refractive surgery. *Survey of Ophthalmology*, 47(6), 533-546.

Goss, D. (1995). *Ocular Accommodation, Convergence and Fixation Disparity: A Manual of Clinical Analysis*. Boston: Butterworth-Heinemann.

Goss, D., and West, R. (2002). *Introduction to the Optics of the Eye*. Boston: Butterworth-Heinemann.

Hammond, S., Puri, A., and Ambati, B. (2004). Quality of vision and patient satisfaction after LASIK. *Current Opinion in Ophthalmo*logy, 15, 328-332.

Kaiserman, I., Hazarbassonov, R., Varssano, D., and Brinbaum, A. (2004). Contrast sensitivity after wave front-guided LASIK. *Ophthalmology*, 111(3), 454-457.

Kaufman, P., and Alm, A. (2003). *Adler's Physiology of the Eye*. St. Louis: Mobsy.

Kimchi, R., Gopher, D., Rubin, Y., and Raij, D. (1993). Performance under dichoptic versus binocular viewing conditions: Effects of attention and task requirements. *Human Factors*, 35(1), 35-55.

Koretz, J.F., Kaufman, P.L., Neider, M.W., and Goeckner, P.A. (1989). Accommodation and presbyopia in the human eye – aging of the anterior segment. *Vision Research*, 29(12), 1685-1692.

Krueger, R., Applegate, R., and MacRae, S. (2004). *Wavefront Customized Visual Correction*. Thorofare, NJ: Slack.

McLaren, K. (1976). The development of the CIE 1976 (L*a*b*) uniform colour-space and colour-difference formula. *Journal of the Society of Dyers and Colourists*, 92, 338-341.

Mrochen, M., and Semchishen, V. (2003). From scattering to wavefronts--what's in between? *Journal of Refractive Surgery*, 19(5), S597-601.

Nadler, D. (1990). *Glare and Contrast Sensitivity for Clinicians*. New York: Springer-Verlag.

Netter, R. (1975). Nervous System. The CIBA Collection of Medical Illustrations, 1. Summit, NJ: CIBA.

Roorda, A., Metha, A., Lennie, P., and Williams, D. (2001). Packing arrangement of the three cone classes in the primate retina. *Vision Research*, 41(12), 1291-1306.

Roorda, A., and Williams, D. (1999). The arrangement of the three cone classes in the living human eye. *Nature*, 397, 520-522.

Salmon, T.O., and van de Pol, C. (2005). Evaluation of a clinical aberrometer for lower-order accuracy and repeatability, higher-order repeatability and instrument myopia. *Optometry*, 76(8), 461-472.

Salmon, T.O., and van de Pol, C. (2006). Normal-eye Zernike coefficients and root-mean-square wavefront errors. *Journal of Cataract and Refractive Surgery*, 32(12), 2064-2074.

Schallhorn, S., Kaupp, S., Tanzer, D., Tidwell, J., Laurent, J., and Bourque, L. (2003). Pupil size and quality of vision after LASIK. *Ophthalmology*, 110(8), 1606-1614.

Schwartz, S. (2004). *Visual Perception, A Clinical Orientation*. New York: McGraw-Hill.

Steinman, S., Steinman, B., and Garzia, R. (2000). *Foundations of Binocular Vision*. New York: McGraw-Hill.

Stockman, A., Sharpe, L., Ripamonti, C., and Smithson, H. (2006). *Color and Vision Database*. London: Color and Vision Research Laboratories.

Swanson, W., and Cohen, J. (2003). Color vision. *Ophthalmology Clinics of North America*, 16, 179-203.

Tauxe, R. (2003). Food Safety and Irradiation: Protecting the public from food borne infections. *Proceedings of* the *Emerging Infection Diseases,* Atlanta, GA.

Thibos, L., Bradley, A., Still, D., Zhang, X., and Howarth, P. (1990). Theory and measurement of ocular chromatic aberration. *Vision Research*, 30(1), 33-49.

Thibos, L., Bradley, A., and Zhang, X. (1991). Effect of ocular chromatic aberration on monocular visual performance. *Optometry and Vision Science*, 68(8), 599-607.

Thibos, L., Hong, X., Bradley, A., and Cheng, X. (2002). Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *Journal of Optical Society of America A*, 19(12), 2329-48.

Tychsen, L. (1992). Binocular vision. In: W. Hart (Ed.) *Adler's Physiology of the Eye* (p. 773-853). St. Louis: Mosby Year Book.

Van Hateren, J.H. (1993). Spatiotemporal contrast sensitivity of early vision. *Vision Research*, 33(2), 257-267.

Westin, E., Wick, B., and Harrist, R.B. (2000). Factors influencing success of monovision contact lens fitting: survey of contact lens diplomates. *Optometry*, 71(12), 757-763.

Yamane, N., Miyata, K., Samejima, T., Hirakoka, T., Kiuchi, T., Okamoto, F., Hirohara, Y., Mihashi, T., and Oshika, T. (2004). Ocular higher-order aberrations and contrast sensitivity after conventional laser in situ keratomileusis. *Investigative Ophthalmology and Visual Science*, 45(11), 3986-3990.

8 BASIC ANATOMY OF THE HEARING SYSTEM

Sumalai Maroonroge
Diana C. Emanuel
Tomasz R. Letowski

Hearing System

Hearing is the sense by which biological systems are aware of the surrounding acoustic environment and perceive sound (see Chapter 11, *Auditory Perception and Cognitive Performance*). It is the primary sense by which various species respond to limited range of physical vibrations in the atmosphere. Human hearing allows for the perception of speech and other acoustic events and for 360° spatial detection and localization of sound sources. However, human hearing is sensitive to a limited range of sound intensities and frequencies and only allows for full 360° of spatial orientation when the listener is not obstructed by any proximal acoustic barriers. Therefore, in order for audio head- and helmet- mounted displays (HMDs) to take full advantage of the wearer's hearing capabilities, HMD designers and the acquisition corps need to have a solid understanding of the anatomy and physiology of human hearing.

The act or process of hearing is called *audition,* and the anatomical structure processing incoming acoustic stimuli is called the *hearing system* or *auditory system*. The human hearing system consists of two ears, located on the left and right sides of the head, the vestibulocochlear nerve, and the central auditory nervous system (CANS) – consisting of auditory centers in the brain and the connecting pathways in the brainstem. Each ear is additionally divided into three functional parts: the outer (external) ear, the middle ear and the inner (internal) ear. The overall anatomical structure of the human ear and its division in three functional parts are shown in Figure 8-1. (A more detailed but more schematic picture of the ear structures is shown in Figure 8-10). The inner ear contains three parts: the vestibule, semicircular canals, and the cochlea and serves as housing for two sensory organs, specifically, the organ of balance and the organ of hearing. The parts of the organ of balance are contained within the vestibule and the semicircular canals. The organ of hearing, the *organ of Corti*, is located in the cochlea. The adjacent locations of the senses of hearing and balance result in some interactions between the sense of hearing and the sense of balance.

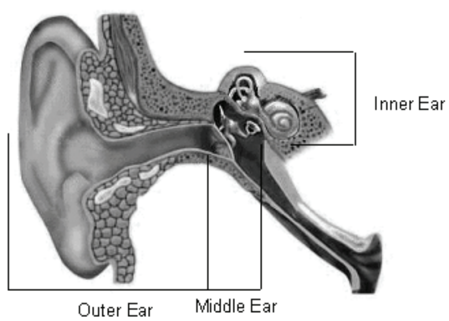


Figure 8-1. Overall structure of the human ear (adapted from http://www.telezdrowie.pl/slysze/english/info.htm).

The structures of the human ear are embedded in the temporal bone of the skull, with only part of the outer ear (the pinna) protruding outside the skull and being visible. The temporal bone is a dense bony structure on either side of the head that forms part of the cranium (cranial vault) around the brain. The *cranium* consist of 8 bones (paired temporal and parietal bones and single frontal, occipital, sphenoid, and ethmoid bones) connected by seams called *sutures.* In addition to the cranium, the skull is comprised of a group of 14 facial bones that make up the facial skeleton. The facial bones include paired nasal, lacrimal, palatine, inferior nasal concha, maxilla, and zygomatic bones; with single mandible and vomer bones. A list of all the main bones of the skull is provided in Table 8-1. The overall structure of the skull and the locations of the main constituent bones of the skull are shown in Figure 8-2. Note that the sphenoid bone appears to be a plate on the side of the skull, but is actually a butterfly-shaped bone that extends the width of the skull from right to left. Only the lateral aspects are visible in the figure.

Table 8-1.
Cranial and facial bones of the human skull
(Henry and Letowski, 2007)

Cranial Bones		Facial Bones	
Single	*Paired*	*Single*	*Paired*
Frontal bone	Parietal bone	Mandible	Maxilla
Occipital bone	Temporal bone	Vomer bone	Palatine bone
Sphenoid bone			Zygomatic bone
Ethmoid bone			Nasal bone
			Lacrimal bone
			(Inferior) nasal concha

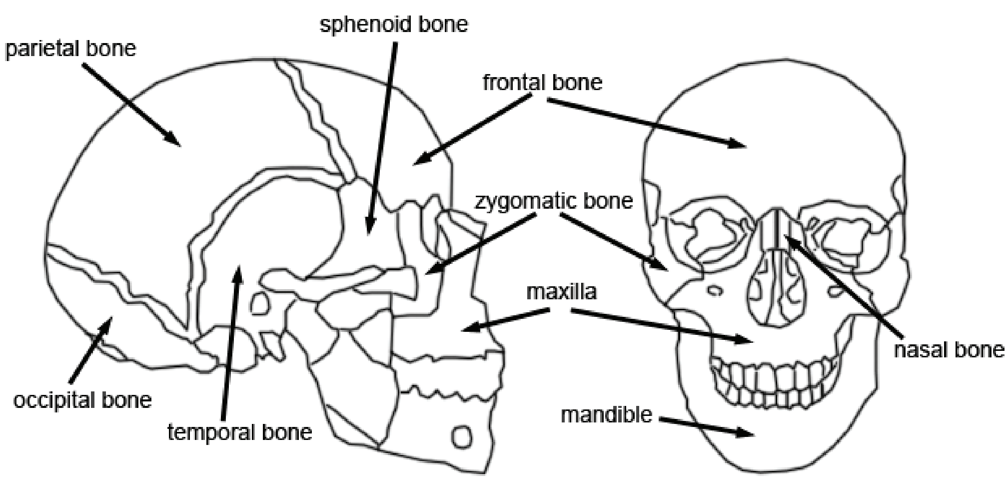


Figure 8-2. Bones of the skull (Howell, Williams and Dix, 1988).

Figure 8-3 shows several important landmarks of the skull for placing bone communication HMDs (see Chapter 5, *Auditory Helmet-Mounted Displays*) including the condyle, mastoid process, forehead, and temple (the bony point above the temple), which are parts of the mandible, temporal bone, frontal bone, and frontal bone again, respectively.

The location of the ear canal in the temporal bone is marked in Figure 8-3 as a small oval in the lower part of the bone. The middle ear and inner ear are located in the petrous portion of the temporal bone, which is a dense core of bone that provides protection for the delicate ear structures. In addition to housing the structures of the ear, the petrous portion of the temporal bone has additional canal, the internal auditory meatus, through which pass the vestibulocochlear (8th cranial) and the facial (7th cranial) nerves. The facial nerve provides sensory and motor innervation to the face.

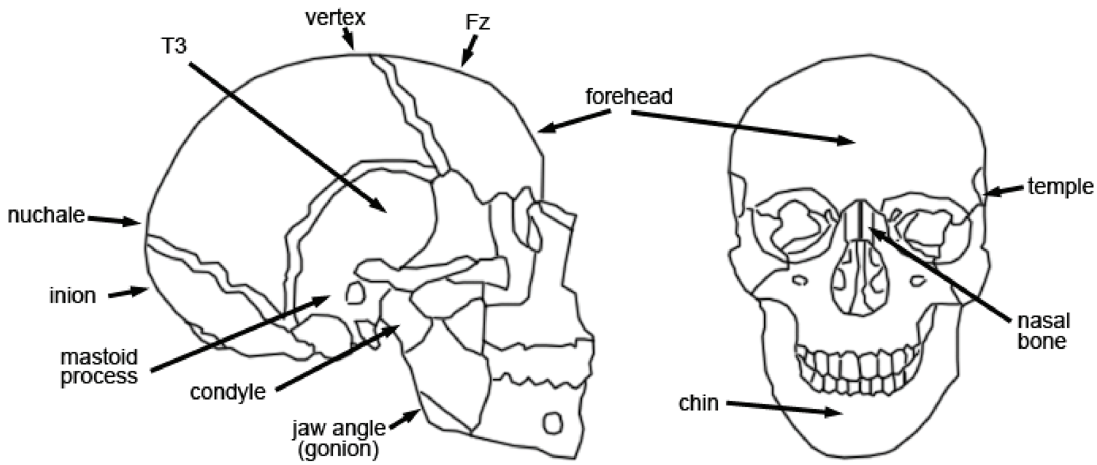


Figure 8-3. Landmark points of the skull used for placing bone conduction vibrators (adapted from (Howell, Williams, and Dix, 1988).

The anterior (front) portion of the temporal bone articulates with the condyle of the mandible, forming the temporomandibular joint (TMJ). The superior (top) portion of the temporal bone is the squamous portion, which is a fan like projection that attaches to the occipital and parietal bones. The posterior (back) portion of the temporal bone is the mastoid portion. The mastoid portion includes a bony protuberance called the mastoid process. The mastoid process is the bony ridge that one can feel on the skull just behind the pinna. The mastoid process is the usual place for attaching bone conduction hearing aids and placing bone conduction vibrators during hearing testing. The condyle, which lies just in front of the visible part of the outer ear, is the very effective location for placing a bone vibrator is speech communication applications (McBride, Tran and Letowski, 2005; 2008).

Outer Ear

The outer ear consists of two major elements: the external flange of the ear (called the pinna) and the ear canal. Both elements are shown in Figure 8-4. The ear canal is terminated by the tympanic membrane (eardrum), which separates the outer ear from the middle ear. The pinna projects from the side of the head at an angle of 25° to 35° (mean value 30°) to the occipital scalp (Glasscock and Shambaugh, 1990; McDonald, 1993; Sclafani and Ranaudo, 2006) and serves as a sound collector. The entrance to the ear canal is located within the pinna, in front of the pinna flap. The ear canal directs sound waves toward the eardrum and protects the eardrum from the external environment (e.g., dust, small flies, and changes in temperature).

Pinna

The pinna (auricle) is an ovoid-shaped structure with an uneven surface filled with numerous grooves and depressions. Humans have two pinnae, one on each side of the head. Similarly to most paired anatomical structures, the two pinnae differ in their specific shape and in their patterns of grooves and pits. In addition, their locations are usually slightly asymmetrical in reference to each other in both vertical and horizontal planes. These differences, together with the fact that acoustic signals are simultaneously received by the two ears, facilitate our ability to localize sounds in space and are responsible for the fact that localization mechanisms are not transferable from one person to another. This is the reason that the Head-Related Transfer Function (see Chapter 11, *Auditory Perception and Cognitive Performance*), of one person cannot be successfully used by another person without some adjustments.

The internal frame of the pinna is composed of a single piece of cartilage that is attached to surrounding tissues and covered with skin. The pinna is innervated by nerve fibers from the great auricular nerve and the auriculotemporal nerve. The blood supply to the pinna is provided by the posterior auricular and superficial temporal arteries. The pinna is connected to the head by ligaments and small muscles. Many species use these muscles to direct the pinna towards incoming sound, but humans lost this ability (although some humans have maintained rudimentary pinna motion ability).

The average length of the pinna is approximately 65 millimeters (mm) (2.6 inches [in]) and the average width is approximately 35 mm (1.4 in). See Table 8-2 for more information about pinnae dimensions. In most adults the width of the pinna is approximately 55% of its length (McDonald, 1993). The main axis of the pinna points upwards with a posterior tilt of 5° to 30° (typical values: 15° to 20°) (McDonald, 1993; Shaw, 1974; Yost and Nielsen, 1977). The highest point or superior aspect (crux) of the pinna is typically even with the brow.

A view of the pinna and its major structures is shown in Figure 8-5. The prominent curved rim is called the helix. The second, internal, ridge is the antihelix, which runs almost parallel to the helix. The helix and antihelix are separated by a narrow curved depression called the scaphoid fossa.

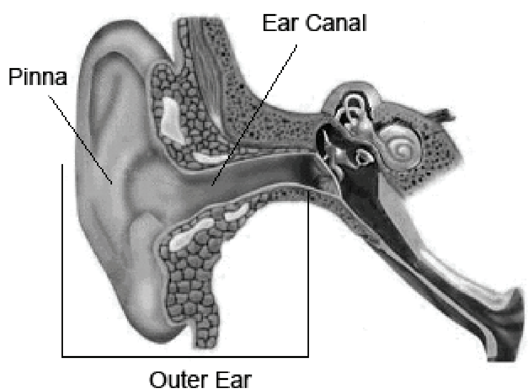


Figure 8-4. The outer ear and its main elements (adapted from http://www.telezdrowie.pl/slysze/english/info.htm).

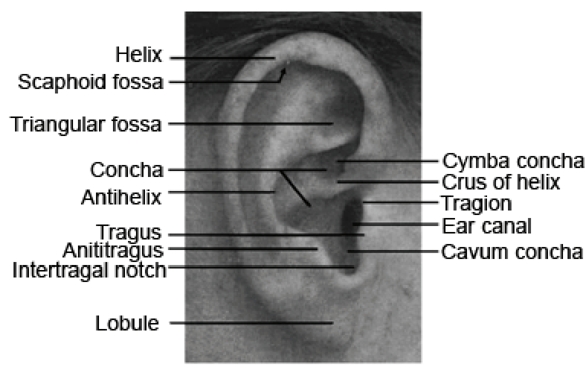


Figure 8-5. The pinna and its major structures (adapted from Rohen and Yokochi, 1983).

The entrance to the ear canal is located in the lower part of the pinna and in the center of the major pinna depression called the concha. The concha is an oval, bowl-like, major depression in the pinna and is divided by the crus of the helix into the cymba concha and cavum concha. The cavum concha surrounds the inlet to the ear canal. The dimensions of the concha vary from person to person but the average diameter of the concha is typically 15 to 20 mm (0.6 to 0.8 in) and its average depth is approximately 13 mm (0.5 in) (Burkhard and Sachs, 1975). The average volume of the concha cavity is approximately 4.0 cm^3 (0.2 in^3) (Teranishi and Shaw, 1968; Zwislocki, 1970; 1971). More detailed information about concha dimension is provided in Table 8-2.

In the front of the entrance to the ear canal, there is a small cartilaginous flap called the tragus. The tragus partially covers the opening of the ear. The notch above the tragus is called the tragion and is frequently used as a point of reference in anatomical measurements. A similar notch located below the tragus and separating the tragus from a second cartilaginous flap, called the antitragus, is the intertragal notch (intertragal incisure). The intertragal notch is used as a reference point for inserting a probe microphone into the ear canal in real-ear measurements (Henry and Letowski, 2007). The distance between the entrance to the ear canal and the intertragal notch is approximately 10 mm (0.4 in) (Hawkins, Alvarez and Houlihan, 1991; Pumford and Sinclair, 2001). At the very bottom of the pinna, below the intertragal notch, there is a large soft flap called the lobule (ear lobe). The lobule does not contain cartilage and is entirely made of fat and skin.

Table 8-2.
Basic average (mean) dimensions of the human ear (pinna).

Dimension	Unit	Male Adult	Female Adult	Source(s)
Ear width (Pinna breath)	mm	37.7 35.5 34.5 29.2 37.0	33.6 33.0	Burkhard and Sachs, 1975 Dreyfus, 1967 Alexander and Laubach, 1968 Algazi et al., 2001 IEC 60318-7, 2009
Ear length (Pinna height)	mm	68.5 63.5 67.0 64.1 66.0	62.4 58.4	Burkhard and Sachs, 1975 Dreyfus, 1967 Alexander and Laubach, 1968 Algazi et al., 2001 IEC 60318-7, 2009
Ear length above tragion	mm	33.0 30.4 30.0	30.7	Burkhard and Sachs, 1975 Dreyfus, 1967 IEC 60318-7, 2009
Ear protrusion distance	mm	22.8 21.0 23.0	20.3	Burkhard and Sachs, 1975 Dreyfus, 1967 IEC 60318-7, 2009
Ear protrusion angle	°	23.3 28.5 20.0	24.9	Burkhard and Sachs, 1975 Algazi et al., 2001 IEC 60318-7, 2009
Ear vertical tilt to the back	°	7.6 24.1 10.0	4.7	Burkhard and Sachs, 1975 Algazi et al., 2001 IEC 60318-7, 2009
Ear vertical tilt to the side	°	3.0 6.0	2.7	Burkhard and Sachs, 1975 IEC 60318-7, 2009
Ear canal offset down	mm	30.3		Algazi et al., 2001
Ear canal offset back	mm	4.6		Algazi et al., 2001
Concha length	mm	27.3 25.9 28.0	25.3	Burkhard and Sachs, 1975 Algazi et al., 2001 IEC 60318-7, 2009
Concha breadth	mm	18.8 15.8 23.0	17.2	Burkhard and Sachs, 1975 Algazi et al., 2001 IEC 60318-7, 2009
Concha depth	mm	12.9 10.2 15.0	12.9	Burkhard and Sachs, 1975 Algazi et al., 2001 IEC 60318-7, 2009
Concha volume	cm^3	4.65	3.94	Burkhard and Sachs, 1975

Note: IEC 60318-7 data refer to a standardized acoustic manikin. See also IEC 60268-7 (2009).

Ear canal

The ear canal (auditory canal; external meatus) is an "S" shaped duct providing an access route for acoustic waves to travel to the tympanic membrane. A general view of the ear canal is shown in Figure 8-4. The outer one third of the ear canal is surrounded by cartilage, whereas the remaining inner two thirds of the canal are surrounded by bone, as the canal enters the temporal bone. The two respective parts of the canal are called the cartilaginous part and the osseous part of the canal. The cartilaginous part is lined with relatively thick layer of skin (0.5 to 1.0 mm [0.02 to 0.04 in] thick) that is continuous with that of the pinnae and contains numerous sebaceous glands, ceruminous (wax) glands, and hair follicles (Lucente, 1995). The cartilaginous part of the ear canal produces cerumen (ear wax), which is comprised of secretions from sebaceous and apocrine glands (Lucente, 1995; Ballachanda, 1995). Cerumen acts to moisturize the skin and, together with hairs, trap dust, debris and other small objects entering the ear canal

The osseous part of the ear canal is covered with relatively thin skin (approximately 0.2 mm [0.01 in] thick) that is continuous with the outer layer of the tympanic membrane (Muller, 2003). There are no hairs or secretion producing glands located in this part of the ear canal (Lucente, 1995). The skin of the ear canal is innervated by the branches of three cranial nerves: the auriculotemporal (mandibular) nerve, the facial nerve, and the vagus nerve.

The outer layer of the skin lining the ear canal and tympanic membrane has lateral migratory properties. The surface cells of the skin move laterally from the tympanic membrane toward the ear canal opening. This property is an active process that enables self-cleaning of the tympanic membrane and the removal of the wax and trapped debris from the inside of the ear canal. The rate of skin cell migration is approximately 100 microns (μm) per day (Muller, 2003).

The average length of an adult ear canal is approximately 25 mm (1.0 in) with a standard deviation of approximately 2 mm (0.2 in) and is approximately 5% longer in males than in females (Alvord and Farmer, 1997; Wever and Lawrence, 1954; Zemlin, 1997; Zwislocki, 1970). The effective acoustic length of the ear canal is approximately 25% larger than its geometrical length due to the “end effect” of the concha and the manner in which the concha is coupled to the ear canal (Teranishi and Shaw, 1968). The canal is oval in shape with an average diameter of 7.0 to 8.0 mm (0.28 to 0.31 in) (Alvord and Farmer, 1997; Békésy, 1932; Zemlin, 1997; Zwislocki, 1970). The shape and cross sectional dimensions of the ear canal change along its length. The oval opening of the canal has average dimensions of 9 mm (0.4 in) by 6.5 mm (0.3 in), and the canal becomes narrower along its length (Shaw, 1974). The cross sectional area of the canal is approximately 0.45 cm^2 (0.07 in^2) at the opening and approximately 0.4 cm^2 (0.06 in^2) in the middle of the canal (Zwislocki, 1970). The final 8 to 10 mm (0.3 to 0.4 in) of the ear canal is slightly tapered and reaches its narrowest point at the isthmus [isthmus (gr.) – passageway], which is located just past the second bend of the ear canal and approximately 4 mm (0.2 in) from the tympanic membrane (Seikel, King, and Drumright, 2000). The tympanic membrane terminates the ear canal at an oblique angle of 45° to 60° in reference to the floor of the canal (Gray, 1918; Stinson and Lawton, 1989; Decraemer, Dirckx and Funnell, 1991; Seikel, King and Drumright, 2000; Sundberg, 2008). This oblique position of the tympanic membrane causes the length of the ear canal to be approximately 6 mm (0.24 in) shorter at its posterior/superior (back/top) portion compared to its anterior/inferior (front/bottom) portion. The cross sectional area of the ear canal in male adults is approximately 10% larger than in female adults (Zwislocki, 1970). The average volume of the canal has a range of approximately 1.0-1.4 cm^3 (0.06 to 0.09 in^3) range (Liu and Chen, 2000; Wever and Lawrence, 1954).

Since the proper selection and fitting of an audio HMD depends on the dimensions of the human head, locations and dimensions of pinnae, and the geometry of the ear canal, the mean values of some of the main dimensions of the pinnae and human head are provided in Tables 8-2 and 8-3, respectively, for reference purposes. The illustration of the extent of each of the main dimensions listed in Table 8-3 is shown in Figure 8-6.

The sources of the values listed in Table 8-3 are large databases or anthropometric studies conducted with large number of participants. For example, the data provided by Burkhard and Sachs (1975) are based on a survey of over 4000 people conducted by Churchill and Truett (1957) and the large set of data accumulated by Dreyfuss (1967). The data for ear dimensions, listed in Table 8-2, are based on a smaller number of measurements compared with data in Table 8-3. These samples include data published by Algazi et al. (2001), which are based on 45 people and data of Burkhard and Sachs (1975) based on 12 male and 12 female adults.

Table 8-3.
Basic average (mean) dimensions of the human head.

Dimension	Unit	Male Adult	Female Adult	Source
Head width (Head breath)	mm	155 152 150 154	147 144	Burkhard and Sachs, 1975 DoD, 2000; IEC 60318-7, 2009 Algazi et al., 2001 ISO 7250-2, 2008
Head length (Head depth)	mm	196 197 220 191 200	180 187 180	Burkhard and Sachs, 1975 DoD, 2000 Algazi et al., 2001 IEC 60318-7, 2009 ISO 7250-2, 2008
Head height	mm	215		Algazi et al., 2001; IEC 60318-7, 2009
Head height from tragion (point at the notch above tragus)	mm	130	130	Burkhard and Sachs, 1975
Head circumference	mm	570 573 576	550 551	NASA, 1978 Algazi et al., 2001 ISO 7250-2, 2008
Menton (chin)-vertex height	mm	232 232 224	211 218	Burkhard and Sachs, 1975 DoD, 2000 IEC 60318-7, 2009
Tragion-to-tragion distance (bitragion diameter)	mm	142 145 143	135 132	Burkhard and Sachs, 1975 DoD, 2000 IEC 60318-7, 2009
Tragion-to-shoulder distance	mm	188 175	163	Burkhard and Sachs, 1975 IEC 60318-7, 2009
Tragion-to-wall (behind the head)distance	mm	102 97	94	Burkhard and Sachs, 1975 IEC 60318-7, 2009
Neck diameter	mm	121 117 113	103	Burkhard and Sachs, 1975 Algazi et al., 2001 IEC 60318-7, 2009
Head-torso forward offset	mm	30		Algazi et al., 2001
Shoulder breadth	mm	455 491 459 440	399 431	Burkhard and Sachs, 1975 DoD, 2000 Algazi et al., 2001 IEC 60318-7, 2009
Chest breadth	mm	305 315 282	277	Burkhard and Sachs, 1975 Algazi et al., 2001 IEC 60318-7, 2009

Note 1: Burkhard and Sachs (1975) data are median values.
Note 2: Algazi et al. (2001) data are overall means for male and female adults.
Note 3: DOD – U.S. Department of Defense.
Note 4: IEC 60318-7 data refer to a standardized acoustic manikin. See also IEC 60268-7 (2009).
Note 5: ISO 7250-2 data are for the U.S. population

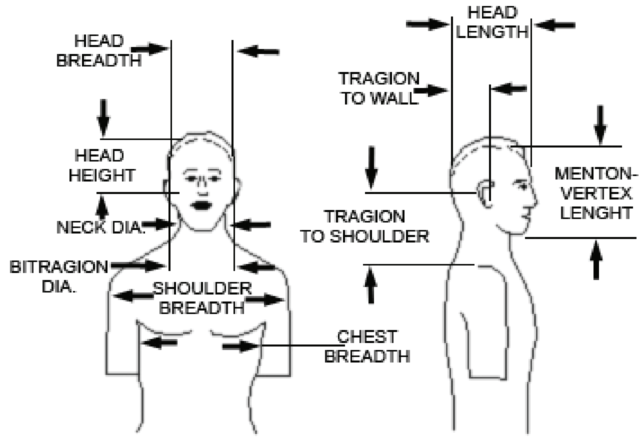


Figure 8-6. Selected anthropometric dimensions of the human head (Burkhard and Sachs, 1975)

Middle Ear

The middle ear is an air-filled cavity called the *tympanic cavity* (tympanum). The walls of the cavity are formed from the temporal bone and the cavity is lined with mucous membrane tissue. The overall volume of the middle ear is approximately 2 cm^3 (0.12 in^3) (Dallos, 1973; Yost and Nielsen, 1977; Zemlin, 1997). The lateral wall of the middle ear contains the tympanic membrane (previously described), and the medial wall is formed by a bony wall that separates the middle ear from the inner ear. This wall contains two membranous windows, called the oval and round windows, which act to anatomically and physiologically connect the middle ear with the inner ear. The air in the middle ear cavity remains just below atmospheric pressure due to the connection between the tympanic cavity and the upper part of throat (nasopharynx) by a narrow duct called the *Eustachian tube* (auditory tube). Within the middle ear cavity are three small bones called the *malleus* (hammer), *incus* (anvil), and *stapes* (stirrup). These bones are collectively called the *ossicles* and form a chain called ossicular chain that connects the tympanic membrane with the oval window. The chain is suspended inside the cavity by middle ear ligaments and two middle ear muscles: the *tensor tympani* and the *stapedius*. The overall structure of the middle ear is shown in the Figure 8-7.

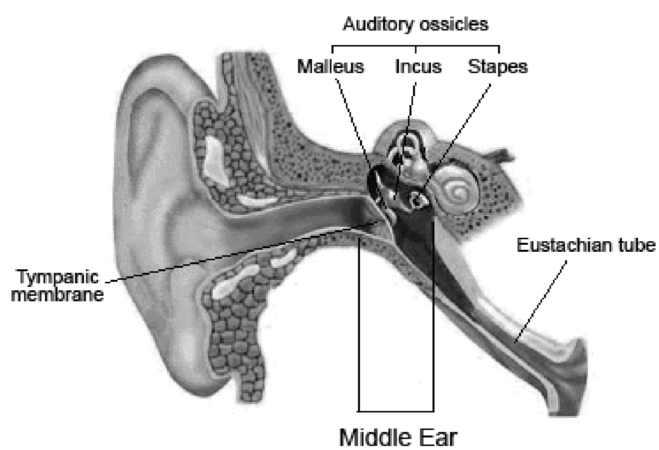


Figure 8-7. The middle ear and its main elements (adapted from http://www.telezdrowie.pl/slysze/english/info.htm).

The tympanic cavity has the shape of a narrow, irregular rectangular box that is narrower in the middle than the edges. A schematic view of the middle ear cavity is shown in Figure 8-8. The largest dimension of the tympanic cavity does not exceed 10 mm (0.4 in). The ossicular chain and middle ear muscles take up most of the space within the cavity. The small empty space above the ossicles is called the *epitympanum recess* (attic). The remaining empty space is referred to as the tympanic *cavity proper*. The superior wall (ceiling) of the tympanic cavity is formed by a thin bone called the *tegmen tympani*, which separates the middle ear from the brain cavity. A small narrow aperture, called the *aditus ad antrum*, located at the top of the posterior (back) wall, connects the middle ear cavity to another small chamber called the *mastoid antrum* (tympanic antrum) that is surrounded by the *mastoid air cells*. The entrance to the Eustachian tube is located in the anterior (front) wall of the middle ear cavity. The oval and round windows of the inner ear form the medial wall of the cavity. The windows are separated by a ridge of bone called the *promontory*. The inferior wall (floor) of the cavity contains the *jugular fossa* that contains the jugular vein. The pulsation of blood in the vein can be a source of ringing noise (tinnitus) in the ear.

.

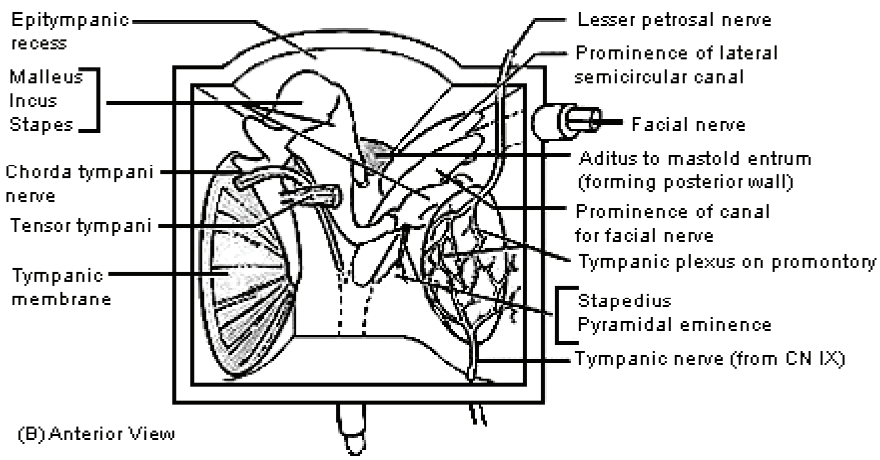


Figure 8-8. A schematic view of the middle ear cavity without the anterior wall (Moore and Dalley, 1999).

The tympanic cavity also contains two middle ear muscles: the tensor tympani muscle and the stapedius muscle. These two muscles are the smallest muscles in the human body, with the tensor tympani being the larger of the two. The tendon of the stapedius muscle emerges from the pyramidal eminence of the posterior (back) wall. The other end of the muscle is connected to the stapes. The *chorda tympani* nerve, a branch of the facial nerve, also emerges from the posterior wall of the cavity. The chorda tympani travels through the middle ear from back to front, joining the lingual nerve and providing taste sensation to part of the tongue. The tendon of the second middle ear muscle, the *tensor tympani*, emerges from the anterior wall of the middle ear and attaches to the malleus bone. A thin plate of anterior bone separates the middle ear from the internal carotid artery (Zemlin, 1997).

Tympanic membrane

The tympanic membrane (eardrum) is a thin, semi-transparent, oval membrane terminating the ear canal. The membrane is shaped like a shallow cone with its tip approximately 1.5 to 2 mm (0.06 to 0.08 in) out towards the middle ear (Alvord and Framer, 1997; Sundberg, 2008). The tip is called the *umbo* and is attached to the tip of *manubrium* ("handle") of the malleus bone (Pickles, 1988). The apex angle of the tympanic membrane is approximately 120° (Fumagalli, 1949). The membrane is attached around most of its circumference to the temporal bone surrounding the ear canal except for the very narrow area called the *notch of Rivinus* (incisura tympanica). The dimensions of the membrane along its two major perpendicular axes are 9 to 10 mm (0.35 to 0.39 in) and 8 to 9 mm (0.31 to 0.35 in) (Gelfand, 1998; Gray, 1918). The total surface area of the tympanic membrane varies from 55 to 90 mm^2 (0.09 to 0.14 in^2) with a mean value of 64 mm^2 (0.10 in^2) (Harris, 1986; Zemlin, 1997). However, due to its conical shape the effective area of the membrane is smaller with a mean value of approximately 55 cm^2 (0.09 in^2) (Seikel, King, and Drumright, 2000; Zemlin, 1997). The membrane is the thinnest at its center and the thickest at its edge. It has an average thickness of approximately 70 μm but can vary from approximately 30 μm to 120 μm (Donaldson and Miller, 1980; Kojo, 1954; Lim, 1970; Wever and Lawrence, 1954). The membrane is composed of four layers of tissue; the outer epithelial layer that is continuous with the skin of the ear canal, two middle fiber layers consisting of radial and concentric fibers and responsible for the stiffness of the membrane, and the inner mucus layer that is continuous with the lining of the middle ear cavity. The external layer of the membrane is innervated by the auriculotemporal nerve. The average mass of the tympanic membrane is approximately 14 milligrams (mg) (0.0005 ounces [oz]) (Lee, 2009; Shennib and Urso, 2000).

The surface of the membrane does not have a uniform stiffness and can be divided into two main regions that differ greatly in their stiffness: a small lax triangular area located at the top of the membrane called the *pars flaccida* (also Shrapnell's membrane) and the much larger and stiffer portion called the *pars tensa.* The function of the pars flaccida is to compensate for small air pressure changes between the middle and outer ear and to allow the tympanic membrane to work like a piston rather than a membrane that is fixed at its circumference (Gray, 1918). Only the pars tensa is involved in the transmission of acoustic energy from the outer ear to the middle ear and therefore the effective area of the tympanic membrane is smaller than the overall area. The modulus of elasticity[1] (Young's modulus) for the tympanic membrane at its center is approximately 0.02-0.03 GPa (Békésy, 1949; Decraemer, Maes, and VanHuyse, 1980).

Ossicular chain

The ossicular chain connects the tympanic membrane with the membrane covering the oval window. The primary purpose of the ossicles and supporting muscles is to transfer sound energy from the tympanic membrane to the inner ear while shielding the inner ear from excessive harmful noise. A diagram of the ossicular chain together with the supporting middle ear muscles is shown in Figure 8-9.

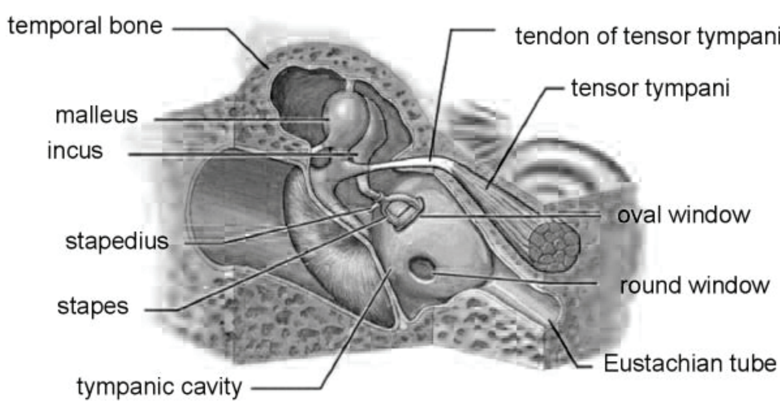


Figure 8-9. The middle ear space and ossicle chain. Petrous part of the temporal bone. (http://www.sfu.ca/ ~saunders/l33098/Ear.f/midear.html).

[1] The modulus of elasticity (also known as Young's modulus [E]) is a measure of stiffness of an elastic material. The common unit of elasticity is the *Pascal*, a unit of pressure equal to 1 Newton per square meter; the unit *GPa* is a gigapascal; 1 GPa = 145,038 pounds per square inch.

The ossicles are the smallest bones in the human body. The malleus, the largest of the three, is approximately 8 to 9 mm (0.31 to 0.35 in) in length, while the stapes, the smallest human bone, is approximately 3 mm (0.12 in) long (Hall and Mueller, 1997; Wever and Lawrence, 1954). The incus is approximately 5 to 7 mm (0.20 to 0.28 in) long. Their respective weights are approximately 25, 3 and 28 mg (0.00088, 0.00011 and 0.00099 oz) (Yost and Nielsen, 1977). The maleus is the first bone in the ossicular chain and the largest of its three processes, the manubrium (the handle), is firmly attached to the tympanic membrane. They both move in unison. The opposite end of the malleus, called the head, is attached to the second bone in the chain, the incus. The incus consists of a body and two processes. The body of the incus connects to the head of the malleus while the longer of the two projections, the lenticular process, connects with the final bone in the chain, the stapes. The shorter projection is connected by a ligament to the back (posterior) wall of the tympanic cavity. The body of the stapes is divided into four parts: a head, a neck, two arms and the footplate. The footplate and the arms form a characteristic shape of the stirrup. The bottom surface of the footplate is tightly attached to the membrane covering the oval window.

The ossicles move in response to the acoustic pressure impinging on the tympanic membrane and in response to the actions of two middle ear muscles: the tensor tympani muscle and the stapedius muscle. The muscles contractions reduce the level of very intense sounds transmitted to the inner ear. The tensor tympani muscle is attached to the manubrium of malleus (see Figure 8-8) and its contraction pulls the handle of the malleus inward and regulates the tension on the tympanic membrane (Gray, 1918; Rodriguez-Velasquez et al. 1998). The stapedius muscle is connected to the stapes and its contraction pulls the stapes in an inferior and lateral (away from the oval window) direction which reduces the range of motion of the stapes (Djupesland and Zwislocki, 1971; Zwislocki, 2002).

Eustachian tube

The Eustachian tube is a thin tube, which connects the middle ear with the nose and throat (Figure 8-6). The tube is approximately 35 to 45 mm (1.4 to 1.8 in) long and it travels downward and inward from the middle ear to the nasopharynx (upper throat). At its upper end, the tube is narrow and surrounded by bone. Nearer the pharynx it widens and becomes cartilaginous.

The Eustachian tube is normally closed, opening only during swallowing and yawning. It is responsible for maintaining the air pressure within the middle ear at approximately ambient pressure. Similar pressure on both sides of tympanic membrane ensures that the tympanic membrane can vibrate maximally when struck by sound waves arriving from the ear canal.

Inner Ear

The inner ear is the final and the most complex part of the ear. It occupies a small bony cavity called the *bony labyrinth* (osseous labyrinth) that is located directly behind the medial wall of the middle ear. The inner ear consists of three main anatomical elements: the *semicircular canals*, the *vestibule*, and the *cochlea*. The structure and main elements of the inner ear are shown in Figure 8-10.

The bony labyrinth of the inner ear has a volume of approximately 2 cm^3 and is lined by the *membranous labyrinth* that closely follows the shape of the bony labyrinth (Buckingham and Valvassordi, 2001). The blood supply to the membranous labyrinth is provided by various small blood vessels extending from the *labyrinthine artery*.

The space between the bony labyrinth and the membranous labyrinth is filled with incompressible body fluid called *perilymph*. The perilymph is high in sodium but low in potassium resembling in its chemical composition in the blood and the cerebrospinal fluid surrounding the brain. The space inside the membranous labyrinth is filled with another incompressible body fluid called *endolymph*. Endolymph is low in sodium but high in potassium and chemically resembles the intercellular fluid found inside cells in the body. The differences in the chemical composition of the perilymph and endolymph create an electric potential difference that like a battery, sustains the physiological activities of the sensory organs located in the inner ear.

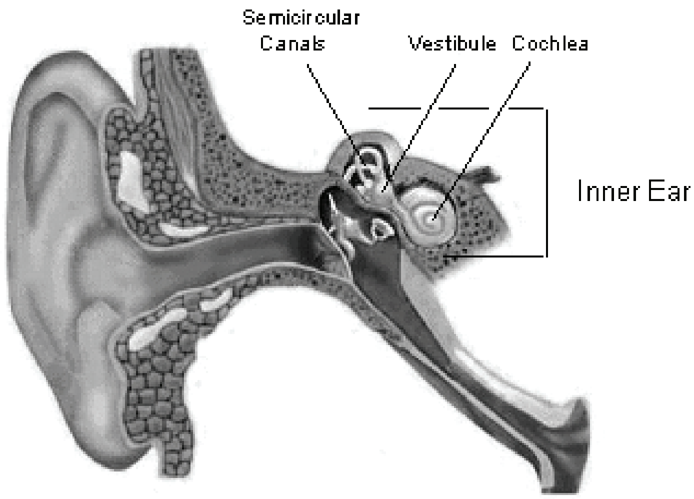


Figure 8-10. The inner ear and its main elements (adapted from [http://www.telezdrowie.pl/ slysze/ english/info.htm).

Functionally, the inner ear consists of two major elements: the *cochlea* and the *vestibular system* (comprising the utricle and saccule of the vestibule and the three semicircular canals). The cochlea contains the organ of hearing (organ of Corti) while the vestibular system contains five balance organs: two maculae (utricular macula and secular macula) and three cristae ampullares (one in each of the three semicircular canals). The cochlea and the semicircular canals are located at the two ends of the inner ear, while the utricle and saccule are parts of the centrally-located vestibule. The oval window is located on the wall of the vestibule and the round window is located at the base of the cochlea. The locations of the windows and the other major parts of the inner ear are shown schematically in Figure 8-11.

Two vary narrow channels; the *vestibular aqueduct* and the *cochlear aqueduct* (not shown in Figure 8-11); connect the inner ear with the cranial cavity surrounding the brain. At their narrowest points, the vestibular aqueduct and the cochlear aqueduct do not normally exceed 0.8 mm (0.03 in) and 0.15 mm (0.006 in) in diameter, respectively. Both aqueducts seem to have little effect on normal ossicular transmission of sound for frequencies above 20 Hertz (Hz) (see, for example, Gopen, Rosowski and Merchant [1997]) and their exact function is still unknown.

Cochlea

The cochlea is a coiled structure that resembles the snail and extends anteriolaterally from the vestibule. Its structural base is the bony *spiral lamina,* which makes 2½ to 2¾ turns around the bony core of the cochlea called the *modiolus*. The external diameter of the cochlea varies from approximately 9 mm (0.35 in) at its base to approximately 5 mm (0.20 in) at its apex (top) and its uncoiled length is 32 to 35 mm (1.25 to 1.38 in). The cochlea is divided along its length into three parallel channels: the *scala vestibuli, scala media,* and *scala tympani*. The scala vestibuli and scala tympani are parts of the bony labyrinth whereas the scala media is a part of the membranous labyrinth. The inside of the cochlea is shown in Figure 8-12.

The scala vestibuli and scala tympani are connected at the apex (top) of the cochlea through a small opening called the *helicotrema*. At the base of the cochlea, the scala vestibuli joins the vestibule. The scala vestibule is terminated at its base by the oval window, the *fenestra ovalis,* while scala media terminates at the round window,

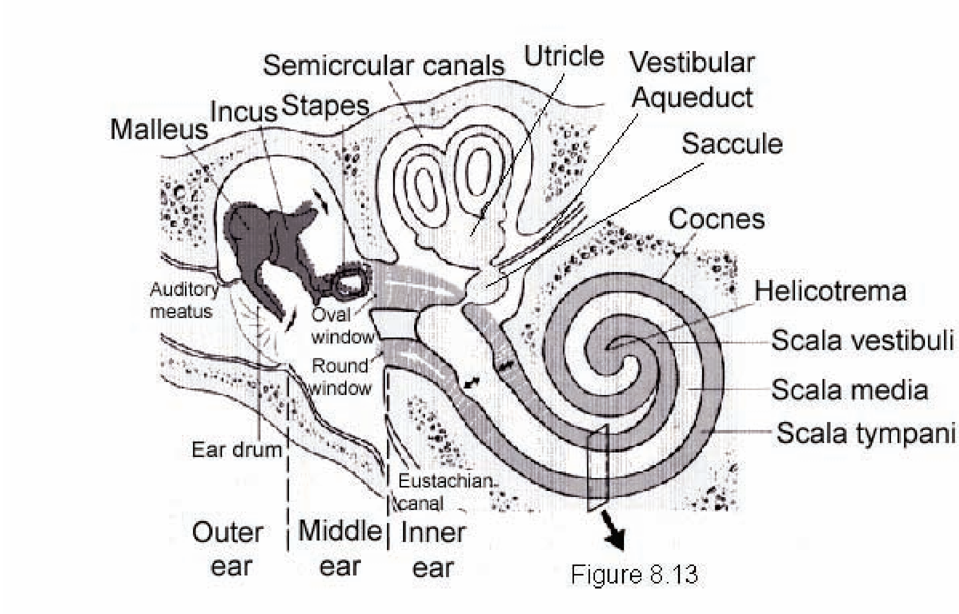


Figure 8-11. A schematic view of the main structures of the inner ear (adapted from Despopoulos and Silbernagl, 1991).

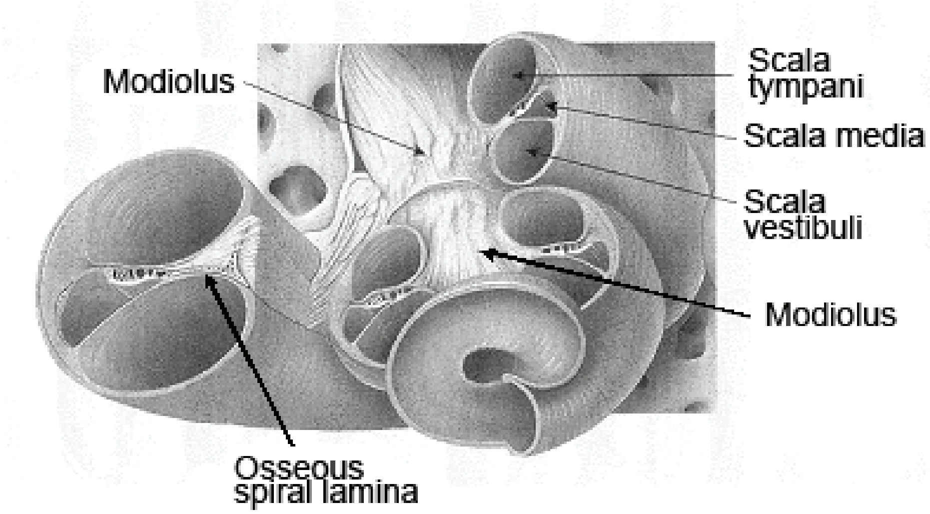


Figure 8-12. View of the cochlea (adapted from Emanuel and Letowski, 2009).

the *fenestra rotunda*. The membrane of the oval window has a surface area of 3.2 to 3.5 mm^2, is completely covered by the footplate of the stapes, and is sealed in the bony opening by the annular ligament. The round window has a surface area of approximately 2 mm^2 and is located inferior and anterior to the oval window in the wall between the middle ear and inner ear and serves as a pressure valve between the two scalae. When an acoustic stimulus causes mechanical vibration of the stapes footplate this movement is translated to the membrane of the oval window. The membrane pushes back and forth on the perilymph of the scala vestibuli and through the helicotrema, the perilymph of the scala tympani. This motion results in alternating outward and inward movement of the membrane of the round window. The membrane bulges outward as the fluid moves from the scala vestibuli to the scala tympani and bulges inward as the fluid moves from the scala tympani to the scala vestibuli.

The central most duct of the cochlea is the membranous scala media (cochlear duct). This duct separates the scala tympani and scala vestibuli. The border between the scala media and the scala vestibuli is *Reissner's membrane* and the border between the scala media and the scala tympani is the *basilar membrane.* Reissner's membrane (vestibular membrane) is attached to the osseous spiral lamina and projects obliquely from it to the outer wall of the cochlea forming a roof of the scala media. Together, the basilar and Reissner membranes would make the scala media into a closed tube if not for the *ductus reunions,* a tiny opening at its base that connects it to the saccule (Figure 8-11). .

The basilar membrane forms the floor of the scala media. The membrane is anchored to the spiral lamina on one end and to the spiral ligament on the other end. When uncoiled, the membrane has an approximate length of 32 to 35 mm (1.25 to 1.38 in), practically the same as the whole cochlea. Blood vessels and nerve fibers supporting the cochlea enter the cochlea through the modiolus and spiral lamina.

The spiral lamina is narrower at the apex of the cochlea and wider at the base. Conversely, the basilar membrane is wider, thicker and more flaccid at the apex and narrower, thinner, and stiffer at the base. These factors affect the vibrational properties of the basilar membrane, which responds to low frequency vibrations at the apex and high frequency vibrations at the base of the cochlea.

The organ of hearing (organ of Corti) is located primarily on the basilar membrane, with a small segment projecting onto the spiral lamina. The organ of Corti is made up of sensory cells (hair cells) and supporting cells. A schematic cross section of the cochlea showing the content of the scala media and the structure of the organ of Corti are shown in Figure 8-13.

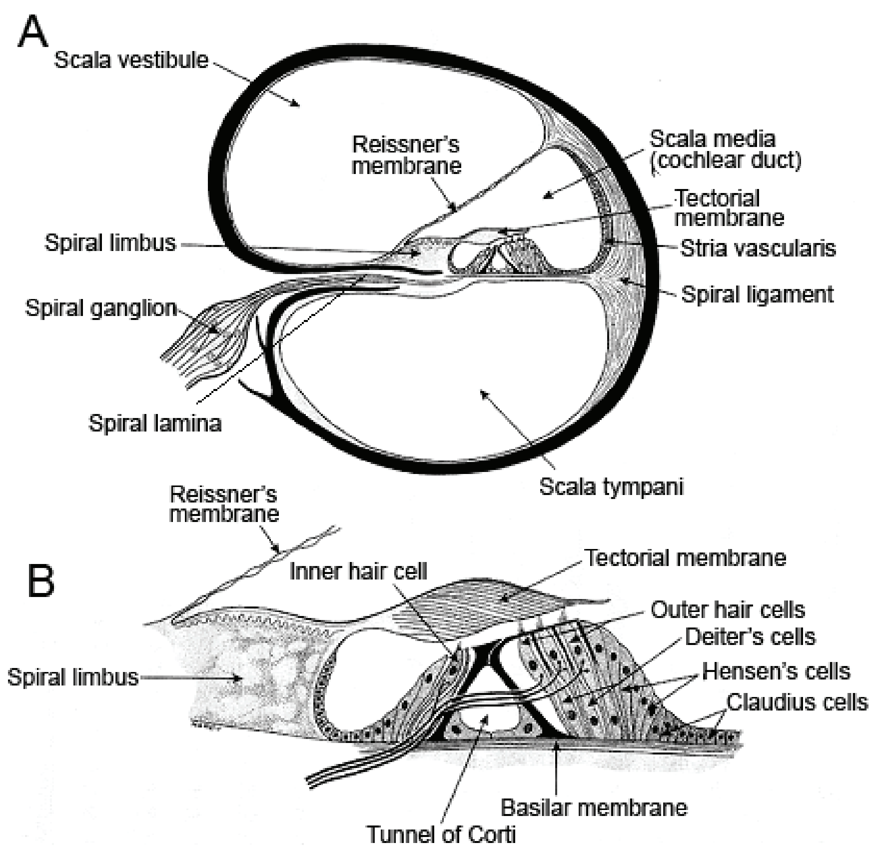


Figure 8-13. Cross section of the cochlea (A) and the structure of the organ of Corti (B) (Emanuel and Letowski, 2009).

Organ of Corti

The organ of Corti converts the mechanical vibrations of the basilar membrane into neural impulses that then travel through the auditory nerve and brainstem to the brain. The organ is composed of sensory cells, called *hair cells*, and several types of supporting cells distributed along the length and width of the basilar membrane. The arrangement of the sensory cells and supporting cells on the basilar membrane is shown in Figure 8-13B. The fibers of the auditory nerves travel from the organ of Corti through a system of small perforations in the spiral lamina collectively called *habenula perforata* that start at the tympanic edge of the spiral lamina and continue further through it. From habenula perforata, nerve fibers travel through a channel in the center of the modiolus (Rosenthal's canal), exit the base of the cochlea, and join vestibular nerve fibers to form the vestibulocochlear nerve.

There are two types of hair cells in the organ of Corti: the inner hair cells (IHCs) and the outer hair cells (OHCs). They are shown in Figure 8-14. Each hair cell has a number of small hair-like projections called *stereocilia* (cilia) extending from the top of the cell. The group of stereocilia at the top of a hair cell is called a *stereocilia bundle*. The stereocilia bundle of each hair cell is organized in several rows forming either a "W" or "V" pattern for OHCs and shallow "U" pattern for IHCs. Stereocilia in each row have graduated heights (like stair steps) and their tips are connected together by thin fibers called *tip links.* Each type of hair cell in the ear is connected to the nervous system by both afferent (ascending) and efferent (descending) nerve endings, but the number and function of these types of connections varies between IHCs and OHCs.

There are altogether approximately 3500 IHCs and approximately 12,000 OHCs distributed along each basilar membrane (thus, each ear contains between 15,000 and 16,000 hair cells). The IHCs are shaped like a flask and form a single row of cells supported by the spiral lamina. The OHCs have a cylindrical shape with a diameter of approximately 9 μm and are organized into three rows located farther away from the spiral lamina. The groups of IHCs and OHCs are separated by two rods (pillars) of Corti, which structurally support the organ of Corti. The inner rod rests on the spiral lamina while the outer rod is attached to the basilar membrane. The rods are attached at their tops and more widely separated at the base, forming a triangular shape called the *tunnel of Corti*. The tunnel is filled with the *cortilymph fluid* that has similar properties to the perilymph fluid found in the bony labyrinth. The tops of the hair cells and supporting cells of the organ of Corti are tightly connected together at their tips to form a continuous layer called the *reticular lamina.* The reticular lamina isolates all of the organ of Corti from the endolymph of the scala media except for stereocilia which project through the reticular lamina into the endolymph.

The OHCs are held in position by the outer rod of Corti on one side and by Deiters cells on the other side. Each Deiters cell holds an OHC at the bottom and through long projections called *phalangeal processes* from above. The middle part of an OHC is not firmly supported and is surrounded by a perilymph-filled space called the *space of Nuel*.

Next to the Deiters cells, moving towards the outer end of the cochlea, there are several groups of supporting cells, called Hensen cells, Claudius cells, outer spiral sulcus cells, and Boettcher cells. Several of these cells are shown in Figure 8-13B. Lateral to these support cells is the *Stria vascularis*, a highly vascular organ attached to the outer surface of the scala media. Stria vascularis recycles potassium and produces endolymph for the scala media, thus maintaining the endocochlear potential (battery) of the inner ear.

The IHCs are structurally supported by the inner rod of Corti on one side and by inner sulcus and pharyngeal cells on the other (Lim, 1986). The inner sulcus cells occupy the region extending from IHCs toward the spiral limbus. The spiral limbus projects from the spiral lamina towards the organ of Corti and provides the attachment point for the *tectorial membrane*. The tectorial membrane is a gelatinous membrane extending above the organ of Corti from the upper surface of the spiral limbus. The largest stereocilia of the OHCs make contact with the tectorial membrane, and this connection is part of the mechanism that leads to the neural responses of the organ of Corti. The basic characteristics of the OHCs and the IHCs are summarized in Table 8-4.

Vestibular system

The vestibular system groups together make up the peripheral organs of balance. The bony structures include three semicircular canals and the vestibule. Within the bony structure, the three semicircular canals contain the three membranous semicircular ducts and the vestibule contains the utricle and saccule. In evolutionary terms, the organs of balance are much older than the organ of hearing, which actually evolved from them (Zemlin, 1997). The arrangement of the components of the vestibular system is shown in Figure 8-11.

Table 8-4.
Summary of the basic characteristics of OHCs and IHCs.

Characteristic	Outer Hair Cells	Inner Hair Cells
Number of hair cells	12,000	3500
Location of hair cells	Further from modiolus	Nearer modiolus
Number of rows	Three to four	One
Shape of hair cells	Cylindrical shape	Flask shape
Number of rows of cilia	6-7 rows per cell	2-4 rows per cell
Stereocilia arrangement	“W” or “V” shape	Shallow “U” shape
Length of stereocilia	Longer and thinner	Shorter and fatter Length of stereocilia increases along the basilar membrane and varies within a single cell
Cell body restriction	Largest cilia contact the tectorial membrane	Cilia are free at the upper end and not in contact with any structure
Motility	Motile	Not motile
Efferent innervation	80% of the efferent fibers terminate directly at the cell bodies of the hair cells Connect to efferent fibers from the medial superior olive Efferent fibers synapse on the base of the hair cell	20% of the efferent fibers synapse with the afferent fibers on the hair cells Connect to efferent fibers from the lateral superior olive Efferent fibers synapse on afferent nerve but not the cell body
Afferent innervation	One afferent fiber supplies multiple hair cells 5% of the afferent fibers supply the hair cells	8-10 afferent fibers supply a single inner hair cell 95% of the afferent fibers supply the hair cells

The three semicircular canals are hoop-shaped structures connected at both ends to the vestibule. One of the ends is almost twice as wide as the other and is called the *ampulla*. The canals are perpendicular to each other and are called the horizontal (lateral), the anterior (superior), and the posterior (inferior) semicircular canal. The horizontal canal is located not exactly horizontally but forms a 30° angle relative to the horizon. The spatial arrangement of the canals is shown in Figure 8-15. Each semicircular canal is sensitive to head motion in the plane of that canal. The canals also form bilateral differential pairs between the ears (e.g., right anterior with left posterior which have their hair cells aligned oppositely). Rotation in one plane will be excitatory to one canal and inhibitory to the other.

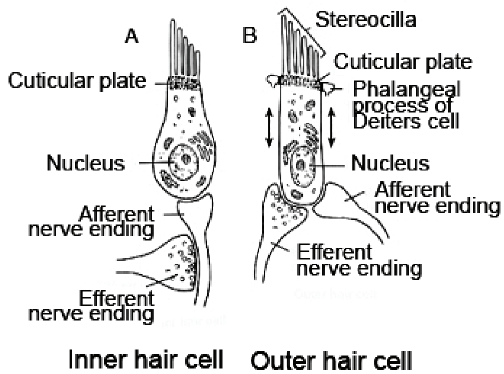


Figure 8-14. Shape and structure of the inner (A) and outer (B) hair cells (Emanuel and Letowski, 2009).

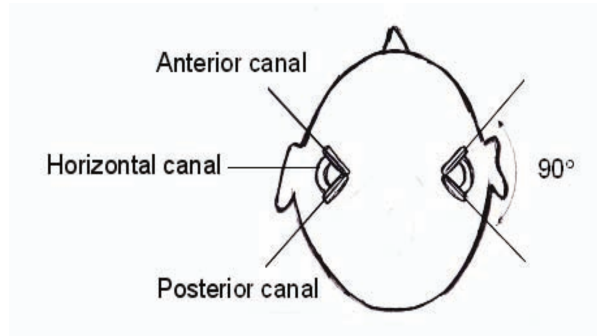


Figure 8-15. Spatial arrangement of the semi-circular canals.

The semicircular canals are filled with perilymph and the semicircular ducts are filled with endolymph. The ampulla (bulge) of each canal contains the *crista ampulliaris*, a saddle-shaped, raised section of wall that is populated with the hair cells whose stereocilia respond to angular acceleration. The stereocilia of the semicircular canals are embedded in a gelatinous mass called a *cupula* that is similar in function to the tectorial membrane of the organ of Corti, that is, its motion bends the hair cells, which then creates a neural impulse.

While the sensory organs in the semicircular canals respond to angular accelaration of the head, two other organs of balance, the sense organs within the utricle and the saccule, respond to gravity and linear acceleration in horizontal (utricle) and vertical (saccule) directions. The sense organs within the utricle and saccule are the maculae. They occupy the concave spaces at the bottom of the utricle and the saccule and contain tiny pieces of calcium carbonate, called *otoliths* (ear stones) or otoconia (ear dust), which are embedded into a gelatinous membrane (*otolithic membrane*) into which the stereocilia of the maculae project. Since the otoliths are quite numerous in the otolithic membrane and they are heavier then the surrounding fluid, the membrane gets displaced towards the Earth during head tilting (due to gravity) and away from the source of motion during linear acceleration (due to inertia); thus, head motions are translated by stereocilia deflection into neural impulses.

There are two types of hair cells in the semicircular canals and the vestibule. Type I hair cells are flask-shaped cells while type II hair cells are cylinder-shaped cells. Type I and type II hair cells are very similar in their structure and innervation to the inner hair cells and the outer hair cells of the organ of Corti, respectively. Each hair cell in the semicircular canals has 50 to 100 small stereocilia and a single larger cilium called a *kinocilium,* which only exists in rudimentary form in the hair cells of the cochlea. The stereocilia are arranged by length, with the longest stereocilia located close to the kinocilium, and are all connected by tip links. Movement of the stereocilia hair bundle toward the kinocilium causes a depolarizing (excitatory) sensory response whereas movement away from the kinocilium causes a hyperpolarization (inhibitory) sensory response.

Bone Conduction System

The inner ear receives mechanical vibrations and converts them into neural impulses by the organ of Corti. As such it can be stimulated by sounds transmitted through the system of outer and middle ear <u>or</u> by skull vibrations. The first mode of stimulation is called air conduction, while the second mode is referred to as bone conduction.

The anatomical system responsible for transmitting skull vibrations to the organ of Corti is called the bond conduction system.

The bone conduction system can be stimulated by either sound waves impinging on the human head or by delivering a vibratory signal by means of a mechanical driver (vibrator) coupled to the head. In the former case, the resulting stimulation is approximately 1000 times (60 decibels [dB]) weaker than the simultaneous air conduction stimulation. This is due to the air-bone mismatch when sound tries to enter the bones of the skull (Chapter 9, *Auditory Function*). Vibration of the whole head may also cause distractive interference between vibrations delivered to various parts of the skull. Therefore, such stimulation only has practical meaning in the case of people wearing hearing protectors or with severe conductive hearing loss (Henry and Letowski, 2007). In the case of people wearing audio HMDs (Chapter 5, *Auditory Helmet-Mounted Displays*) or heavily sound-attenuating head gear, even though their air conduction pathways may be completely blocked, they may still hear very intense sounds, such as explosions, jet engine sounds, pile driver impact sounds, etc. due to the stimulation received through bone conduction.

When the vibrations are delivered directly to the head by a mechanical driver, the driver-bone is much smaller than the air-bone mismatch and the person may hear even very weak vibrations of the driver. Thus, this mode of stimulation can be effectively used for speech communication, human-robot interaction, and delivery of tactical signals during military operations.

In order to understand the potential applications and limitations of bone conduction hearing it is necessary to understand the basic elements of the human head and how they interact with one another. The typical weight of the human head is approximately 3.5 kg and the basic dimensions of the male and female heads are given in Table 8-3. The head is a complex structure made of bones, cartilage, several types of soft tissue, and fluids (e.g., cerebrospinal fluid), which differ in their mechanical properties. These different forms of matter transmit sound with different speeds and with various degrees of attenuation. Densities for select components of the head together with their associated speeds of sound transmission are listed in Table 8-5. According to Evans and Lebow (1951) and Sauren and Classens (1993) the average density of the bones of the skull is 1412 kg/m^3, the Young modulus is 6.5×10^9 N/m^2, and the Poisson ratio[2] is 0.22. All these values, however, are highly dependent on the amount of water in the bones and other tissue and the boundary conditions. The more dry and less constrained the bone is, the higher the speed of sound through the bone. In addition, solid matter, like the bone, can simultaneously propagate longitudinal, transverse (traveling), and surface waves that have different speeds and can interact with one another.

The bones of the skull are listed in Table 8-1 and their arrangement is shown in Figure 8-2. The manner in which the skull and associated tissue respond to mechanical stimulation depends on the point of stimulation and the frequency of the signal. Two typical driving points described in the literature are the mastoid process and the forehead. The distance from the mastoid process to the cochlea is approximately 30 mm (1.2 in) and is the shortest distance between the cochlea and the head periphery (Tonndorf and Jahn, 1981). The main stimulation pathway from the mastoid process lies wholly within the respective temporal bone, which results in relatively low attenuation of the initial stimulus. In addition, the direction of this stimulation is the same as that of the air conduction pathway, which causes elements of the latter pathway to also become excited by bone stimulation.

The forehead is the relatively flat surface of the frontal bone, which is the largest bone of the skull. It is a fairly symmetrical and deeply extended bone that can easily transfer its vibration to many other bones of the skull. Another effective bone conduction pathway in the skull is from the condyle of the mandible (jaw bone), which is located on the side of the head just in front of the entrance to the ear canal. Stimulation of condyle activates bones and cartilage surrounding the ear canal and tympanic cavity creates a secondary pathway through the air conduction system even more effectively than stimulation at the mastoid process. The most common excitation points for bone conduction communication are shown in Figure 8-3.

---

[2] Poisson's ratio is the ratio of the relative contraction strain, or transverse strain (normal to the applied load), divided by the relative extension strain, or axial strain (in the direction of the applied load).

Table 8-5.
Material components of the human head with their densities and corresponding speeds of sound transmission. (O'Brien and Liu, 2005)

Material	Speed of Sound (m/s)	Density (kg/m^3)
Air	340	1.2
Water	1500	1000
Soft Tissues	1520-1580	980-1010
Lipid-based tissues	1400-1490	920-940
Collagen-based tissues	1600-1700	1020-1100
Blood	1580	1040-1090
Brain – grey	1532-1550	1039
Brain – white		1043
Skull – compact inner and outer tables	2600-3100	1900

Depending on the frequency of stimulation, the skull has several modes of vibration that differ in the phase relationship between the vibrations at different locations on the skull. At low frequencies, e.g. 200 Hz, the skull driven at the forehead location moves as a whole in a back and forth pattern (Bekesy, 1932). This type of vibration is called inertial vibration and its corresponding mode of vibration is called the inertial mode. At higher frequencies, e.g., 800 Hz, the direction of vibration remains the same, but the front and the back of the skull vibrate 180° out-of-phase. Vibration where different parts of the skull vibrate out-of-phase is referred to as compressional vibration. Out-of-phase vibration of the front and back of the head is called the first compressional mode. More complex compressional modes are elicited at even higher frequencies. Vibration patterns for the inertial and first two compressional modes of vibration are illustrated in Figure 8-16. A more detailed discussion of how different points and frequencies of stimulation affect bone conduction hearing is included in Chapter 9, *Auditory Function*.

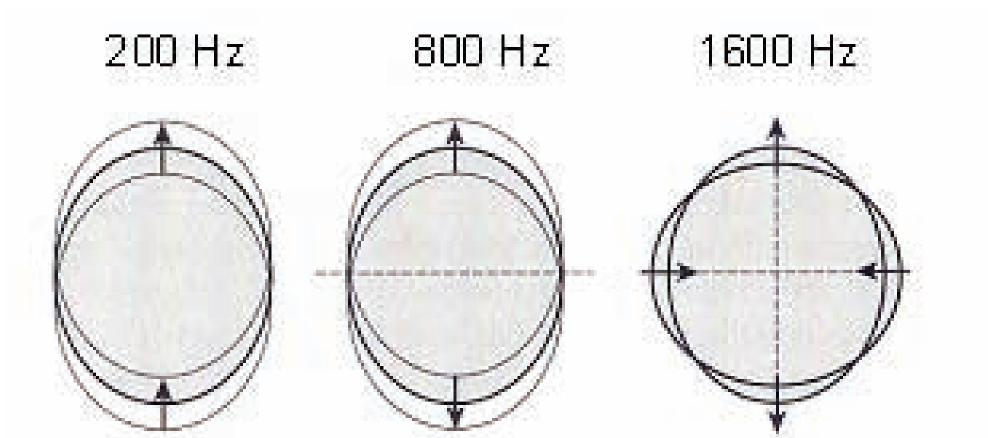


Figure 8-16. Modes of bone vibration at three different frequencies (adapted from Bekesy, 1932).

Vestibulocochlear Nerve

The vestibulocochlear nerve (the VIII cranial nerve) is the nerve connecting the organs of hearing and balance to the brain. It consists of two parts: the *auditory nerve* (cochlear nerve) and the *vestibular nerve*. Sensory signals that produce sensations of sound travel from the ear to the brainstem via the auditory nerve. The ascending neurons of the auditory nerve innervate the hair cells of the organ of Corti, exit the organ of Corti through the habenula perforata, and then form the *spiral ganglion* in the modiolus. The ascending neurons, called the *afferent neurons*, carry information from the sensory cells toward the brain. The descending neurons, called the *efferent neurons*, carry information from the brain to the sensory cells and other cells of the peripheral nervous system.

A neuron consists of a neuron cell and input (dendrites) and output (axon) projections extending from the neuron cell. These projections are called nerve fibers. Depending upon their location and function, neurons may be any length from 1 inch to 4 feet (2.5 cm to 1.2 m) long. Neurons transmit electrochemical signals to and from the brain via their nerve fibers. Upon receiving a signal, one neuron sends information to its adjacent neuron, through a junction called a synapse

The nerve fibers of the auditory nerve originate all along the cochlea, from its apex to its base, and project to the cell bodies of these nerves, which form the spiral ganglion. The fibers extending from the apex follow a straight course and form the core of the spiral ganglion, while the fibers from the base are twisted to form the outside surface of the ganglion. After leaving the cochlea, the auditory nerve joins the vestibular nerve, another bundle of fibers supporting the vestibular system, and they together form a bundle of approximately 30,000 afferent and efferent nerve fibers called the vestibulocochlear nerve. The vestibulocochlear nerve exits the inner ear through the *internal auditory meatus*, approximately 1 cm long channel in the temporal bone, which also houses the facial nerve, and enters the brainstem, where the auditory and vestibular parts of the vestibulocochlear nerve separate and take different pathways through the central nervous system. The structural representation of the vestibulocochlear nerve is shown in Figure 8-17.

The ascending pathways of the vestibulocochlear nerve that support the hair cells of the organ of Corti involve two types of afferent neurons: *inner radial neurons* (type I afferent neurons) and *outer spiral neurons* (type II afferent neurons). The inner radial neurons constitute approximately 95% of the ascending neurons in the cochlea and the outer spiral neurons make up the remaining 5% (Gelfand, 1998). The inner radial neurons, which are myelinated (insulated with a fatty substance) and larger of the two, innervate the IHCs. The innervation pattern is many-to-one and approximately 8 to 10 afferent fibers supply one IHC (Gelfand, 1998). The outer spiral neurons, which are unmyelinated and thinner, innervate the OHCs. The innervation pattern is one-to-many with one neuron making synapse connections with approximately 10 OHCs (Gelfand, 1998).

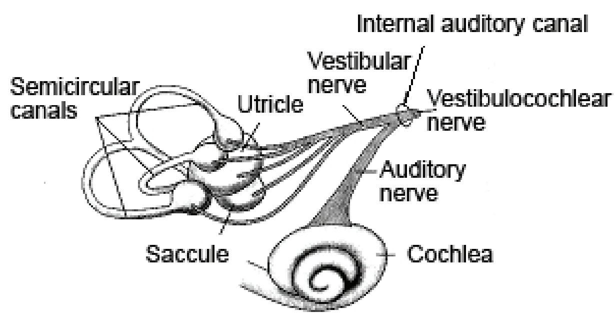


Figure 8-17. General structure of the vestibulocochlear nerve (adapted from Lass and Woodford, 2007).

Similar to the ascending pathways, the descending pathways of the vestibulocochlear nerve also involve two types of efferent neurons. They are the lateral olivocochlear neurons and the medial olivocochlear neurons, both of which descend from the superior olivary complex in the brainstem. The lateral olivocochlear neurons are myelinated and the larger, more numerous of the two, and they synapse with the projections of the afferent neurons connected to the IHCs. They constitute approximately 20% of the efferent neurons in the cochlea. The remaining 80% of efferent neurons in the cochlea are medial olivocochlear neurons (Gelfand, 1998). They are thin and unmyelinated and synapse to the OHCs. The distribution of the efferent fibers on the OHCs heavily favors the base of the cochlea. A schematic view of the innervation pattern for IHCs and OHCs is shown in Figure 8-18.

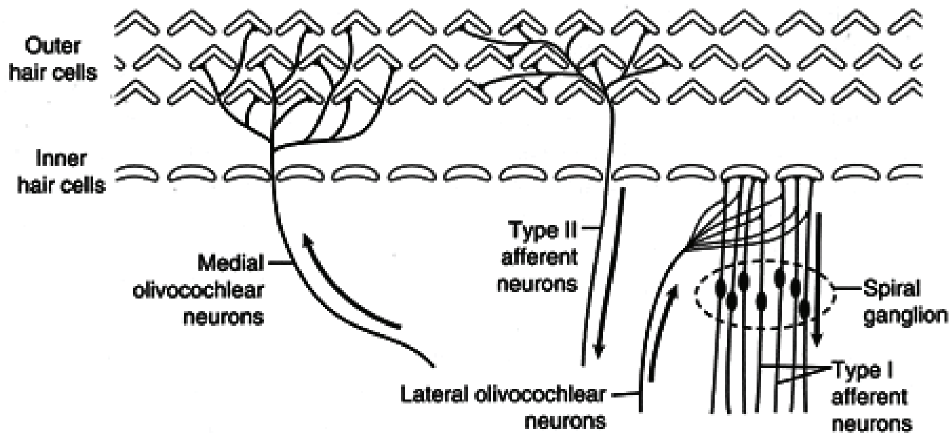


Figure 8-18. The innervation pattern of the hair cells of the organ of Corti (Emanuel and Letowski, 2009).

Central Auditory Nervous System

The central auditory nervous system (CANS) is a system of neural structures and connections within the brain that processes neural impulses transmitted from the vestibulocochlear nerve and converts them into auditory sensations. It is a subsystem of the *central nervous system* (CNS), which includes the entire brain and the spinal cord. The CNS is a dynamic system composed of various types of nerve cells (neurons), which form an extraordinary network of neural connections reaching out from the brain to every part of the body. The human brain is estimated to contain 100 billion (10^{11}) neurons[3] and a quadrillion (10^{15}) synapses[4] (Kimball, 2005). The anatomical organization of the main structural elements of the human brain is shown in Figure 8-19.

The most inferior (lowest) portion of the brain is the *brainstem*. The brainstem is approximately 10 cm long and 2.5 cm wide at the central core (Seikal, King, and Drumright, 2000). It is the superior extension of the spinal cord and the place where the vestibulocochlear nerve enters the brain. The main anatomical elements of the brainstem are the *medulla oblongata*, *pons*, and *midbrain*. The medulla oblongata, pons, and *cerebellum* form the posterior (back) part of the brain called the *hindbrain*. The midbrain is the most superior (top) part of the brainstem and it is connected to and located just below the *forebrain* (*cerebrum*), the largest and the most advanced part of the brain. The main parts of the forebrain are the telencephalon (including the *cerebral hemispheres*, basal nuclei, and medullary center of nerve fibers) and the diencephalons (including the *thalamus*, and *hypothalamus).*

Neural pathways in the CANS consist of various nuclei (groups of cell bodies) and fiber tracts (bundles of nerve fibers) which carry information between and among the nuclei. Each nucleus serves as a relay station for dispatching neural information from one nucleus to the next. The neurons comprising a specific neural pathway travel through several nuclei in the brainstem before reaching the auditory cortex. The nuclei involved in the classical ascending auditory pathway are the cochlear nucleus, superior olivary complex, inferior colliculus, and medial geniculate body. Neural fibers carrying specific information may synapse with nuclei on the same side or

---

[3] A *neuron* is a cell specialized to conduct and generate electrical impulses and to carry information from one part of the brain to another.
[4] A *synapse* is the junction between two nerve cells across which a nerve impulse is transmitted.

decussate (cross from one side to the other) and synapse with nuclei on the other side of the brainstem. The pathway that connects the nuclei on the same side of the brainstem is called the *ipsilateral pathway* and the pathway that crosses from one side to the other is called the *contralateral pathway*. A general view of the ascending auditory pathway is shown in Figure 8-20. Note the size of the lateral lemniscus, which is the largest fiber tract in the CANS.

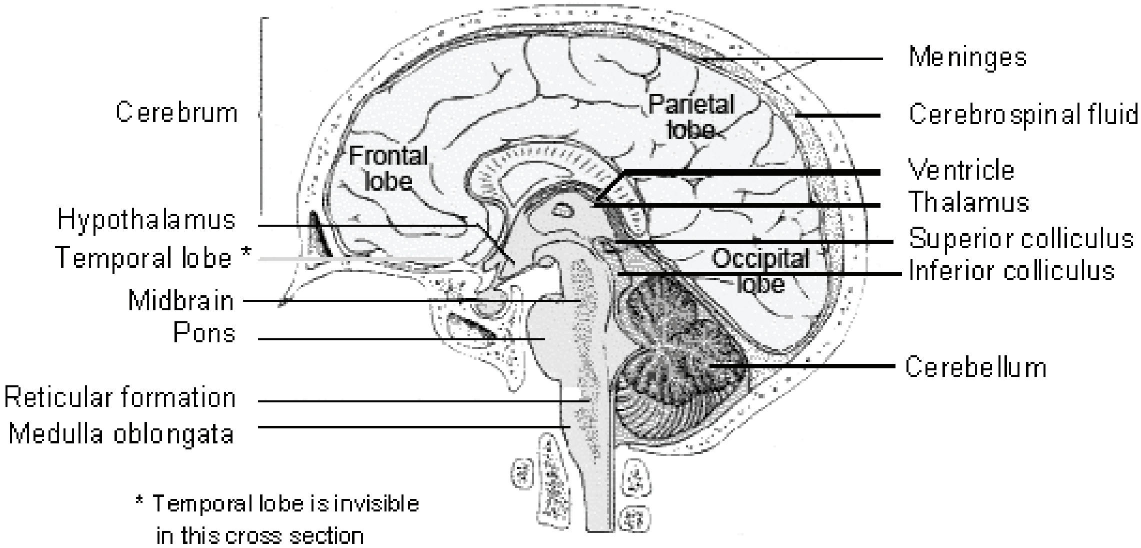


Figure 8-19. Main elements of the human brain (adapted from Kimball, 2005).

The *cochlear nucleus,* which spans the pons and medulla, is the first processing center and relay station of the ascending auditory pathways. The two other major relay stations in the brainstem are the superior olivary complex in the pons and the inferior colliculus in the midbrain. From the cochlear nucleus, the nerve fibers project to the *superior olivary complex* (SOC) or directly to the lateral lemniscus. Approximately 75% of the ascending CANS fibers leaving the cochlear nucleus cross over to the contralateral side of the brain to terminate at the SOC on the opposite side of the brainstem or project to the lateral lemniscus. The remaining 25% of the fibers follow the pathway on the ipsilateral side of the brainstem and terminate at the SOC or the lateral lemniscus (Pickles, 1988). The SOC consists of a cluster of nuclei including the lateral superior olive (LSO), the medial superior olive (MSO), and the ventral nucleus of the trapezoid body (VNTB). The SOC is also the site at which afferent auditory neurons connect with facial nerve, which innervate the stapedius muscle in the middle ear. When a very intense sound traveling via the vestibulocochlear nerve arrives at the cochlear nucleus, the signal is sent to SOC via several different pathways (including ipsilateral and contralateral SOC) to the facial nerve nucleus. From the nucleus the signal travels via the facial nerve to the stapedius muscle, which contracts. Contraction of the stapedius muscle pulls the stapes posteriorly increasing stiffness of the ossicles and tympanic membrane and decreasing the effective level of transmission for loud low frequency sounds (Deutsch and Richards, 1979; Moller, 1965; Stach and Jerger, 1990; Wilber, 1976). In addition, VNTB projects to many cochlear nuclei and forms efferent medial olivocochlear bundle (MOCB) innervating ipsilateral and contralateral OHCs (Guinan, 2006; Warr and Beck, 1996) decreasing gain of the cochlear amplifier (see Chapter 9, *Auditory Function*).

The organization of superior olivary complex and the connecting neural fibers are shown in Figure 8-21. Ascending projections from both cochlear nuclei and both superior olivary complexes travel via the largest fiber tract in the CANS, the lateral lemniscus, to the inferior colliculi (one colliculus on each side), located on the posterior surface of the midbrain. Similar to the decussation seen at earlier levels in the brainstem, the two *inferior colliculi* are connected by fibers that allow crossover of signals from one side of the brainstem to the other. The connections between the two sides of the brainstem, from the SOC to the inferior colliculi, are important for directional hearing. From the inferior colliculi, all fibers ascend to the *medial geniculate body* in the thalamus. The thalamus is located immediately above the midbrain and it directs all sensory information (except smell) to the appropriate area of the cerebrum. The cerebellum, or "little brain," is primarily responsible for coordinating motor commands with sensory inputs in order to control movement and communicates with the brainstem, spinal cord and cortex.

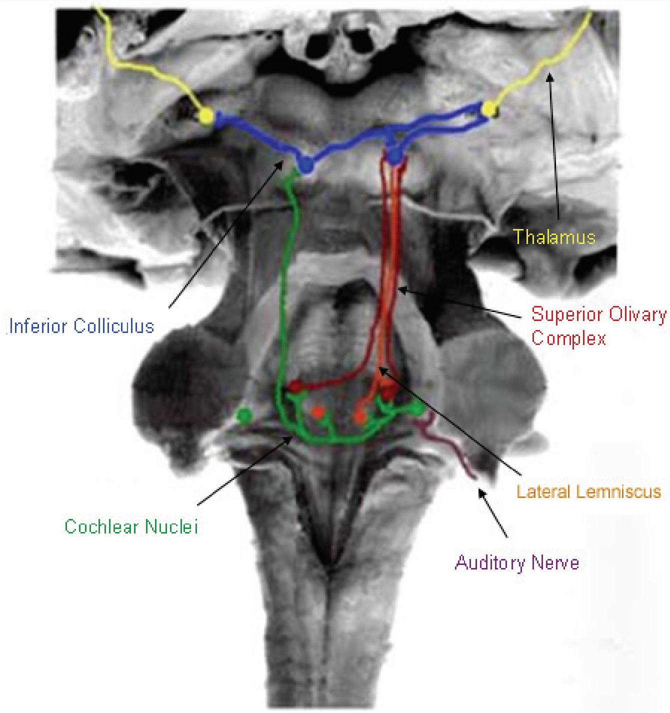


Figure 8-20. Ascending pathway of the central auditory nervous system. Different parts of the auditory pathway are color coded (adapted from http://serous.med.buffalo.edu/hearing/).

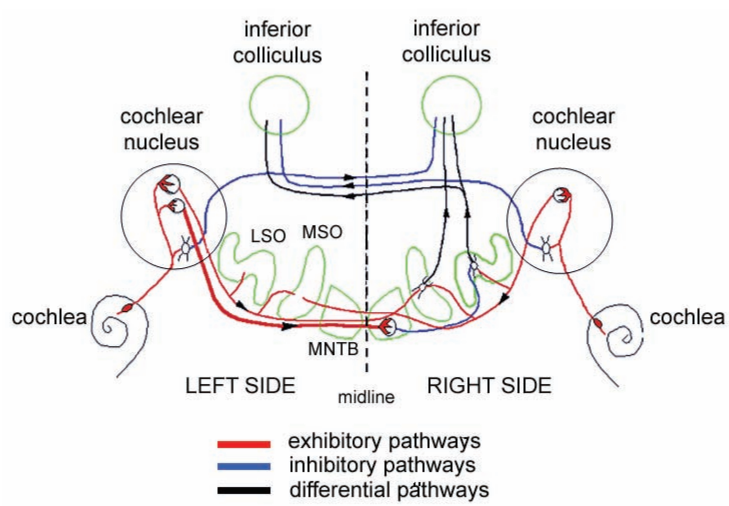


Figure 8-21. Nuclei and pathways of the left and right complex olivary complexes; LSO - lateral superior olivary nucleus, MSO - medial superior olivary nucleus, MNTB – medial nucleus of the trapezoid body (adapted from Johnson, 1997)

The cerebral hemispheres make up the largest portion of the forebrain. There are a number of connections between the medial geniculate body and the cerebral hemispheres, but the main ascending auditory pathway travels from the medial geniculate nucleus to the ipsilateral transverse temporal gyri (Heschl's gyrus) of the cerebrum and then to auditory association areas in other areas of the brain

The outermost segment of the cerebrum is called the cerebral cortex; commonly referred to as the "gray matter." The cortex is 2-6 mm (0.08 to 0.23 in) thick and is made up of the "gray-looking" nerve cell bodies. It is supported from underneath by the "white matter." which consists of the myelinated nerve fibers (axons) connecting various gray matter areas of the brain to each other. The surface of the cortex contains numerous peaks (gyri) and valleys (sulci) that serve to increase the overall area of the cortex. Extremely deep sulci are called fissures. The deepest fissure of the brain, the *longitudinal fissure* divides the cerebrum into two cerebral hemispheres. Cerebral hemispheres are only connected by a narrow structure called the *corpus callosum,* which is the only communication link between the hemispheres.

Each hemisphere is divided into four basic anatomical areas called lobes. They are called the frontal lobe, temporal, parietal, and occipital lobes. The frontal lobe takes up 1/3 of the cortex and is associated with executive functions such as the planning and initiation of motor actions. The parietal lobe is the primary reception area for somatic sensory data, while the occipital lobe is the main visual processing center of the brain (Seikel, King and Drumright, 2000). The main site of auditory and receptive language (Wernicke's area) processing centers is the temporal lobe. The four lobes are additionally divided into smaller functional areas based on the type and organization of neurons occupying these areas. These areas are called Brodmann areas and numbered from 1 to 48. Many of them have also been found to be responsible for specific cortical activities and so are labeled by these activities. For example, auditory activity in the cortex has been found to be concentrated in Brodmann areas 41 and 42, which are called the *primary auditory cortex*, and in area 22, called the *secondary auditory cortex*. Both these regions are located in the posterior (back) part of the *superior temporal gyrus* and descend into the *lateral sulcus* (Sylvain fissure) as the *transverse temporal gyri*, known also as the *Heschl's gyri*. A schematic view of various part of the cortex with a map of the Brodmann areas is shown in Figure 8-22.

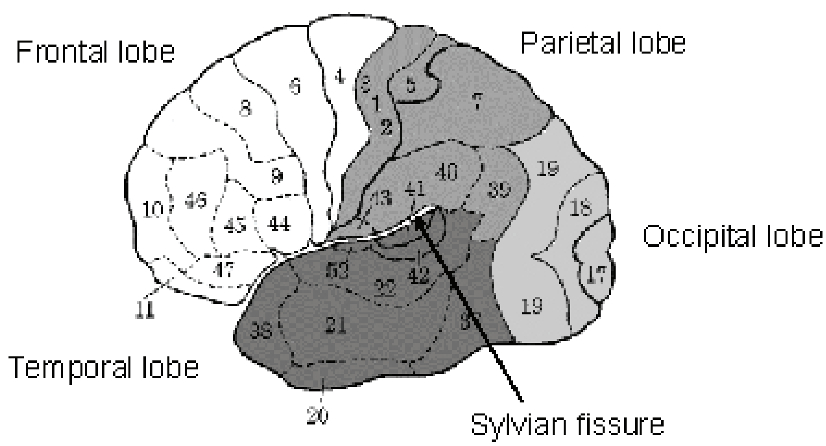


Figure 8-22. Sagittal view of the cerebral cortex and the Brodmann areas (adapted from http://www.umich.edu/ ~cogneuro/jpg/Brodmann.html).

The cortex, and thus the auditory cortex, is organized in six neural layers numbered from I to VI (Emanuel and Letowski, 2009). Auditory information arriving at the thalamus is further relayed to nonpyramidal neurons located in layer IV of the primary auditory cortex. Layers V and VI have efferent connections to the medial geniculate nucleus and the inferior colliculus, respectively. Other layers are involved in motor function (layer II and III) and have connections to other parts of the brain. Through theses connections all information entering the brain creates a synergistic perceptual image of the surrounding environment together with corresponding emotional state created by this image.

References

Alexander, M., and Laubach, L.L. (1968). Anthropometry of the human ear. Aerospace Medical Research Laboratories, Report TR-67-203. Dayton, OH: Wright Patterson Air Force Base.

Algazi, V.R., Duda, R.O., and Thompson, D.M. (2001). The CIPIC HRTF Database. *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics.* New Paltz, NY: IEEE.

Alvord, L.S., and Farmer, B.L. (1997). Anatomy and orientation of the human external ear. *Journal of the American Academy of Audiology*, 8, 383-390.

Ballachanda, B.B. (1995). Cerumen and the ear canal secretory system. In: Ballanchanda, B.B. (Ed.). *The Human Ear Canal: Theoretical Considerations and Clinical Applications Including Cerumen Management*. San Diego: Singular Publishing Group, Inc.

Békésy, G. von. (1932). Über den Einfluss der durch den Kopf und den Gehörgang bewirkten Schallfeldverzerrungen auf die Hörschwelle. Ann. Physik 14, 51-56. [Also in Békésy, G. von (1960). Effects of the head and external auditory meatus on the sound field. In: *Experiments in Hearing* (translated and edited by Wever, E.G.), p. 267-271. New York: McGraw-Hill].

Békésy, G. von. (1949). The structure of the middle ear and the hearing of one's own voice by bone conduction. *Journal of the Acoustical Society of America*, 21, 217-232.

Buckingham, R., and Valvassordi, G. (2001). Inner fluid volumes and resolving power of magnetic resonance imaging: Can it differentiate endolymphatic structures? *Annals of Otology, Rhinology, and Laryngology*, 110, 113-117.

Burkhard, M.D., and Sachs, R.M. (1975). Anthropometric manikin for acoustic research. *Journal of the Acoustical Society of America*, 58, 214-222.

Churchill, E., and Truett, B. (1957). Metrical relations among dimensions of the head and face. WADC Technical Report 56-621. ASTIA Document No. AD 110629 (June). Dayton (OH): Wright Air Development Center.

Dacraemer, W.F., Maes, M.A., and Vanhuyse, V.J. (1980). An elastic stress-strain relation for soft biological tissues based on a structural model. *Journal of Biomechanics,* 13, 463-468.

Decraemer, W.F., Dirckx, J.J., and Funnell, W.R. (1991). Shape and derived geometrical parameters of the adult, human tympanic membrane measured with a phase-shift moiré interferometer. *Hearing Research,* 51, 107-121.

Dallos, P. (1973). *The Auditory Periphery*. New York: Academic Press.

Department of Defense. (2000). *Human Engineering Design Data Digest*. Washington, D.C.: Department of Defense.

Despopoulos, A., and Silbernagl, S. (1991). *Color Atlas of Physiology* (4th Ed.). New York: Thieme Medical Publishers.

Deutsch, L.J., and Richards, A.M. (1979) *Elementary Hearing Science*. Boston: Allyn and Bacon.

Djupesland, G., and Zwislocki, J.J. (1971). Effect of temporal summation on the human stapedius reflex. *Acta Otolaryngologica,* 71, 262-265.

Donaldson, J.A., and Miller, J.M. (1980). Anatomy of the ear. In: *Otolaryngology*, Pararella, M., and Shumrick, D. (Eds.), *Basic Sciences and Related Disciplines*, 1, 26-42. Philadelphia: Sounders.

Dreyfus, H. (1967). *The Measure of Man*. New York: Whitney Library of Design.

Emanuel, D., and Letowski, T. (2009*). Hearing Science*. Baltimore: Lippincott, Williams, and Wilkins.

Evans, F., and Lebow, M. (1951). Regional differences in some of the physical properties of the human femur. *Journal of Applied Physiology,* 3, 563-572.

Fumagalli, Z. (1949). Recherche morfologische sull' apparato di transmissione del suono (Morphological research of the sound-transmitting apparatus). *Archives of Italian Otology, Rinology and Laringology*, 60(1), 322.

Gelfand, S.A. (1998). *Hearing: An Introduction to Psychological and Physiological Acoustics*. New York: Marcel Decker.

Glasscock, M.E., and Shambaugh, G.E. (1990). *Surgery of the Ear* (4th Ed.). Philadelphia: W.B. Sounders.

Gopen, Q., Rosowski, J.J., and Merchant, S.N. (1991). Anatomy of the normal human cochlear aqueduct functional implications. *Hearing Research*, 107, 9-22.

Gray, H. (1918). *Anatomy of the Human Body,* 20th edition. Philadelphia: Lea and Febiger. Retrieved 16 October 2008 from: www.bartleby.com/107/

Guinan, J.J. (2006). Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans. *Ear and Hearing*, 27, 589-607.

Hall, J.W., and Mueller, H.G. (1997). *Audiologists' Desk Reference*. San Diego: Singular Publishing Group, Inc.

Harris, J.D. (1986). *Anatomy and Physiology of the Peripheral Auditory Mechanism*. Austin, TX: Pro-Ed.

Hawkins, D., Alvarez, E., and Houlihan, J. (1991). Reliability of three types of probe microphone measurements. *Hearing Instruments,* 42, 14-16.

Henry, P., and Letowski, T. (2007). Bone conduction: Anatomy, physiology, and communication. ARL Technical Report ARL-TR-4138. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Howell, P., Williams, M., and Dix, H. (1988). Assessment of sound in the ear canal caused by movement of the jaw relative to the skull. *Scandinavian Audiology,* 17, 93-98.

International Electrotechnical Comission (IEC). (2009). IEC 60318-7, Electroacoustics – Simulators of human head and ear – Part 7: Head and torso simulator for acoustic measurements of air-conduction hearing aids (Revision of IEC/TR 60959:1990). Geneva: International Electrotechnical Commission.

International Electrotechnical Commission (IEC). (2008). ISO 7250-2 Basic human measurements for technical design – Part 2: Statistical summaries of body measurements from individual ISO populations. Geneva: International Organization for Standardization.

Johnson, D.H. (1997). The binaural pathway. Rice University. Retrieved 10 January 2009 from http://www-ece.rice.edu/!dhj/binaural.html

Kimball, J. (2005). Kimball's Biology Pages: The human nervous system. Retrieved 25 June 2005 from: http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/C/CNS.html

Kojo, Y. (1954). Morphological studies of the human tympanic membrane. *Journal of Oto-Rhino-Laryngological Society of Japan,* 57, 115-126 (in Japanese).

Lee, K.J. (2008). *Essential Otolaryngology and Head and Neck Surgery* (9th Ed). New York: McGraw-Hill Publishing.

Lim, D.J. (1970). Human tympanic membrane: An ultrastructural observation. *Acta Otolaryngologica,* 70, 176-186.

Lim, D.J. (1986) Functional structure of the organ of Corti: A review. *Hearing Research,* 22, 117-146.

Liu, T-C., and Chen, Y-S. (2000). Aging and external ear resonance. *Audiology,* 39, 235-237.

Lucente, F.E. (1995). Anatomy, histology, and physiology. In: *The External Ear*, Lucente, F.E., Lawson, W., and Novick, N.L. (Eds.). Philadelphia: W.B. Sounders Company.

McBride, M., Letowski, T., and Tran, P. (2005).Bone conduction head sensitivity mapping: Bone vibrator. ARL Technical Report ARL-TR-3556, July 2005.

McBride, M., Letowski, T., and Tran, P. (2008). Bone conduction reception: Head sensitivity mapping. *Ergonomics,* 55(5), 702-718.

McDonald, K. (1993). Surgical procedures of the external ear canal and ear. Grand Rounds paper. Houston, TX: Department of Otolaryngology, University of Texas. Retrieved 16 October 2008 from: http://www.ndcrt.org/data/Health_facts/Auricle

Moller, A. (1965). An experimental study of the acoustic impedance of the middle ear and its transmission properties. *Acta Otolaryngologica*, 60, 129-149.

Moore, K.L., and Dalley, A.F. (1999). *Clinical oriented anatomy* (4th Ed.). Baltimore: Lippincott Williams and Wilkins.

Muller, C. (2003). Tympanoplasty. Grand Rounds paper. Houston: Department of Otolaryngology, University of Texas. Retrieved 16 October 2008 from: http://www.ndcrt.org/data/Health_facts/Auricle.html

National Aeronautics and Space Administration. (1978). *Anthropometric Source Book* (Vol. 1-3). Yellow Springs, OH: NASA.

O'Brien, W., and Liu, Y. (2005). Evaluation of acoustic propagation paths into human head. *Proceedings of the RTO-MP-HFM-123 Meeting "New Directions for Improving Audio Effectiveness,"* p. 15-1 to 15-24 (paper 15). Neuilly-sur-Seine (France): RTO.

Pickles, J.O. (1988). *An Introduction to the Physiology of Hearing*. Berkeley, CA: Academic Press.

Pumford, J., and Sinclair, S. (2001). Real-ear measurement: basic terminology and procedures. Audiology Online (May 7, 2001). Retrieved 16 October 2008 from: http://www.audiologyonline.com/articles/arc_disp?article_id=285

Rordriguez-Vasquez, J.F., Merida-Velasco, J.R., Merida-Velasco, J.A., and Jimenez-Collado, J. (1998). Anatomical considerations on the discomalleolar ligament. *Journal of Anatomy*, 192, 617-621.

Rohen, J.W., and Yokochi, C. (1983). *Color Atlas of Anatomy. A Photographic Study of the Human Body*. New York: Igaku-Shoin.

Sauren, A., and Classens, M. (1993). Finite element modeling of head impact: The second decade. *Proceedings of the International IRCOBI Conference on Biomechanics of Impacts*. Eindhoven (The Netherlands). September 8-10.

Sclafani, A., and Ranaudo, J. (2006). Otoplasty. *The Medscape Online Journal (eMedicine)*. Retrieved 16 October 2008 from: http://www.emedicine.com/ent/topic110.thm

Seikel, J.A., King, D.W., and Drumright, D.G. (2000). *Anatomy and Physiology for Speech, Language, and Hearing*. San Diego: Singular Publishing Group.

Shaw, E.A.G. (1974). The external ear. In: *Handbook of Sensory Physiology (Vol. VI): Auditory System*, Keidel, W.D., and Neff, W.D. (Eds.). New York: Springer-Verlag.

Shennib, A., and Urso, R. (2000). Direct tympanic membrane excitation via vibrationally conductive assembly. United States Patent No. 6137889.

Stach, A.B., and Jerger, F.J. (1990). Immittance measures in auditory disorders. In: Jacobson, J.T., and Northern, J.L. (Eds.), *Diagnostic audiology*. Needham Heights, MA: Allyn and Bacon.

Stinson, M.R., and Lawton, B.W. (1989). Specification of geometry of the human ear canal for the prediction of sound-pressure level distribution. *Journal of the Acoustical Society of America*, 85, 2492-2503.

Sundberg, M. (2008). Optical methods for tympanic membrane characterization. Linköping Studies in Science and Technology, Dissertation No. 1173. Linköping (Sweden): Liu-Tryck.

Teranishi, R., and Shaw, E.A.G. (1968). External ear acoustic models with simple geometry. *Journal of the Acoustical Society of America*, 44, 257-263.

Warr, W.B., and Beck, J.E. (1996). Multiple projections from the ventral nucleus of the trapezoid body in the rat. *Hearing Research*, 93, 83-101.

Waver, E., and Lawrence, M. (1954). *Physiological Acoustics*. Princeton, NJ: Princeton University Press.

Wilber, A.L. (1976). Anatomy and physiology of the middle ear. In: Feldman, A.S., and Wilber, L.A. (Eds.), *Acoustic impedance and admittance: The measurement of middle ear function*. Baltimore: Williams and Wilkins.

Yost, W., and Nielsen, D. (1977). *Fundamentals of Hearing: An Introduction*. New York: Holt, Rinehart, and Winston.

Zemlin, W.R. (1997). *Speech and Hearing Science: Anatomy and Physiology* (4th Ed.). Boston: Allyn and Bacon.

Zwislocki, J.J. (1970). An acoustic coupler for earphone calibration. Special Report LCS-S-7. Syracuse, NY: Syracuse University.

Zwislocki, J.J. (1971). Ear-like coupler for earphone calibration. Special Report LSC-S-9. Syracuse, NY: Syracuse University.

Zwislocki, J.J. (2002). Auditory system: Peripheral nonlinearity and central additivity in the human stapedius-muscle reflex. *Proceedings of the National Academy of Sciences U.S.A,* 99, 14601-14606.

9 AUDITORY FUNCTION

Diana C. Emanuel
Sumalai Maroonroge
Tomasz R. Letowski

Physiology and Function of the Hearing System

The hearing system, also called also the auditory system, consists of the outer ear, middle ear, inner ear, and central auditory nervous system. The overall function of the hearing system is to sense the acoustic environment thus allowing us to detect and perceive sound. The anatomy of this system has been described in Chapter 8, *Basic Anatomy of the Hearing System*. The current chapter describes the function and physiology of the main parts of the hearing system in the process of converting acoustic events into perceived sound.

In order to facilitate perception of sound, the hearing system needs to sense sound energy and to convert the received acoustic signals into the electro-chemical signals that are used by the nervous system. A schematic view of the processing chain from the physical sound wave striking the outer ear to the auditory percept in the brain is shown in Figure 9-1.

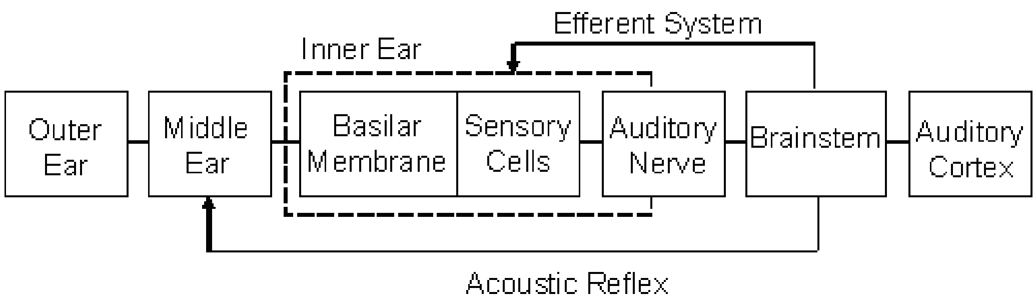


Figure 9-1. A schematic view of the hearing system.

The hearing system shown in Figure 9-1 has two functions: sound processing and hearing protection. Sound processing by the hearing system starts when the sound wave arrives at the head of a person. The head forms a baffle that reflects, absorbs, and diffracts sound prior to its processing by the hearing system. The first two sound processing elements of the hearing system are the outer and middle ears that form together a complex mechanical system that is sensitive to changes in intensity, frequency, and direction of incoming sound. Acoustic waves propagating in the environment are diffracted, absorbed, and reflected by the listener's body, head, and the pinnae and arrive through the ear canal at the tympanic membrane of the middle ear. After the acoustic wave strikes the eardrum, its acoustic energy is converted into mechanical energy and carried across the middle ear. At the junction of the middle ear and the inner ear, the mechanical energy of the stapes is transformed into the motion of the fluids of the inner ear and thence into the vibrations of the basilar membrane. The motion of the basilar membrane affects electro-chemical processes in the organ of Corti and results in generation of electric impulses by the array of the hair cells distributed along this membrane. The electrical impulses generated by the hair cells affect the inputs to the nerve endings of the auditory nerve and are transmitted via a network of nerves to the auditory cortex of the brain where the impulses are converted into meaningful perception.

A secondary function of the hearing system is to provide some protection for the organ of Corti and the physical structures of the middle ear from excessive energy inputs and subsequent damage by modulating the reactivity of the mechanical linkages. The anatomy of the outer ear also protects the tympanic membrane from harmful effects of wind, dust, and changes in temperature and humidity while the muscles of the middle ear provide some protection of the inner ear and organ of Corti. The text of this chapter covers the sound transmission function of the hearing system but the main protective structures of the ear also will be discussed as they are mentioned.

The Outer Ear

Directional properties of the hearing system

As sound waves arrive at a listener's head, the energy of sound entering the ear is affected by the presence of the human body and by the acoustic properties of the outer ear. Some sounds are attenuated and reflected away by the barriers caused by the head structures while others are reflected toward the ear canal and even amplified by the ear cavities. The shape of the head and of the upper torso and the locations and shapes of the two pinnae serve as a direction cueing system that modifies incoming sound depending on the location of the sound source. The difference between the sound arriving at the listener and the sound that enters the ear canal is called the Head Related Transfer Function (HRTF) and it varies as a function of the direction from which sounds arrive and with the frequency of the sound (see Chapter 5, Audio *Helmet-Mounted Display Design* and Chapter 11, *Auditory Perception and Cognitive Performance*).

The directional system of the human head operates throughout full three-dimensional spherical space and is sensitive to a wide range of acoustic frequencies. Directional cues generated by the hearing system generally can be divided into binaural and monaural cues, depending whether they involve both ears or just one ear. The two main binaural cues are the interaural time difference (ITD) and the interaural intensity difference (IID).

If a sound source is located in the median sagittal plane (midline, dividing right and left sides) of the head, the two ears receive approximately the same acoustic signal. However, if the sound is approaching the head from one side, the ear closest to the sound source will receive the sound earlier and with greater intensity than the other ear. ITDs are the differences in the onset of sound and are equivalent to phase differences in the case of continuous periodic sound with no perceived onset. IIDs are caused by the absorption and reflection of incoming sound by the body and head structures and creation of an "acoustic shadow" affecting the ear farther away from the sound source. The ITD cues operate most effectively at low auditory frequencies whereas the IID cues are most effective at high frequencies but fail at low frequencies since these sound waves diffract around the human body.

Binaural cues are supported by monaural cues resulting from the specific positions and shapes of the two pinnae. The pinna has the shape of an irregular funnel that is attached to the head at an angle of 15 to 30° (see Chapter 8, *Basic Anatomy of the Hearing System*). Its main function is to collect sound and to channel it to the ear canal. However, the frontal orientation of the pinna favors sounds coming from the front and helps to differentiate between the sounds arriving from various locations along the front to back axis. In addition, the configuration of the ridges and depressions on the surface of the pinna provides a complex system of resonating cavities and reflecting surfaces that differently affects sounds arriving from various locations along both the vertical axis and the horizontal plane. The effect of pinna reflection on the same sound arriving from different vertical directions is shown in Figure 9-2. Depending on the angle of sound arrival, different ridges of the pinna are involved in sound reflections causing angle-dependent changes in the overall acoustic spectrum of the sound entering the ear canal (Batteau, 1967; Hebrank and Wright, 1974; Lopez-Poveda and Meddis, 1996; Roffler and Butler, 1968).

The relatively small dimensions of the pinna and its features compared to the wavelengths of sound perceived by humans cause the directional function of the pinna to operate primarily in the mid- and high-frequency regions of perceived sound (Wright, Hebrank, and Wilson, 1974). Low-frequency sounds have wavelengths longer than the dimensions of the pinna and are easily diffracted. As a result, the locations of low-frequency sound sources are difficult to localize using pinna mechanisms, that is, along top-down and front-back axes.

The brain uses monaural cues for sound localization on the vertical plane and both monaural and binaural cues for sound localization on the horizontal plane. Both binaural and monaural cues are additionally enhanced by different positions of both pinnae on the head and in relation to the torso, which contributes to the three-dimensional directional characteristic of the human head by causing time delays and intensity changes for both direct and body-reflected sounds entering the ears. In addition, the different patterns of pinna convolutions in the left and right ear of each human affect the directional properties of the head, creating a very unique and non-transferable HRTF for each individual. The detailed discussions of the directional properties of the human auditory system and the limitations of the specific directional cues are presented in Chapter 11, *Auditory Perception and Cognitive Performance*.

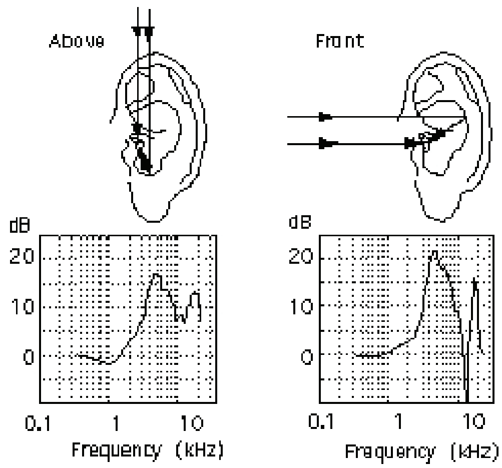


Figure 9-2. Sound spectra at the ear canal of the same sound arriving from two different directions (Duda, 2000).

Selective amplification of sound

After the pinna collects, modifies and channels sound toward the ear canal, the sound is further altered by the resonances of the concha and ear canal. As the sound enters the concha and the ear canal, the sounds of some frequencies are relatively amplified while others are correspondingly suppressed resulting in distinct spectral shaping of the incoming sound by the outer ear. The function of selective amplification of sound by external ear cavities is to enhance the sounds that are important to human behavior and speech communication.

The cavity of the ear canal forms a tube that acts as a ¼ wavelength resonator. This type of resonator enhances sounds of certain frequency and damps others depending on the relationship between the wavelength of sound and the length of the resonator. A ¼-wavelength resonator increases sound pressure at the blocked end of a tube for sound waves that have a wavelength four times the length of the tube and any odd whole number multiple of this wavelength. The resonance frequencies of the ¼ wavelength resonator can be calculated as

$$f_n = \frac{(2n-1)c}{4L} \qquad \text{Equation 9-1}$$

where f_n is an n^{th} resonance frequency of the resonator, c is the speed of sound, L is the length of the tube, and n is the resonance frequency number.

The average effective length of the ear canal (which is the ear canal plus some of the depth of concha) is approximately 30 mm. This means if the ear canal were a hard walled tube with uniform cross sectional diameter, the average ear canal would increase the relative sound pressure at the tympanic membrane at approximately 2833 hertz (Hz) (assuming standard temperature and pressure conditions such that c = 340 meters/second [m/s]). However, the ear canal is lined with soft tissue, does not have a uniform cross-sectional diameter, and varies in size based on age, gender, and genetic factors. Therefore, the specific resonance characteristics of the ear canal vary with people and age. The specific amplification effects of the ear canal and other parts of the external ear measured by Shaw (1974) are shown in Figure 9-3.

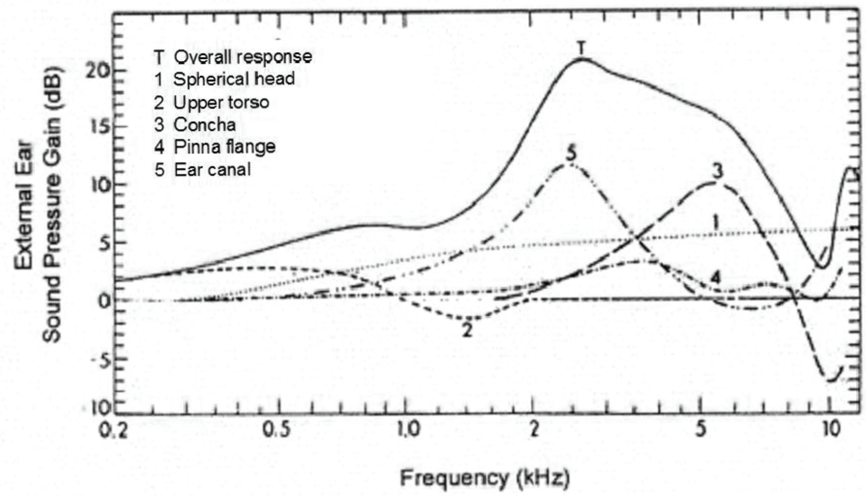


Figure 9-3. The sound pressure gain at the tympanic membrane due to the contribution of the outer ear, head, and body. Direction of sound arrival: azimuth: 45°; elevation: 0° (adapted from Shaw, 1974).

Examination of Figure 9-3 indicates that the resonance properties of the concha increase the sound pressure around 5,000 Hz, the helix and antihelix provide some lesser amount of amplification across a broader frequency range, and the largest resonance is provided by the ear canal. All together the peak of the resonance of the outer ear occurs between 2,000 and 3,000 Hz and is between approximately 15 and 20 decibels (dB). This is the frequency region that is most important for speech communication.

The functions presented in Figure 9-3 show amplification characteristics for sound waves propagation in the horizontal plane and arriving at a 45° angle relative to the front of the head. At this angle of arrival the overall sound pressure gain caused by the head and outer ear is the greatest. For other angles of arrival, the specific gain functions have slightly different shapes providing information approximately the direction of the incoming sound. Some examples of the overall gain functions for sounds propagating in the horizontal plane and arriving at different azimuth angles are shown in Figure 9-4.

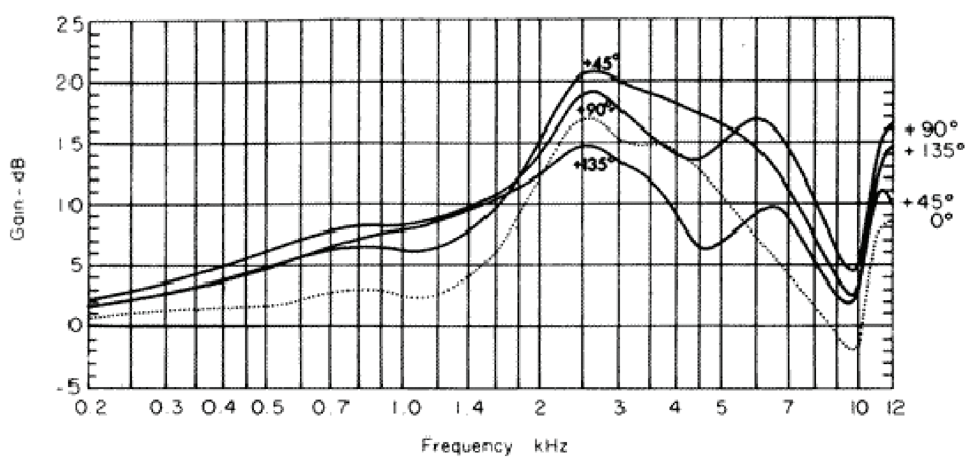


Figure 9-4. Sound pressure gain functions for sound waves propagating in the horizontal plane and arriving at different azimuth degree at the listener ear (Pickles, 1988).

In addition to the frequency-specific selective filtering of sound, the anatomical features of the ear canal (e.g. a long, curved, narrow tunnel) provide a protective barrier between the outside world and the middle ear structures. The length and shape of the ear canal serve to isolate the tympanic membrane from the changes in external temperature, humidity, and the effects of wind. They also make it difficult for dust, small flies, and other debris to reach the tympanic membrane. The skin of the ear canal contains hair follicles and glands that secrete the oily substances that protect the canal from drying and act to repel dust particles and insects (see Chapter 8, *Basic Anatomy of the Hearing System*). Even the shape, position and cartilaginous form of the pinna contribute to hearing system protection by providing a cushion against physical impact to the head.

The Middle Ear

Acousto-mechanic transduction

The primary function of the middle ear is to act as an impedance matching element between the air-filled outer ear and the fluid-filled inner ear. Impedance is the opposition of a system to the energy flow through the system, e.g., to a change in the velocity of motion, and defines the ability of the system to store and transfer energy. Impedance is a vector quantity that has two parts, resistance (real part) and reactance (imaginary part), that are responsible for the transfer and storage of energy, respectively. The transfer of energy from one system to another is most efficient when both systems have the same impedance (Emanuel and Letowski, 2009).

When sound waves reach the tympanic membrane their acoustic energy is converted into physical vibrations of the ossicular chain, which is attached at one end to the tympanic membrane to the membrane covering the oval window of the inner ear at the other, providing an anatomical bridge between the outer and inner ears. Without this conversion of sound energy into mechanical energy, the amount of energy delivered to the inner ear would be significantly less than the amount of energy arriving at the tympanic membrane.

When the tympanic membrane vibrates in response to changes in sound pressure in the ear canal, its vibration is confined primarily to the *pars tensa*, which constitutes approximately two-thirds of the membrane's surface (see Chapter 8, *Basic Anatomy of the Hearing System*). However, both parts of the tympanic membrane, the *pars tensa* and the *pars flaccida*, are responsible for the tension of the membrane. The tension of the tympanic membrane directly affects hearing sensitivity. If the membrane is too flaccid, most sound energy is absorbed by the membrane itself, that is, it goes into stretching the membrane. If the membrane is too tense, too much sound energy is reflected back into the environment. Therefore, the tympanic membrane must be at the appropriate tension for sound energy to be converted efficiently into the mechanical movement of the ossicles in the middle ear.

The middle ear is normally filled with air, and under normal operation conditions, the static air pressure in the middle ear is the same as the atmospheric pressure in the ear canal. Equal air pressure on both sides of the tympanic membrane is needed to establish proper tension on the membrane. The pressure in the middle ear cavity is maintained by the periodic opening and closing of the eustachian tube (auditory tube). This tube connects the middle ear with the nasopharynx (back of the throat) and can be opened or closed by the action of the tensor veli palatine muscles (muscles from the velum and palate). The tube is normally closed, but it pops open during yawning and swallowing. A swelling of the surrounded tissue can cause tube malfunction and consequently a lack of proper air pressure in the middle ear. If the air pressure in the middle ear cavity is significantly different from the pressure in the ear canal, this may cause over- or under-stretching of the tympanic membrane, which leads to inefficient sound transmission, pain and can also produce middle ear diseases.

Impedance matching

The acousto-mechanical transformation of the received sound energy serves to match the high acoustic impedance [1]of the fluid-filled inner ear with the low acoustic impedance of the air in which sound waves propagate and to optimize energy transfer between these two systems. In order to calculate the actual mismatch one needs to know the input impedance of the oval window and the impedance of the source from which the sound impinges on the window (Killion and Dallos, 1979). The acoustic input impedance of the oval window has been calculated by Zwislocki (1975) to be approximately 350,000 acoustic Ω [dyne · s/cm^5]. This value is based on Békésy's low frequency impedance data corrected for postmortem effects (Békésy, 1942, 1949, 1960). At higher frequencies this impedance is probably closer to 1,200,000 acoustic Ω measured for a cat's ear (Lynch, Nedzelnitsky and Peake, 1982). Both these impedances are much higher than the characteristic impedance of air, which is approximately 41.5 centimeter-gram-second system (cgs) rayls [dyne · s/cm^3] at 30°C temperature.

The specific acoustic impedance of air in the ear canal is the characteristic impedance of air normalized by the cross-sectional area of the canal (0.45 cm^2) and equal approximately 100 acoustic Ω (Shaw, 1974, 1997; Zwislocki, 1957, 1970, 1975). The power transmission index η describing power transmission from the ear canal to the tympanic membrane has been reported to be approximately η=0.75 and fairly similar across all mammals (Hemilä, Nummela and Reuter, 1995; Møller, 1974; Voss and Allen, 1994). Thus, assuming that both the input impedance of the tympanic membrane (Z_t) and the characteristic acoustic impedance of air in the ear canal (Z_{ec}) are resistive, the relationship between Z_t, Z_{ec}, and η can be written as:

$$\eta = \frac{4 Z_t Z_{ec}}{(Z_t + Z_{ec})^2} \qquad \text{Equation 9-2}$$

If Z_{ec}=100 acoustic Ω, and η = 0.75, then the impedance of the tympanic membrane equals $Z_t = 3Z_{ec} = 300$ acoustic Ω, and the impedance ratio of the inner ear fluid (cat's data) and the tympanic membrane is approximately 4000.

In order to ensure an efficient transfer of energy between the acoustic system of the ear canal and the hydraulic system of the inner ear, the middle ear must compensate for this mismatched impedance by increasing the pressure between the tympanic membrane and oval window by approximately 63 times ($63^2 \sim 4000$). This is equal to a 36 dB increase in sound pressure level (SPL). In other words, the pressure acting on the fluids in the inner ear must be 36 dB higher than pressure acting on the tympanic membrane to ensure the most efficient transfer of acoustic energy to the inner ear. This is the role of the middle ear transformer, consisting of the tympanic membrane, ossicles, and oval window membrane. Without the impedance matching function of the middle ear, more than 99.9% of the acoustic energy acting on the tympanic membrane would be reflected by the tympanic membrane back into the ear canal and not used. If the human middle ear matching function is not functioning properly, sound can only be transmitted via a shunt pathway (tympanic membrane to the air in the middle ear to the fluid of the inner ear), which results in the transmission of less than 0.1% of the input energy.

The impedance matching system of the middle ear consists of three separate mechanisms:

1. Area ratio transformation
2. Ossicular chain lever action
3. Catenary lever action

[1] Acoustic impedance is defined as the ratio of effective acoustic pressure averaged over a given surface to effective volume velocity of acoustic energy flowing through this surface. The units for impedance are Pa-s/m^3 or dyne-s/cm^5, which are called the acoustic ohm (Ω).

All three mechanisms contribute to the overall pressure transformation, however, the first mechanism, the area ratio transformer, is the most essential to the impedance matching process. The presence of the last mechanism, the catenary lever action, is still somewhat controversial, and there is some disagreement in the literature regarding its contribution (Goode, 2006).

Area ratio transformation

The area ratio (pressure) transformer is the first and most effective of the three impedance matching mechanisms. It results from the difference in surface area between the tympanic membrane and the membrane covering the oval window. The principle of this mechanism is illustrated in Figure 9-5.

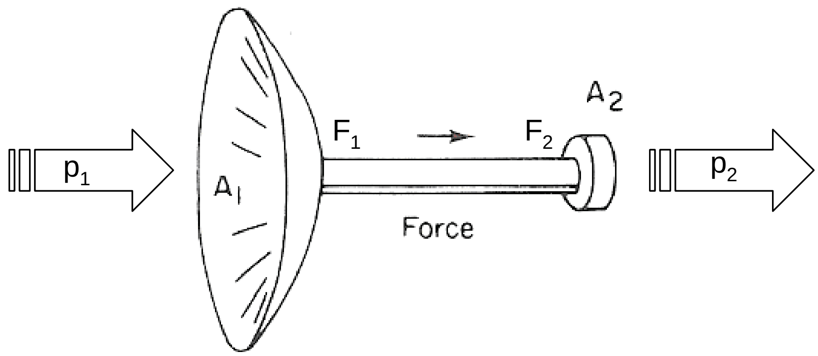


Figure 9-5. Schematic drawing of surface area mismatch between tympanic membrane and oval window membrane (adapted from Pickles, 1988).

In Figure 9-5, a pressure p_1 acts over the surface of the tympanic membrane and results in a force F_1. Assuming that the ossicular chain is a lossless system, the force F_2 acting on the oval window is equal to force F_1, that is, $F_1 = F_2 = F$. Since force (F), surface area (A), and pressure (p) are related by the equation p=F/A, then

$$F = p_1 \times A_1 = p_2 \times A_2 \qquad \text{Equation 9-3}$$

and

$$p_2 = p_1 \times \frac{A_1}{A_2}. \qquad \text{Equation 9-4}$$

Since the vibrating area of the tympanic membrane ($A_1 = 55mm^2$) is approximately 17.2 times larger than the vibrating area of the oval window membrane ($A_2 = 3.2mm^2$), this results in an increase in SPL at the oval window of approximately 25 dB.

Ossicular chain lever action

The second impedance matching mechanism of the middle ear, the ossicular chain lever action, involves the rotational motion between the malleus and stapes. This kind of motion is possible because the ossicles are fixed at the junction between the malleus and incus while being suspended in the middle ear cavity by the anterior ligament of the malleus (anteriorly) and the posterior ligament of the incus (posteriorly). This arrangement creates a central pivot point (fulcrum) and allows for the relative rotational motion of the malleus and stapes, thereby forming a lever mechanism. The principle of this mechanism is shown in Figure 9-6.

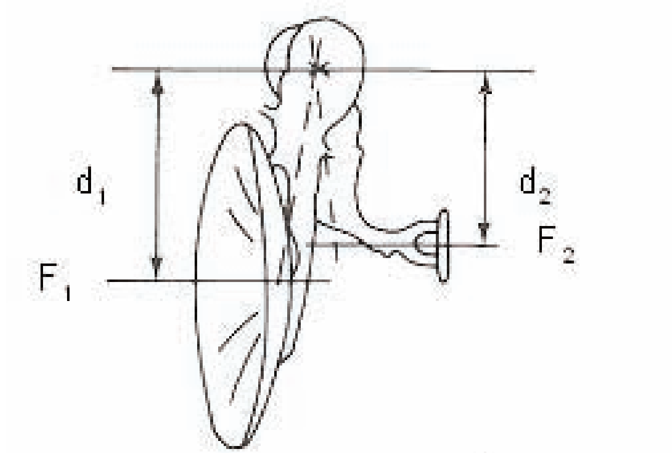


Figure 9-6. Schematic drawing of ossicular chain level action (adapted from Pickles, 1988).

In a lever system, a force F_1 applied at effort arm d_1 results in force F_2 acting on effort arm d_2, that is

$$F_1 \, x \, d_1 = F_2 \, x \, d_2. \qquad \text{Equation 9-5}$$

In the case of the ossicular chain lever, the forces F_1 and F_2 are the forces acting at the malleus and stapes and the distances d_1 and d_2 are the lengths of the malleus and stapes, respectively. Since the length of the malleus is approximately 1.3 times longer that the length of the stapes, this increases the force between the tympanic membrane and the oval window membrane by approximately 2 dB (Békésy, 1941; Wada, 2007). It should be noted that some authors recommend using 1:1.15 (1.2 dB increase) rather than 1:1.3 (2. dB increase) ratio in order to compensate for the fact that the malleus and the tympanic membrane act as a coupled system (Battista and Esquivel, 2003; Pickles, 1988).

Catenary lever action

The third impedance matching mechanism, the catenary lever action, curved membrane effect, or buckling effect of the tympanic membrane, was first explained by Helmholtz (1868), who observed that the umbo of the tympanic membrane is displaced less than the remaining surface of the tympanic membrane. Since the outside edge of the membrane is firmly attached to the annulus and curves medially to attach to the umbo, the displacement of the membrane between the annulus and umbo is larger than at the umbo (Khanna and Tonndorf, 1970; Tonndorf and Khanna, 1970). This creates a lever action which increases the force acting at the umbo by approximately 2 times or 6 dB (Rosowski, 1996). The principle of this mechanism is shown in Figure 9-7.

Overall transformation

The 63 (36 dB) ratio needed to compensate for the air-to-cochlea impedance mismatch is called in the literature the *ideal transformer prediction* (Rosowski, 1996) but is not fully realized by the middle ear system. The three impedance matching mechanisms together increase the sound pressure at the footplate of the stapes by approximately 40 to 45 times (32 to 33 dB) in comparison to the sound pressure acting on the tympanic membrane. This increase is approximately 3 to 4 dB short of completely making up for the impedance mismatch. However, the 32 to 33 dB effectiveness of impedance matching mechanisms agrees with physiological findings that completely disconnected ossicular chain causes a hearing loss (due to the air-bone gap) of approximately 32 dB (Rosowski, Mehta and Merchant, 2004), indicating that the transformer model described above adequately represents the real-world operation of the middle ear system.

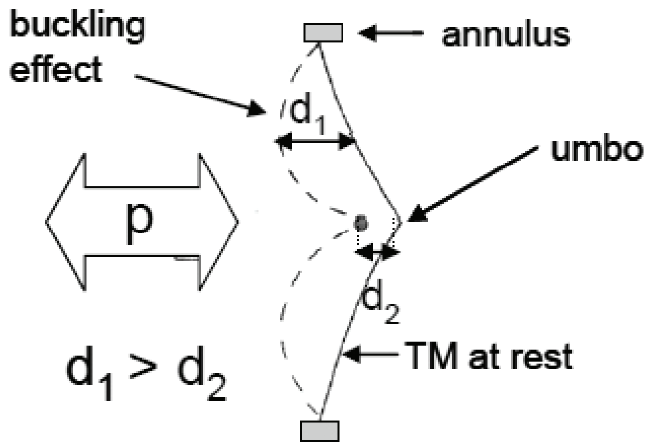


Figure 9-7. Schematic drawing of the catenary lever action of the tympanic membrane (TM); p-acoustic pressure, d–membrane displacement; (adapted from Pickles, 1988).

The impedance matching provided by the middle ear is most effective between approximately 500 and 3000 Hz but becomes less effective as the sound frequency is further away from this region (Battista and Esquivel, 2003; Nedzielnitsky, 1980; Puria, Peake and Rosowski, 1997). At low frequencies, the impedance of tympanic membrane becomes reactive and impedes the transfer of energy. Above 1000 Hz, the tympanic membrane changes its vibration pattern, resulting in a decrease in the area of the membrane contributing to the vibration (Tonndorf and Khanna, 1972). In addition, the ossicles vibrate less efficiently at frequencies above 2000 to 3500 Hz, affecting the lever mechanism and resulting in a decrease in energy transfer for higher frequencies (Battista and Esquivel, 2003). Sound transmission through the middle ear also can be affected by the air pressure in the middle ear cavity, abnormal inner ear impedance, and air coupling between the oval and round window membranes. The non-ideal operation of the middle-ear transformer at higher frequencies is, however, greatly ameliorated by the outer ear resonances. Thus, the combined effects of the outer and middle ear systems overcome their individual limitations, which would otherwise result in a large loss in the amount of energy transferred from the air to the inner ear fluid.

A natural way to assess the effect of the middle ear impedance transformer on sound transmission through the hearing system is to measure directly the input impedance of the tympanic membrane. Resistance and reactance of the tympanic membrane measured by Zwislocki (1975) for male and female populations are shown in Figure 9-8.

An examination of Figure 9-8 indicates that stiffness reactance (primarily due to the stiffness of the tympanic membrane) is the primary component in middle ear reactance. It offers the greatest opposition to the flow of energy for sounds below approximately 500 Hz and becomes negligible above approximately 800 Hz. Mass reactance (primarily due to the mass of tympanic membrane and ossicular chain) is negligible for the mid frequencies but is the primary contributor to reactance above approximately 5,000 Hz. Most importantly, the resistance of the middle ear (tympanic membrane), which affects energy transmission of most sounds within the auditory range of frequencies, varies between 200 and 400 acoustic Ω and is relatively independent of frequency across the entire range of measured frequencies (200 Hz to 8,000 Hz) (Shaw, 1997; Zwislocki, 1975). The average value of the middle ear impedance in this frequency range is approximately 300 acoustic Ω (Shaw, 1974), which agrees with the value calculated earlier in this chapter on the basis of energy reflected from the tympanic membrane.

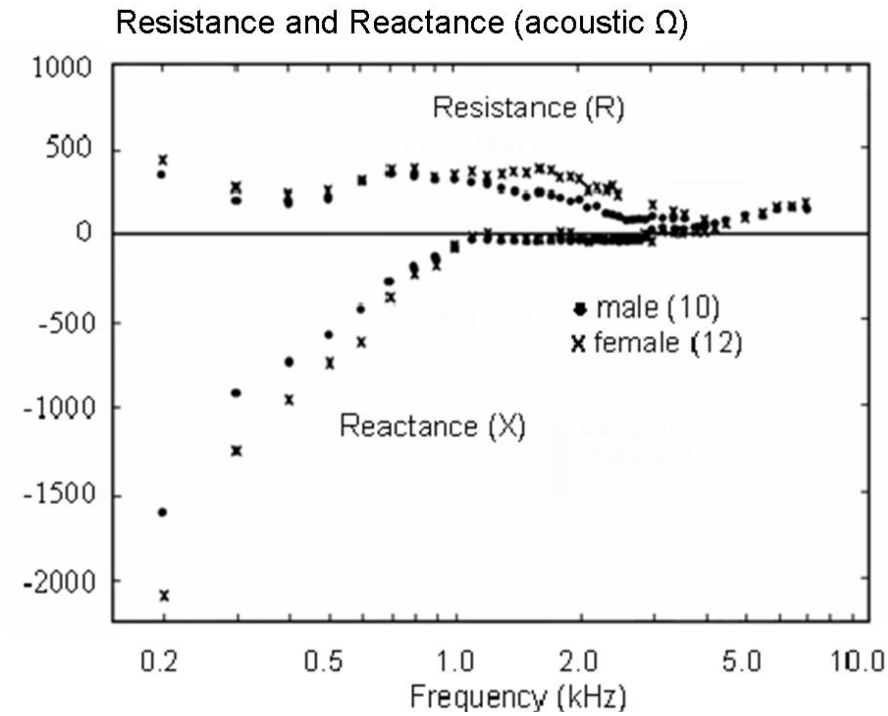


Figure 9-8. The effects of gender on resistance and reactance of middle ear impedance (adapted from Zwislocki, 1975).

Acoustic reflex

In addition to acousto-mechanical energy transformation and impedance matching, the middle ear also provides some limited protection to the inner ear against very strong stimulation. When very high acoustic pressures arrive at the tympanic membrane, the tensor tympani and stapedius muscles of the middle ear contract and temporarily stiffen the middle ear system, thereby decreasing the efficiency of energy flow through the middle ear. When the muscles are activated, the tensor tympani stiffens the tympanic membrane by pulling it toward the middle ear, and the stapedius muscle stiffens the stapes by rotating it away from its normal axis of action. This protective mechanism is known as the *acoustic reflex* or middle ear reflex and causes a 15- to 20-dB attenuation in the transmitted sound (Bess and Humes, 1990). While the role of the stapedius muscle in the acoustic reflex is generally accepted, the role of the tensor tympani has recently been questioned due to the long response latency of the tensor tympani action (approximately 100 ms) (Bosatra, Russolo and Semerano, 1997).

The SPL that triggers the acoustic reflex varies among people but is generally approximately 80-90 dB HL (above threshold of hearing). An important limitation of the acoustic reflex is that it primarily operates at low frequencies (below 4000 Hz), long latency of 35 to 150 ms (MØller, 1962), and that its contraction can be limited in duration to as little as a few seconds (e.g., at 4000 Hz). However, it is noteworthy that middle ear muscles are also activated before an onset of vocalization or chewing and remain contracted for the duration. Therefore, in addition to protecting inner ear from too intense low frequency external stimulation, the middle ear muscles may be protecting the inner ear from noise generating by the muscles activated during vocalization and jaw movements (Simmons, 1964). It is also possible that another goal of attenuation of low frequency energy by acoustic reflex is to improve audibility high frequency stimuli that are subjected to masking by low frequency stimuli (see Chapter 11, *Auditory Perception and Cognitive Performance*.

The Inner Ear

Cochlear mechanism

The cochlea of the inner ear is the system that converts the mechanical energy of the stapes motion into electro-chemical impulses that can be transmitted by the central auditory nervous system to the auditory centers of the brain. The first stage in this process is the conversion of the stapes' motion into motion of the fluids of the cochlea and the subsequent creation of a traveling wave moving along the basilar membrane. The induced movement of the basilar membrane affects the motion of the stereocilia of the outer and inner hair cells. The outer hair cells provide an amplification function, increasing the amplitude of the incoming sound wave, while the inner hair cells are sensory receptor cells, changing the mechanical motion of the stereocilia into the release of a neurotransmitter chemical that communicates with the auditory portion of the vestibulocochlear nerve. Therefore, the sound reception process in the inner ear is an active process that dissipates some of its energy in the form of otoacoustic emissions.

The complex process by which the cochlea breaks down the mechanical motion of the basilar membrane and translates it into a series of nerve impulses that can be transmitted, reassembled, and interpreted has been theorized for over a century but is still under investigation.

Traveling wave

The oval and round windows of the cochlea are both covered with elastic membranes that can bulge in and out of the cochlea. An inward motion of the stapes into the scala vestibuli causes movement of the incompressible perilymph from the scala vestibuli into the scala tympani. After the fluid passes through the helicotrema at the apex of the cochlea, the round window membrane bulges out to accommodate the increased amount of fluid in the scala tympani. A motion of the stapes away from the scala vestibuli causes perilymph to move from the scala tympani into scala vestibule and the membrane of the round window to consequently bulge inwards. The motion of the inner ear fluid caused by the inward and outward motions of the stapes creates a traveling (transverse) wave motion along the basilar membrane. The basilar membrane responds differently to sound stimuli of different frequencies, making the location where it reaches its maximum displacement depend on the frequency of the sound wave. There is a systematic shift in the point of maximal vibration from the apex toward the base as the frequency increases. Thus, the basilar membrane is said to be tonotopically organized. A view of a traveling wave moving along an uncoiled basilar membrane is shown in Figure 9-9.

The traveling wave was first described by Békésy (1953; 1955), who worked with cadaver ears and ear models. He found that the point at which the displacement of the basilar membrane was the greatest was dependent on the frequency of the sound wave and considered the traveling wave mechanism to be responsible for the sound analysis done by the cochlea. The theory of hearing, or of sound perception, based on this concept is called the traveling wave theory. There is however another competing theory of hearing, called the resonance theory, that views the basilar membrane as an array of sequentially tuned tiny resonators distributed along the membrane. This theory was originally proposed by Helmholtz (1885) and states that the tiny resonators of the basilar membrane are set directly into motion by sound pressure changes in the perilymph without needing the traveling wave to set them off. Both of these theories belong to a larger group of theories of hearing, called the place theories, which support the tonotopic organization of the basilar membrane. Place theories as well as the other group of theories of hearing, called periodicity theories, will be addressed in the frequency coding section of this chapter.

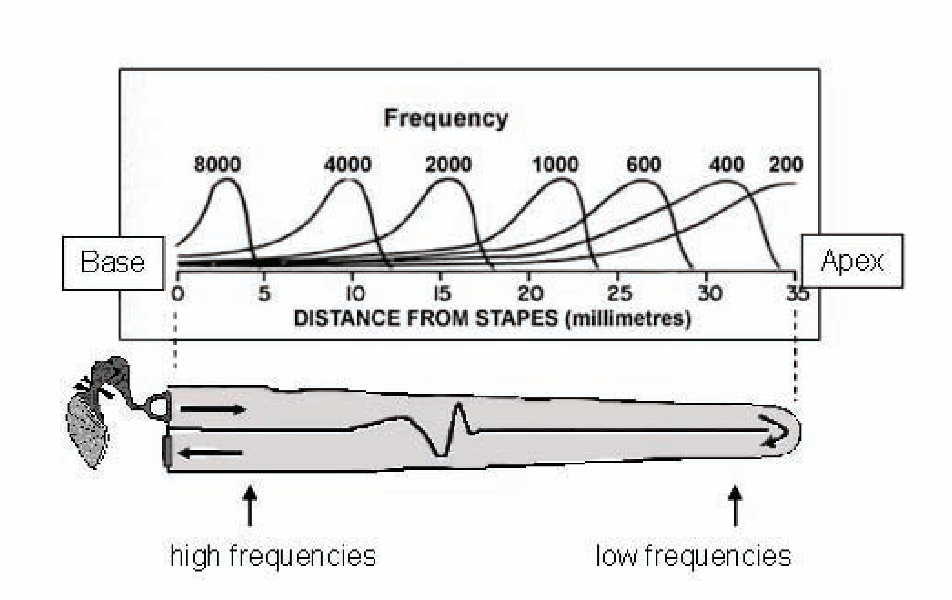


Figure 9-9. View of an uncoiled cochlea and the traveling wave (adapted from Bear, Connor and Paradiso, 2001).

There are a number of experimental studies which support either the traveling wave or the resonance theory of hearing, but there is still a lack of complete agreement in the literature as to whether the traveling wave is directly responsible for the basilar membrane motion or whether it is a secondary effect caused by the direct stimulation of basilar membrane's resonators by sound pressure propagating through the perilymph (Bell, 2004; Bell and Fletcher, 2004). Regardless of the theory behind the true manner in which the cochlea analyzes sound, the basilar membrane is the element responsible for sound analysis, and its tonotopic organization is mirrored within the auditory nervous system.

Current research, as well as some older studies (e.g., Gold, 1948, 1987; Gold and Pumphrey, 1948), supports the notion that the basilar membrane is not a passive element of the auditory system and that its action is amplified by an active mechanism in the cochlea called the *cochlear amplifier*. Thus, the fairly small motion of the basilar membrane caused by low and mid intensity sounds is amplified by the cochlea prior to the transmission of these signals from the cochlea to the vestibulocochlear nerve. The action of the cochlear amplifier has been attributed to the motility function of the outer hair cells, which expand and contract along their long axis in response to voltage changes across the cell membrane.

Electric potentials in the inner ear

The neural activity of the inner ear is dependent on electro-chemical processes and initial electric potentials between the fluids occupying the various structures of the inner ear. An electric potential is created when there is a difference in electric charge between two different locations. The area of higher charge is said to be positively polarized while the area of lower charge is said to be negatively polarized. In a biological system, such as the human ear, a difference in chemical charge between two areas is called a bioelectric potential. When the two polarized areas are connected, the charged particles move from one area to another. This occurs because of the electromotive force that is created by the difference in electrical charge. Common charged particles in the human ear include positively-charged potassium ions (K^+), negatively-charged chloride ions (Cl^-), positively-charged sodium ions (Na^+) and positively-charged calcium ions (Ca^{2+}). In the inner ear, the endolymph contains a large amount of potassium ions and the perilymph contains a large amount of sodium ions. A static bioelectric potential that involves the separation of charged particles by a cell membrane is called a *resting potential*. The resting potential of the endolymph in the scala media, called the *endocochlear potential* (EP**)**, is + 80 millivolts (mV) in reference to the resting potential of the perilymph in the two other cochlear channels. The resting potential of the inner hair cells is -40 mV and of the outer hair cells is -70 mV compared with the perilymph (Jahn and Santos-Sacchi, 2001). Therefore, the difference in potential between the endolymph and the inner hair cell is 120 mV and between the endolymph and the outer hair cell is 150 mV. This 150-mV potential is a biological battery that supports all inner ear processes. It is a very efficient system that consumes only approximately 14 microwatts (μW) of power while carrying out the equivalent of approximately one billion floating-point operations per second (1Gflops) (Zhak, Mandal and Sarpeshkar, 2004).

A dynamic bioelectrical potential that involves the movement of charged particles from one area to another in response to a stimulus is called a *stimulus related potential*. There are three stimulus related potentials that are commonly observed in the inner ear in response to an auditory stimulus. These are the summating potential (SP), cochlear microphonic (CM), and compound action potential (CAP). The former two are generated by the hair cells and the latter is generated by the vestibulocochlear nerve

The SP is a direct current potential that causes a positive or negative change in the endocochlear potential for the duration of a signal. It is the driving force for moving the charged ions through the stereocilia and membrane separating hair cell from the surrounding endolymph.

Both the CM and CAP are alternating current potentials that vary in polarity based on changes in the phase of the signal. The CM is a pre-neural electric potential that mimics the incoming sound signal; it is considered to be a reflection of receptor currents flowing through the hair cells. The CAP is the actual event related potential (ERP) that is generated when the auditory nerve “fires” (transmits a signal) in response to a stimulus. The CAP results from the firing of the auditory portion of the vestibulocochlear nerve in response to the release of neurotransmitter from the hair cells. Various ERPs and their changes in time and space can be measured in the ecenral nervous system using electroencephalography (EEG). One example of such potentials is the mismatch negativity (MMN) potential generated in the auditory cortex and having a latency of 150 to 250 ms post-stimulus. The MMN is a negative, task independent neural potential generated in response to an infrequent change in a repetitive sound sequence.

Hair cell action

Up and down movement of the basilar membrane causes shearing force acting on the cilia of the hair cells of the organ of Corti. The shearing force is a result of different points of attachment of the basilar membrane and the tectorial membrane to the cochlear wall. The force bends the cilia to the left and to the right of the basilar membrane axis. The stereocilia of a hair cell have gradually changing height and are held together by tip-to-side links that cause the whole bundle to move together when stimulated. Tilting movements of the stereocilia affect the tension on the fiber in the tip link. When the stereocilia are bent toward the largest stereocilium, the tip-to-side links cause mechanically-gated ion channels in the stereocilia membranes to open. The opening of the ion gates allows positively-charged ions (K+) of potassium, which are the main cations in the endolymph, to flow from the positively-charged endolymph into the negatively-charged hair cell. As the fiber tension increases, the flow of ions into the hair cell also increases. When the stereocilia bundle is bent in the direction away from the largest stereocilium, the ion channels close and the excess of K+ in the cell is pushed out of the cell through a semi-permeable membrane via active pumping processes restoring natural negative polarization of the cell (Geisler, 1998). However, the effects of stereocilia bending and the in-and-out flow of K+ ions are different in the inner and the outer hair cells.

When the K+ ions enter the inner hair cell, they depolarize the content of the cell, that is, they change to zero the difference in electric potentials between the areas inside and outside the cell. As a result, when the gates are open, the cell becomes depolarized (excited), and when the gates are closed, the cell becomes hyperpolarized (inhibited). These actions are shown in Figure 9-10.

The change in the electric potential across the membrane of the inner hair cell opens voltage-dependent calcium (Ca+) channels in the cell membrane. The flow of Ca+ ions triggers a release of a chemical neurotransmitter into the synaptic cleft between the hair cell and the afferent nerve ending at the basal end of the cell. The release of the neurotransmitter excites the dendrite endings of the afferent neurons connected to the inner hair cells and results in the generation of action potential in the neurons. Thus, the change in the resting potential of the inner hair cell due to K+ ions influx results in generating a stimulus related potential at its synaptic juncture with the afferent nerve fiber and in subsequent firing of the fiber.

When K+ ions enter and leave the outer hair cells, these cells alternately contract and expand in response to the alternating current generated by the polarization and depolarization of the cell walls. Outer hair cells are anchored both at their top (at the base of the stereocilia) and at their bottom but are not firmly attached at their sides. Thus, the expansion and contraction of the outer hair cells pushes up on the reticular lamina and down on the support cells and basilar membrane. Active motions of outer hair cells (OHCs) increases the range of motion of the basilar membrane, thereby causing larger deflections of inner hair cells (IHCs), and “sharpens” the shape of the traveling wave motion along the basilar membrane.

Each hair cell is connected to both afferent and efferent neurons. When activated by a hair cell, the afferent neuron conducts stimulus related potential up to the central nervous system. The brain-controlled action of the efferent neuron is to release a neuro-inhibitor acetylocholine (Ach) in a synaptic juncture with the hair cell or with a respective afferent neuron to impede hair cell action.

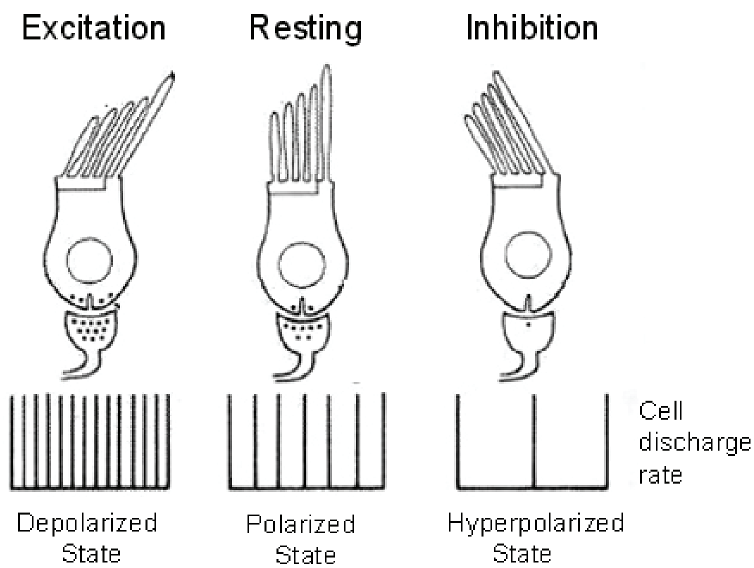


Figure 9-10. Inner hair cell response to bending of the stereocilia.

Cochlear amplifier and otoacoustic emissions

The amplification of basilar membrane’s vibration caused by the motility of the outer hair cells is not linear and varies based on the intensity of the incoming signal. Low intensity signals are amplified more than high intensity signals. It is generally accepted that an active mechanism of amplification of the IHC responses by OHC motility operates in the range to 50 dB SL (Stebbins et al. 1979).

A by-product of non-linear activity of the outer hair cells is nonlinear distortion of the basilar membrane movements. The distortion products generated in the cochlea travel from the inner ear through the middle ear and into the ear canal in a transmission process that is the reverse to the process of hearing. They have a form of *cochlear echoes*, more formally called otoacoustic emissions (OAEs), which can be observed in the ear canal. It is still unclear, however, exactly how the waves travel backwards along the cochlea to the oval window, that is, whether they are slow traveling waves along the basilar membrane or faster compression waves in the perilymph. The results of recent studies favor compression waves; however, the debate is still unsettled.

The OAEs produced by nonlinear actions of OHCs have been first reported by Kemp (1978) and can be measured in the ear canal using sensitive microphones and applying signal averaging techniques to the collected data. The presence of the OAEs in the outer ear canal is an evidence of an active amplification process within the cochlea.

There are two types of OAEs that can be observed in the ear: spontaneous OAEs and evoked OAEs. Spontaneous OAEs are present in approximately 50% of people with normal hearing and result from random processes ongoing in the inner ear. Evoked OAEs are the ear responses to external stimulation and are present in nearly 98% of the normal ears. They are usually evoked for audiologic assessment of the ear function with either a click stimulus or two simultaneous pure tones f_1 and f_2, whose ratio is between 1.1 and 1.3. The most prominent and commonly observed evoked OAE component is the cubic difference distortion product denoted as $f_{cd} = 2f_1\text{-}f_2$.

Auditory nerve: The link

The transduction of the mechanical actions of the shearing of the stereocilia to the electro-chemical signal transmitted by the nervous system begins when a neurotransmitter is released from the base of the inner hair cell. This neurotransmitter crosses the synaptic cleft and binds to specialized receptor sites located on the post-synaptic membrane of the peripheral processes of the nerve fibers connecting the inner ear with the brainstem. The bundle of nerves connecting the inner ear with the brainstem is called the auditory nerve. If sufficient neurotransmitter is released, the afferent nerve fibers will fire in response, sending an electric signal down the length of the auditory nerve towards the brainstem. The innervations of the inner and outer hair cells by the fibers of the auditory nerve are shown in Figure 8-18 of Chapter 8, *Basic Anatomy of the Hearing System*.

The auditory nerve (cochlear nerve, acoustic nerve) is a part of the vestibulocochlear (VIII) cranial nerve comprised of the vestibular nerve (not discussed in this chapter) and the auditory nerve, The auditory nerve is a bipolar nerve with cell bodies (collectively called the spiral ganglia) located within Rosenthal's canal in the cochlea. The peripheral projections (dendrites) of the cells synapse with the hair cells and the central projections (axons) synapse with other nerve cells located in the cochlear nucleus of the brainstem. The fibers of the auditory nerve that transfer information from the cochlea to the brainstem are mostly myelinated, i.e., covered with an insulating lipid membrane that wraps around the nerve fiber and acts as an electrical insulator.

The tonotopic organization of the basilar membrane is passed to the auditory nerve via a tonotopic arrangement of auditory nerve fibers. The fibers that originate in the low frequency area (apex) of the basilar membrane run in the center of the auditory nerve while the fibers that originate in the high frequency area (base) of the basilar membrane are in the periphery of the auditory nerve. Therefore, damage to the outside of the auditory nerve will result in high frequency hearing loss, while damage to the central core of the nerve trunk will cause low frequency hearing loss.

The nerve conduction velocity or the speed at which the neural impulse travels along the neural pathways depends on the size of the neuronal fiber (axon or dendrite) – larger fibers conduct impulses faster than smaller ones. In addition, nerves with myelin sheathing have faster conduction times than the un-myelinated ones. It is only at the gaps in the myelin (nodes of Ranvier) that electrically-gated ion channels will open and close in response to the impulse traveling down the nerve. When the impulse travels along an unmyelinated axon, the ion channels open and close along the entire axon, slowing the conduction time. The type I bi-polar neurons responsible for conveying sound information to the central auditory system have myelin coating most of their length and thus are relatively efficient conductors of nerve impulses. The size of the auditory nerve fibers, at least in children, is relatively uniform, suggesting that nerve conduction velocity in the auditory nerve is similar, which preserves the timing coding created at the level of the cochlea (Spoendlin and Schrott, 1989).

Frequency coding in the cochlea

There are a number of theories that attempt to explain how the motions of the hair cells work to code sound. In other words, trying to explain how the firing pattern created by the periodic release of neurotransmitter from the inner hair cells to the auditory nerve results in a signal that the brain can interpret as sound. Two major classes of the theories of hearing that try to explain how signal frequency is encoded by the cochlea are place theories and periodicity (frequency, time) theories.

Place theories of hearing

Place theories of hearing, such as the traveling wave theory and the resonance theory, assume that sound frequency is coded along the basilar membrane by the place at which the membrane vibrates the strongest in response to the acoustic stimulus. Recall that the physical properties of the basilar membrane change gradually along the length of the membrane and the high and low frequency vibrations reach their maximum amplitudes at the base and the apex of the membrane, respectively. The place along the basilar membrane that vibrates the strongest maximally stimulates the hair cells at that location and these, in turn, stimulate the auditory neurons at that location. However, since the vibration of the basilar membrane extends always over a certain finite area, it does not just activate a single discrete row of neurons but also those in the surrounding region. Therefore, one of the important properties of the basilar membrane treated as a place-dependent coder of signal frequency is frequency selectivity of nerve fibers located along the basilar membrane. This selectivity can be measured and expressed by tuning curves of various fibers.

A tuning curve represents the changes in the minimal response threshold of auditory fibers as a function of frequency. The tuning curve for a single auditory fiber resembles the letter “V” where the tip of the tuning occurs at the characteristic frequency (CF) of the fiber. The CF is the sound frequency at which a fiber has its lowest threshold, i.e., it fires in response to the sound at a fairly low intensity level compared to sounds with other frequencies. This same nerve will also fire in response to other frequencies, but the threshold will be much higher. Generally, the further away in frequency from the characteristic frequency, the greater the intensity is required for the nerve to fire. The tuning curves can be measured using both psychoacoustic and physiologic methods. Some examples of psychoacoustic tuning curves are shown in Figure 9-11.

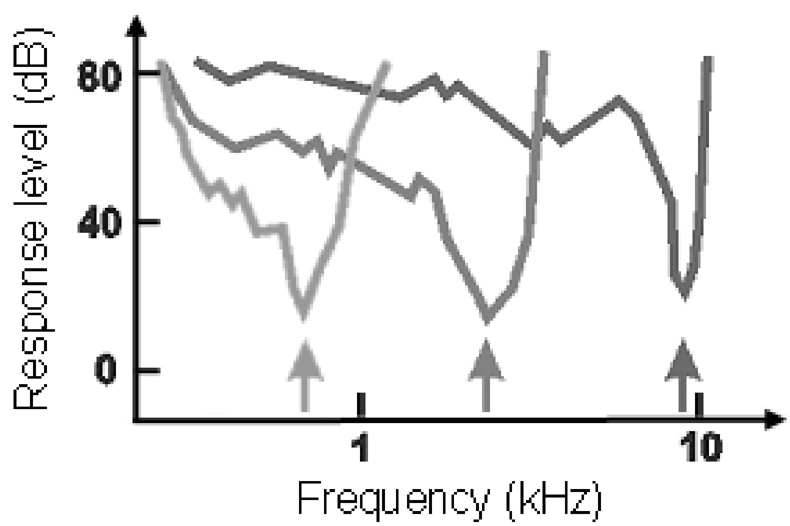


Figure 9-11. Psychoacoustic tuning curves.

Periodicity theories of hearing

Periodicity (frequency, time) theories of hearing, such as the telephone theory (Rutherford, 1886) or volley theory (Wever, 1949), state that the perception of sound depends on the temporal pattern of the sound wave and that the sound frequency is coded by the number of neural impulses per second that are fired by a group of nerve fibers along the basilar membrane. Periodicity theories claim that single nerve fibers of the basilar membrane need not respond to every successive wave of the sound stimulus but could respond only to every second, third, or fourth wave. Each wave is thought to excite a separate group of nerve fibers and the pattern of neural impulses reaching the brain by the successive "volley of impulses" represents the frequency of the sound.

Periodicity coding assumes that during the positive phase of a sound wave the stereocilia are sheared in one direction and in the negative phase of the wave, the stereocilia are shearing in the opposite direction. All nerve fibers have some amount of random, spontaneous firing. When stimulated, the firing rate increases and when inhibited from firing, the firing rate decreases. When the shearing of the outer hair cell stereocilia is in an excitatory direction, neurotransmitter is released and the receiving neuron is stimulated (assuming sufficient neurotransmitter). When the shearing is in the inhibitory direction, the firing rate decreases from the resting firing rate. Thus there is a distinction in the firing code between phases with positive polarity and phases with negative polarity. Sound waves containing a periodicity will have a firing rate that follows the period of the sound wave up to approximately 4,000 to 5,000 Hz (Kiang et al., 1965; Rose et al., 1967). Individual neurons cannot fire at such high rate, but groups of neurons acting together can, a phenomenon known as population coding.

Newer research indicates that place coding and periodicity coding actually work together in coding signal frequency. According to this concept the place coding mechanism acts like a series of tuned filters which divide the signal into pieces that are more easily transmitted in the form of periodicity codes (d'Cheveigne, 2005). The most common current view on frequency coding, holds that low frequencies up to approximately 400 Hz are coded by signal periodicity and high frequencies above approximately 4000 Hz by place of excitation. The frequencies in 400 to 4000 Hz range are coded by a combination of place and periodicity codes.

Intensity coding in the cochlea

Place coding and *periodicity coding* are the two ways in which sound frequency is coded into neural impulses. These theories, however, do not explain how the ear codes the intensity of the sound.

According to current literature, there are two basic ways in which the intensity of the sound can be coded in the cochlea. First, the more intense the stimulus, the larger the excitation of the basilar membrane and the larger the number of neural fibers that can fire simultaneously. Thus, a high intensity sound may stimulate a broad range of cells and the number of cells that fire may represent the intensity of the stimulus.

Second, auditory neurons, like all nerve fibers, fire spontaneously even in the absence of stimulation. There is evidence that auditory neurons can be grouped into three basic types on the basis of how frequently they exhibit this spontaneous activity and the range of sound levels that they respond to when stimulated (Lieberman, 1978). These different types of nerve fibers are called high, medium, and low spontaneous activity fibers. High spontaneous activity fibers have, as their name implies, a high rate of firing in the absence of stimulation and their firing rate increases in response to stimuli between approximately -10 and 30 dB SPL. At approximately 30 dB SPL, these fibers reach their saturation point and their response plateaus. Mid spontaneous activity fibers have a lower rate of firing in the absence of stimulation than do the high spontaneous activity rate fibers and their firing rate increases in response to somewhat louder stimuli between approximately 5 and 35 dB SPL. Low spontaneous activity fibers have the lowest rate of firing in the absence of stimulation and their firing rate increases in response to stimuli between approximately 25 and 40 dB SPL. It is hypothesized that the combination of these fibers working together allows for a large dynamic range of intensity to be coded by the cochlea and carried along the auditory nerve.

Bone Conduction

Sound waves and mechanical vibrations acting on the surface of the human head are absorbed by the soft tissue and bones of the head and propagate through the head's structures. They also induce complex physical vibrations of the skull that can be transmitted to the brain and sensory organs of the head. Both these mechanisms together can affect the hearing system and evoke auditory responses analogous to those caused by sound waves arriving at the outer ear of the hearing system. Note that all important head resonances have frequencies below transmission range of bone conducted stimuli to protect supporting them mechanisms from potential harmful effects of body vibrations and acoustic stimulation. They include, but are not limited to, eyeball resonance, jaw resonance, neck resonance, and head and shoulder vibrations

For the human communication purposes only bone vibrations created by directly applied physical vibrations (via a mechanical vibrator) have sufficient energy to be used as a carrier of information. Sound waves arriving to the surface of the head are either captured by the outer ear and delivered through the hearing system to the organ of Corti or are mostly reflected back into the environment due to the impedance mismatch between the impedance of the skull and impedance of surrounding air. Sound waves can only be heard through bone conduction when the arriving sound has very high intensity and the person's ears are occluded by hearing protectors or head mounted audio displays (audio HMD). However in such cases, perceived sound often constitutes the harmful noise that leaked to the hearing system by bone conduction pathways rather than a communication signal.

The first modern theory of bone conduction hearing was proposed by Herzog and Krainz in 1926 (Herzog and Krainz, 1926). According to this theory the bone conduction hearing is a combination of two phenomena:

1. Relative motion of the middle ear bones caused by head vibrations
2. Compression waves in the cochlea resulting from the transmission of vibrations through the skull

Two landmark publications on bone conduction by Békésy (1932) and Barany (1938) provided further evidence for and expanded this theory. They also provided clear evidence that air conduction and bone conduction mechanisms were two different hearing mechanisms resulting in the same excitation of the basilar membrane.

Current theory of bone conduction hearing is mostly based on the comprehensive studies by Tonndorf (1966; 1968) who expanded Herzog and Krainz' work and identified seven potential mechanisms that can contribute to human hearing. The four main mechanisms proposed by Tonndorf are:

1. Inertial Mechanisms
    a. Middle ear inertial mechanism involving relative and delayed movement of the ossicular chain in reference to the surrounding temporal bone (cochlear promontory).
    b. Inner ear inertial mechanism involving transmission of temporal bone vibrations on the inner ear fluids.
2. Compression Mechanisms
    a. Outer ear compression mechanism involving radiation of bone-conducted energy from the osseous walls of the ear canal back to the ear canal.
    b. Inner ear compression mechanism involving compression and decompression of the inner ear fluids by compression vibrations of the bony cochlea.

The most dominant of the above mechanisms seems to be the inner ear inertial mechanism (Stenfelt and Goode, 2005a) although several other mechanisms contribute as well. The effectiveness of the individual mechanism depends on the frequency of the signal, place and direction of vibration application, and the status of the outer ear. The two first factors affect the modes of vibrations of the skull whereas the third depends on the type and quality of the ear occlusion. The last factor affects dramatically the effectiveness of the outer ear compression mechanism. Ear occlusion boosts the effectiveness of bone conduction in a low frequency range by as much as 30 dB at 100 Hz and 5 dB at 2000 Hz (Stenfelt, 2007).

The middle ear inertial mechanism and the inner ear compression mechanism seem to operate in the high frequency range above 1 kHz (Stenfelt, 2006; Stenfelt and Goode, 2005b). However, the contribution of the latter mechanism has been questioned recently due to a lack of convincing experimental evidence. For example, clinical findings for otosclerosis and semicircular canal dehiscence cases do not provide support for the presence of this mechanism for frequencies below 4 kHz (Stenfelt, 2007).

In summary, the bone conduction mechanisms are still not very well understood, although hearing through bone conduction has been known and used for centuries (Henry and Letowski, 2007). The process is linear and most likely cochlear (Steinfelt, 2007). However, some new theories of bone conduction hearing point to the movement of the otolith stones in the saccule of the inner ear (Lenhardt, 2002; Todd and Cody, 2000; Welgampola et al., 2003) or to the action of the cochlear and vestibular aqueducts that exchange perilymph between cerebrospinal cavity and the cochlea (Freeman, Seichel and Sohmer, 2000; Sohmer et al., 2000).

One of the limitations of the bone conduction hearing is limited spatial perception due to the lack of pinnae cues and high velocity of acoustic waves through the bones. However, it permits sound source lateralization within the head and for spatial perception of audio signals delivered through bone conduction audio HMDs if air conduction HRTFs are used (MacDonald, Henry and Letowski, 2006).

Vestibular System Function

In addition to the cochlea, the inner ear houses the organs of balance. They include three cristae ampullares, located in three bony channels called the semicircular canals, and two maculae, located in two connected sacks, called utricle and saccule, within the bony cavity of the vestibule, between the semicircular canals and the cochlea. The cristae ampullares convey information approximately angular acceleration of the head to the central nervous system. The maculae convey information about linear acceleration and head position relative to gravity. The utricular macula is oriented horizontally and the saccular macula is oriented vertically. Tilting head to the side stimulates saccular macula and titling head forward or to the back stimulates the utricular macula. All these sensory organs contain hair cells with their stereocillia responding to the head motion analogous to the way the inner hair cells in the cochlea respond to the acoustic signal. Depending on the head position and the direction of the head movement the endolymph flow in the semicircular canals and the vestibule stimulates the hair cells of the organs of balance. For example, the cilia of the maculae are embedded in the gelatinous membrane containing relatively heavy calcium carbonate (Na) otoliths (otoconia); movements of the head cause the otoliths to bend the cilia causing depolarization/hyperpolarization of the hair cells depending on the direction of movement.

The signals from the organs of balance are transmitted through the vestibular portion of the vestibulocochlear nerve to four vestibular nuclei (superior, lateral, medial, and inferior nuclei) within the brainstem and further to the brain (cerebellum). The fibers from the vestibular nuclei also crossover to the contralateral nuclei from which they project, among others, to oculomotor nuclei that drive eye muscle activity resulting in vestibule-ocular reflex that helps maintain fixation of the eyes on the object moving in relation to the head position. Thus, human balance involves a complex coordination between the vestibular system, visual system, proprioceptors (sensors in muscles and joints), and structures within the cerebellum, brainstem, and the whole cortex.

Central Auditory Nervous System Processing

Cochlear nucleus

Auditory nerve fibers arrive at the brainstem by forming synapses with large groups of neurons in the cochlear nuclei located in the border between the pons and medulla. The fibers from each ear terminate on the nucleus located on the same (ipsilateral) side of the brainstem from where most of the fibers cross to the opposite (contralateral) side of the brainstem and either connects to contralateral superior olivary complex or ascends directly to contralateral inferior colliculus in the midbrain. Type I nerve fibers with large myelinated neurons are responsible for transporting the coded auditory signal from the peripheral to the central nervous system. The function of the smaller and less numerous type II fibers is still largely unknown.

The neural cells in the cochlear nuclei have several complex firing patterns and wider dynamic ranges than the neurons in the auditory nerve. In response to simple tonal stimuli, several response patterns have been recorded in the various cells of the cochlear nuclei. These response patterns include a "primary" pattern that is similar to that of the auditory nerve ("primary-like" neurons); a "chopper" pattern that consists of repeated bursts of firing followed by short pauses ("chopper" neurons) (however, this periodicity does not match the periodicity of the stimulus; an "on" pattern, in which the cell fires only when a stimulus begins ["on" neurons]; and a "pauser" pattern in which the cell fires only at the onset of the stimulus, pauses, and then continues until the stimulus is turned off ("pauser" neurons) (Pfeiffer, 1966). Examples of peristimulus (PST) histograms, i.e., the histograms of the times at which neurons fire, as a function of latency following tonal stimuli, are shown in Figure 9-12.

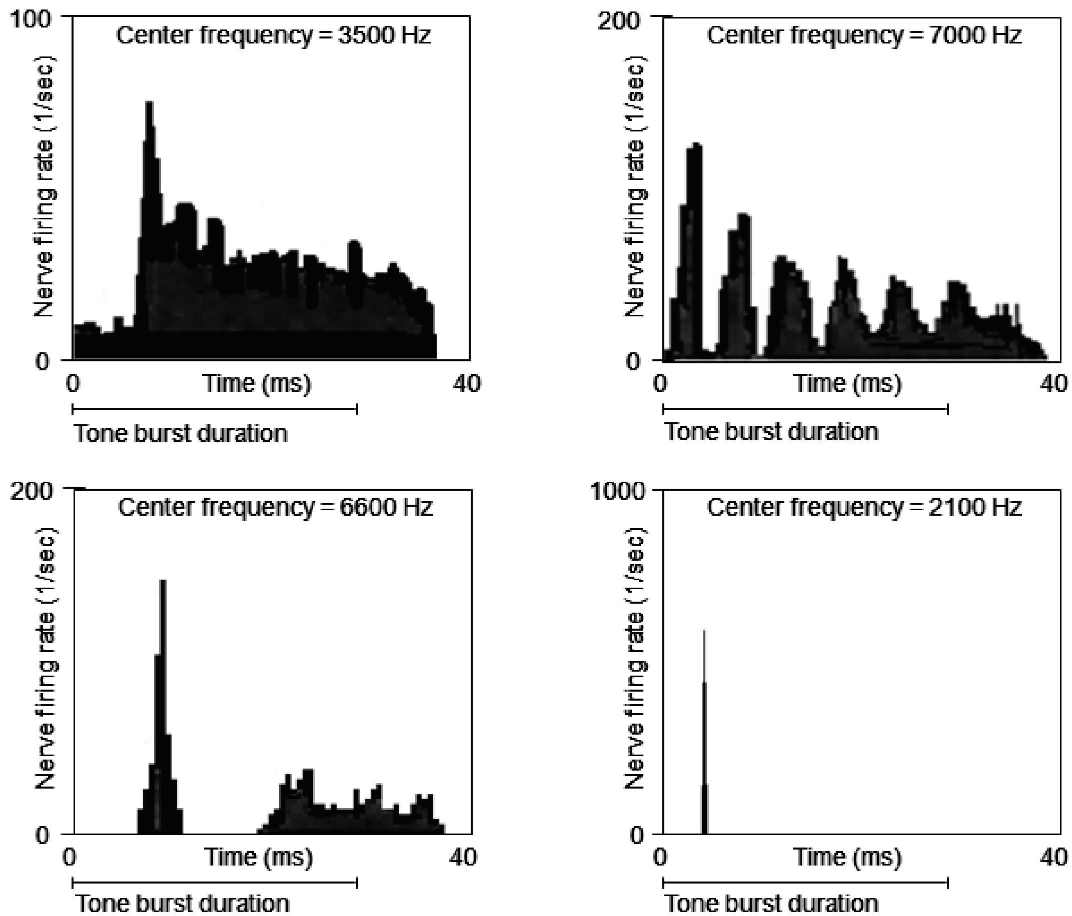


Figure 9-12. PST histograms illustrating different types of neuron firing patterns observed in the cochlear nucleus including: (a) primary like, (b) chopper, (c) pauser and (d) on (from Pfeiffer, 1966).

These firing patterns shown in Figure 9-12 are the most commonly reported firing patterns. However, there also are other firing patterns observed in response to simple stimuli. Further, there are reports of many different subcategories of firing pattern under each of these main categories, and the response of the cochlear nucleus to complex stimuli varies from the response to simple stimuli.

The type of response recorded from the neurons in the cochlear nucleus depends on a number of physical features associated with the cells (e.g. characteristics of the membrane, type of cell), the connection between the auditory nerve and the cochlear nucleus cells (e.g. many or few axon endings contacting many or few dendrites), and the presence of inhibitory input from other cells (Cant, 1982; Ostapoff, Feng and Morest, 1994; Rhode and Smith, 1986). For example, in the anterior ventral cochlear nucleus (AVCN), the most common cell types are the global and spherical bushy cell. These cells receive very few axonal connections from auditory nerve fibers from a localized frequency area of the cochlea and are the most likely contributors to the primary response pattern seen in the cochlear nucleus. Their frequency specificity may also be enhanced by their function as coincidence detectors, which reduce the random noise level from spontaneous activity of the auditory nerve (Joris et al., 1994; Joris, Smith and Yin, 1994; Louage, van der Heijden and Joris, 2005). Other cells may specialize in transmitting intensity of sound (multipolar cells) or temporal order of sound events (octopus cell).

In the posterior ventral cochlear nucleus (PVCN), the octopus cell is a common cell type, so-called because these cells resemble an octopus with long tentacle-like dendrites. These dendrites receive many more connections, across a broader frequency range of the cochlea, compared with the AVCN bushy cells, and they are thus more broadly tuned. These cells have been reported to respond well to amplitude modulated tones (Oertel et al., 2000) and clicks, but have a reduced activity in response to steady state noise (Levy and Kipke, 1998). In some species, the dorsal cochlear nucleus (DCN) has been recorded to respond to spectral differences that may indicate they provide some coding in response to monaural localization cues in a vertical plane (Spirou et al., 1999).

Superior olivary complex

The superior olivary complex (SOC) receives inputs from both cochlear nuclei (contralateral and ipsilateral) and has an important role in sound localization. The two largest nuclei in the SOC are the lateral SOC (LSOC) and the medial SOC (MSOC). The MSOC and LSOC are binaural integration sites where information from each ear comes together and the input from the right and left ears are compared against each other. Therefore, these two centers play a primary role in the creation of coding cues for the sound localization in the horizontal plane. The differences between the ears are used to create codes that represent ITDs, primarily in the MSOC (Masterton, Jane and Diamond, 1967; Masterton et., 1975) and IIDs, primarily in the LSOC. The MSOC, similar to the cochlear nucleus, appears to use coincidence detection as a mechanism for coding the different arrival times between cells. The actions performed by the MSOC and LSOC neurons are shown schematically in Figure 9-13.

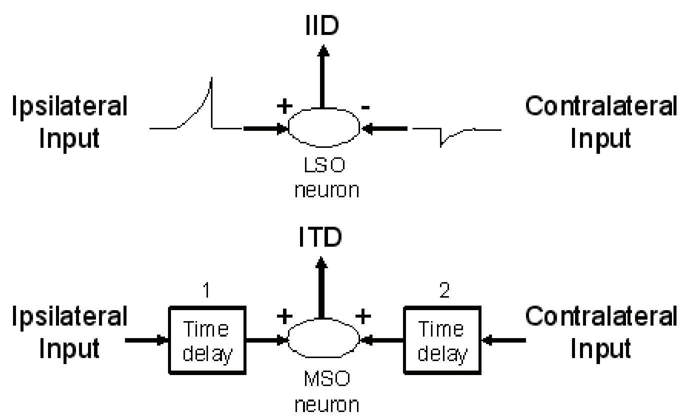


Figure 9-13. Coding of IID and ITD cues in the superior olivary complex.

Higher auditory centers

All of the ascending fibers from the cochlear nucleus and superior olivary complex travel in a large fiber tract called the lateral lemniscus and synapse in the inferior colliculus. The inferior colliculus appears to be another site that processes intensity and timing differences between ears important for spatial sound perception (Masterton, Jane and Diamond, 1967; Møller, 2002). It also processes sound onset and duration and is generally believed to be the first site at which complex sounds are encoded.

Auditory fibers projecting from the inferior colliculus ascend to the medial geniculate body of the thalamus and from here to the neocortical structures of the primary and secondary auditory cortex located in the transverse temporal gyrus (Heschl's gyrus) of the brain. The medial geniculate body acts primarily as a relay station for all ascending auditory signals passing them bilaterally onto primary auditory cortex. However, because of the decussating of a majority of fibers prior to the lateral lemniscus, the primary signal from the right ear arrives in the left hemisphere of the brain and the primary signal from the left ear arrives in the right hemisphere. The primary auditory cortex has tonotopic organization with characteristic frequencies (CFs) of the neurons increasing from caudal to rostral locations. Somewhere in the center of the area there are also patches of more specialized neurons (Ehret, 1997).

Auditory signals are subsequently sent from Heschl's gyrus to other parts of the brain for feature extraction and further analysis by auditory, multisensory, and cognitive areas. One of such prominent other areas is Wernicke's area, located at and adjacent to the posterior end of the superior temporal gyrus. Wernicke's area is responsible for decoding a speech signal into a semantically recognizable message. The *planum temporale* of Wernicke's area is also involved in music and timbral changes perception. In most individuals, Wernicke's area is more developed on the left side, thus a speech signal directed to the right ear arrives directly to Wernicke's area and the signal from the left ear must travel from the left temporal lobe to the right temporal lobe via the large corpus callosum fiber track. Thus, most individuals demonstrate a right ear advantage, in which speech perception is superior when the signal is directed to the right ear. This effect is seen most clearly in children and diminishes with age. Another important area connected with primary auditory cortex is Broca's area that control speech production.

Descending (efferent) nerve fibers run from the auditory cortex back to the cochlea forming ipsilateral and contralateral synaptic junctions in reverse order to the ascending fibers (Figure 8-20). The corticothalamic projections descend from the auditory cortex to the thalamus and medial geniculate body whereas another group of projections, called the corticocollicular projections, descends directly to the inferior and superior colliculi. Most of the descending pathways from the colliculi terminate at the lateral and medial periolivary nuclei of the SOC but some others project directly to cochlear nuclei. The descending olivo-cochlear projections connect the periolivary nuclei to the outer hair cells and the radial fibers in the cochlea modifying their responses (Oliver, 1997).

Although the complete understanding of the efferent system is still elusive, its main function seems to be to provide a gain control mechanism for the auditory system. The descending fibers conduct neural impulses that control the neuromuscular feedback system of the middle ear, the amplification function of the outer hair cells, and the sensitivity of the inner hair cells. Stimulation of the outer hair cells by signals descending from the medial periolivary nuclei through MOCB (see Chapter 8, *Basic Anatomy of the Hearing System*) decreases motility of the outer hair cells affecting sensitivity of the cochlear amplifier and improves speech recognition in noise (Giraud et al., 1997; Kumar and Vanaja, 2004). It has also been reported that efferent system affects detection of tones of unexpected frequencies (Scharf, Magnan and Chays, 1997) and controls an attentive state of a person (Froehlich et al., 1990; Yost, 2000).

It has to be stressed that the properties of the individual neurons and their synaptic connections do not represent a fixed structure in the central nervous system. They are affected by immediate surroundings, experience, selective attention, learning, and emotional states of a person and they are continuously modified throughout the lifetime.

References

Batteau, D.W. (1967). The role of the pinna in human localization. *Proceedings of the Royal Society, Series B, Biological Sciences,* 168, 158-180.

Battista, R.A., and Esquivel, C. (2003). Middle Ear, Ossiculoplasty. E-medicine (WEB Journal, J.A. Shohet, [Ed.]): Retrieved 16 January 2009 from http://www.emedicine.com/ent/topic219.htm.

Bear, M.F., Connors, B.W., and Paradiso, M.A. (2001). *Neuroscience: Exploring the Brain*. Baltimore: Lippincott, Williams, and Wilkins.

Barany E. (1938). A contribution to the physiology of bone conduction. *Acta Otolaryngologica,* 26(Suppl.), 1-223.

Békésy, G. von (1932). Zur theorie des Horens bei der Schallaufnahme durch Knochenleitung. *Annalen der Physik*, 12, 111–36.

Békésy, G. von (1941). Über die Messung der Schwingugsamplitude der Gehörknöchelchen mittels einer kapazitiven Sonde. *Akustische Zeitschrift*, 6, 1-16.

Békésy, G. von (1942). Über die Schwingungen dfer Schneckentrennwand des Ohres. *Akustische Zeitschrift*, 7, 173-186.

Békésy, G. von (1949). The vibration of the cochlear partition in anatomical preparations and in models of the inner ear. *Journal of the Acoustical Society of America*, 21, 233-245.

Békésy, G. von (1953). Description of some mechanical properties of the organ of Corti. *Journal of the Acoustical Society of America*, 25, 770-785.

Békésy, G. von (1955). Paradoxical direction of wave travel along the cochlear partition. *Journal of the Acoustical Society of America*, 27, 155-161.

Békésy, G. von (1960). *Experiments in Hearing*. New York: McGraw-Hill.

Bell, A., and Fletcher, N.H. (2004). The cochlear amplifier as a standing wave: “Squirting” waves between rows of outer hair cells? *Journal of the Acoustical Society of America*, 116, 1016-1024.

Bell, A. (2004). Hearing: Traveling wave or resonance? *PLoS Biology,* 2(10), 1521-1523.

Bess, F.H., and Humes, L.H. (1990). *Audiology: The Fundamentals*. Baltimore, MD: Williams and Wilkins.

Bosatra, A., Russolo, M., and Semerano, A. (1997). Bipolar electric stimulation to elicit and isolate tensor tympani reflex. *Acta Otolaryngologica*, 84, 391-392.

Cant, N.B. (1982). Identification of cell types in the anteroventral cochlear nucleus that project to the inferior colliculus. *Neuroscience Letters*, 32, 241-246.

d'Cheveigné, A. (2005). Pitch perception models. In: Plack, C.J., Oxenham, A.J., Fay, R.R., and Popper, A.N. (Eds.). *Pitch*. Berlin: Birkhäuser.

Duda O.R. (2000). 3-D Audio form HCI: Elevation cues. Retrieved on January 22, 2008: http://interface.cipic.ucdavis.edu/CIL_tutorial/3D_psych/elev.htm

Ehret, G. (1997). The auditory cortex. *Journal of Computational Physiology*, 181, 547-557.

Emanuel, D.C. and Letowski, T. (2009). *Hearing Science*. Baltimore, MD: Lippincott, Williams and Wilkins.

Freeman, S., Seichel, J.Y., Sohmer, H. (2000). Bone conduction experiments in animals: Evidence for a non-osseous mechanism. *Hearing Research*, 146, 72-80.

Froehlich, P., Collet, L., Chanal, M., and Morgon, A. (1990). Variability of the influence of a visual task on the active micromechanical properties of the cochlea. *Brain Research*, 508, 286-288.

Geisler, C.D. (1998). *From Sound to Synapse*. New York: Oxford University Press.

Giraud, A.L., Garnier, S., Micheyl, C., Lina, G., Chays, A., and Chery-Croze, S. (1997). Auditory efferents involved in speech-in-noise intelligibility. *Neuroreport*, 8, 1779-1783.

Gold, T. (1948). Hearing II. The physical basis of the action of the cochlea. *Proceedings of the Royal Society of London. Series B (Biological Sciences),* 135, 492-498.

Gold, T. (1987). The theory of hearing. In: Messel, H. (Ed.), *Highlands in Science*. Sydney, Australia: Pergamon Press.

Gold, T., and Pumphrey, R.J. (1948). Hearing I. The cochlea as a frequency analyzer. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 135, 462-491.

Goode, R. (2006). The catenary principle of tympanic membrane function – time to put it to rest? *Proceedings of the 4th International Symposium on Middle Ear Mechanics in Research in Otology (MEMRO 2006)*. Zurich, (27-30 July).

Hebrank, J., and Wright, D. (1974). Spectral cues used in the localization of sound sources on the median plane, *Journal of the Acoustical Society of America*, 56, 1829-1834.

Helmholtz, H. von (1885). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. London: Longmans.

Helmholtz, H. von (1868). Die Mechanic der Gehorknöchelchen und des Trommelfells. *Pflügers Archiv für die gesamte Physiologie*, 1, 1-60.

Hemilä, S., Nummela, S., and Reuter, T. (1995). What middle ear parameters tell about impedance matching and high frequency hearing. *Hearing Research*, 85, 31-44.

Henry, P., and Letowski, T. (2007). Bone conduction: Anatomy, physiology, and communication. Aberdeen, MD: U.S. Army Research Laboratory. ARL Technical Report ARL-TR-4138.

Herzog, H., and Krainz, W. (1926). Das Knochenleitungsproblem. *Zeitschrift Hals- Nas- Ohrenheilk*, 15, 300-311.

Jahn, A., and Santos-Sacchi, J. (2001). *Physiology of the Ear* (2nd Ed.). New York: Singular Press.

Joris, P.X., Carney, L.H., Smith, P.H., and Yin, T.C.T. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *Journal of Neurophysiology*, 71(3), 1022-1036.

Joris, P.X., Smith, P.H., and Yin, C.T. (1994). Enhancement of neural synchronization in the anterovental cochlear nucleus. II. Responses in the tuning curve tail. *Journal of Neurophysiology,* 71(3), 1037-1051.

Kemp, D.T. (1978). Stimulated acoustic emissions from within the human auditory system. *Journal of the Acoustical Society of America,* 64, 1386-1391.

Khanna, S.M., and Tonndorf, J**.** (1972)**.** Tympanic membrane vibrations in cats studied by time-averaged holography. *Journal of the Acoustical Society of America*, 51, 1904-1920.

Kiang, N.Y-S., Watanabe, T., Thomas, E.C., and Clark, L.F. (1965). *Discharge patterns of single fibers in the cat's auditory nerve.* Cambridge, MA: MIT Press.

Killion, M., and Dallos, P. (1979). Impedance matching by a combined effect of the outer and middle ears. *Journal of the Acoustical Society of America*, 66, 599-602.

Kumar, U.A., and Vanaja, C.S. (2004). Functioning of the olivocochlear bundle and speech perception in noise. *Ear and Hearing*, 25(2), 142-146.

Lenhardt, M.L. (2000). Eardrum saccule coupling; novel form of hearing. In: Vossoughi, J. (Ed.), *Biomedical Engineering: Recent Developments*. Washington, D.C.: Medical and Engineering Publishers.

Levy, K.L., and Kipke, D.R. (1998). Mechanisms of the cochlear nucleus octopus cell's onset response: Synaptic effectiveness and threshold. *Journal of Acoustical Society of America,* 103(4), 1940-1950.

Lieberman, M.C. Auditory nerve response from cats raised in a low-noise chamber. *Journal of the Acoustical Society of America,* 63, 442-455.

Lopez-Poveda, E.A., and Meddis, R. (1996). A physical model of sound diffraction and reflections in the human concha. *Journal of the Acoustical Society of America,* 100, 3248-3259.

Louage, D.H.G., van der Heijden, M., and Joris, P.X. (2005). Enhanced temporal response properties of anteroventral cochlear nucleus neurons to broadband noise. *Journal of Neuroscience,* 25(6), 1560-1570.

Lynch, T.J., Nedzelnitsky, V., and Peake, W.T. (1982). Input impedance of the cochlea in cats*. Journal of the Acoustical Society of America,* 72, 108-130.

MacDonald, J., Henry, P., and Letowski, T. (2006). Spatial audio through a bone conduction interface. *International Journal of Audiology*, 45, 595-599.

Masterton, B., Jane, J.A., and Diamond, I.T. (1967). Role of brainstem auditory structures in sound localization. I: Trapezoid body, superior olive, and lateral lemniscus. *Journal of Neurophysiology,* 30, 341-359.

Masterton, B., Thompson, G.C., Bechtold, J.K., and RoBards, M.J. (1975). Neuroanatomical basis of binaural phase-difference analysis for sound localization: A comparative study. *Journal of Comparative and Physiological Psychology,* 89(5), 379-386.

Møller, A.R. (1962). Acoustic reflex in man. *Journal of the Acoustical Society of America,* 34, 1524-1534.

Møller, A.R (1974). Function of the middle ear. In: Keidel, W.D., and Neff, W.D. (Eds.), *Handbook of Sensory Physiology,* Volume V, *Auditory System*. Berlin: Springer Verlag.

Møller, A.R. (2000). *Hearing. Its Physiology and Pathophysiology*. New York: Academic Press.

Nedzelnitsky, V. (1980). Sound pressure in the basal turn of the cat cochlea. *Journal of the Acoustical Society of America,* 68, 1676-1689.

Oertel, D., Bal, R., Gardner, S.M., Smith, P.H., and Joris, P.X. (2000). Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. *Proceedings of National Academy of Science,* 97(22), 11773-11779.

Oliver, D. (1997). Anatomy of the central auditory nervous system. In: Crocker, M.J. (Ed.), *Encyclopedia of Acoustics*, 1, 1381-1388.

Ostapoff, E.M., Feng, J.J., and Morest, D.K. (1994). A physiological and structural study of neuron types in the cochlear nucleus. II. Neuron types and their structural correlation with response properties. *Journal of Comparative Neurology,* 346, 19-42.

Pfeiffer, R.R. (1966). Classification of response patterns of spike discharges for units in the cochlear nucleus: tone burst stimulation. *Experimental Brain Research,* 1, 220-235.

Pickles, J.O. (1988). *An introduction to the physiology of hearing* (2nd Ed.). New York Academic Press.

Puria, S., Peake, W., and Rosowski, J. (1997). Sound-pressure measurements in the cochlear vestibule of human-cadaver ears. *Journal of the Acoustical Society of America*, 101, 2754-2770.

Rhode, W.S., and Smith, P.H. (1986). Encoding timing and intensity in the ventral cochlear nucleus of the cat. *Journal of Neurophysiology,* 56(2), 261-286.

Roffler, S.K., and Butler, R.A. (1968). Localization of tonal stimuli in the vertical plane. *Journal of the Acoustical Society of America,* 43, 1260-1266.

Rose, J.E., Brugge, J. F., Anderson, D.J., and Hind, J.E. (1967). Phase locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey, *Journal of Neurophysiology,* 30, 769-793.

Rosowski, J.J. (1996). Models of external- and middle-ear function. In: Hawkins, H.L., McMullen, T.A., Popper, A., and Fay, R. (Eds.), *Auditory Computation*. New York: Springer Verlag.

Rosowski, J.J., Mehta, R.P., and Merchant, S.N. (2004). Diagnostic utility of Laser-Doppler Vibrometry in conductive hearing loss with normal tympanic membrane. *Otology and Neurology*, 25(3), 323-332.

Rutherford, W. (1886). The sense of hearing. *Journal of Anatomy and Physiology,* 21, 166–168.

Scharf, B., Magnan, J., and Chays, A. (1997). On the role of the olivocochlear bundle in hearing: 16 case studies. *Hearing Research*, 103, 101-122.

Shaw E.A.G. (1974). The external ear. In: Keidel, W. D., and Neff, W.D. (Eds.). *Auditory System: Anatomy Physiology (Ear)*. Berlin: Springer-Verlag.

Shaw, E.A.G. (1997). Acoustic characteristics of the outer ear. In: Crocker, M.J. (Ed.), *Encyclopedia of Acoustics*, 1, 1325-1336.

Simmons, F.B. (1964). Perceptual theories of middle ear muscle function. *Annals of Otology, Rhinology, and Laryngology,* 73, 724-739.

Sohmer, H., Freeman S, Geal-Dor M, Adelman, C., and Savion, I. (2000). Bone conduction experiments in humans: A fluid pathway from bone to ear. *Hearing Research,* 146, 81–88.

Spirou, G.A., Davis, K.A., Nelken, I., and Young, E.D. (1999). Spectral integration by type II interneurons in dorsal cochlear nucleus. *Journal of Neurophysiology,* 82, 648-663.

Spoendlin, H., and Schrott, A. (1989). Analysis of the human auditory nerve. *Hearing Research,* 43, 25-38.

Stebbins, W.C., Hawkins, J.E., Johnsson, L-G., and Moodty, D.B. (1979). Hearing thresholds with outer and inner hair cell loss. *American Journal of Otolaryngology*, 1, 15-27.

Stenfelt, S. (2006). Middle ear ossicles motion at hearing thresholds with air conduction and bone conduction stimulation. *Journal of the Acoustical Society of America*, 119, 2848-2858.

Stenfelt, S. (2007). Overview and recent advances in bone conduction physiology. In: A. Huber, and A. Eiber (eds.), *Middle Ear Mechanics in Research and Otology*. Zurich: World Scientific Publishing Company.

Stenfelt, S., and Goode, R.L. (2005a). Bone conducted sound: Physiological and clinical aspects. *Otology and Neurotology*, 26, 1245–1261.

Stenfelt, S., and Goode, R.L (2005b). Transmission properties of bone conducted sound: Measurements in cadaver heads. *Journal of the Acoustical Society of America*, 118, 2373-2391.

Todd, N.P.M., and Cody, F.W. (2000). Vestibular responses to loud dance music: A physiological basis of the "rock and roll threshold"? *Journal of the Acoustical Society of America*, 107, 496-500.

Tonndorf, J. (1966). Bone conduction. Studies in experimental animals. *Acta Otolaryngologica*, 213(Suppl.), 1-132.

Tonndorf, J. (1968). A new concept of bone conduction. *Archives of Otolaryngology*, 87, 595-600.

Tonndorf, J., and Khanna, S.M. (1970). The role of the tympanic membrane in middle ear transmission. *Annals of Otology, Rhinolaryngology, and Laryngology*, 79, 743-753.

Tonndorf, J., and Khanna, S.M. (1972). Tympanic membrane vibrations in human cadaver ears studied by time-averaged holography. *Journal of the Acoustical Society of America*, 52, 12221-1233.

Voss, S.E., and Allen, J.B. (1994). Measurements of acoustic impedance and reflectance in the human ear canal. *Journal of the Acoustical Society of America*, 95, 372-384.

Wada, H. (2007). Recent findings on our auditory system: It is highly sensitive owing to the motility of sensory cells. *International Journal of Acoustics and Vibration*, 12, 75-82.

Wever, E.G. (1949). *Theory of Hearing* (6th Ed). New York: Wiley and Sons.

Welgampola, M., Rosengren, S., Halmagyi, G., and Colebatch, J. (2003). Vestibular activation by bone conducted sound. *Journal of Neurology, Neurosurgery, and Psychiatry*, 74(6), 771–778.

Wright, D., Hebrank, J.H., and Wilson, B. (1974). Pina reflections as cues for localization. *Journal of the Acoustical Society of America*, 56, 857-962.

Yost, W.A. (2000). *Fundamentals of Hearing*. New York: Academic Press.

Zhak, S., Mandal, S., and Sarpeshkar, R. (2004). A proposal for an RF cochlea. *Proceedings of the Asia Pacific Microwave Conference (2004APMC)*. New Delhi: IEEE.

Zwislocki, J. (1957). Some impedance measurements on normal and pathological ears. *Journal of the Acoustical Society of America*, 29: 1312-1317.

Zwislocki, J. (1970). *An acoustic coupler for earphone calibration.* Special Report LSC-S-7, Laboratory of Sensory Communication, Syracuse University, Syracuse, NY: Syracuse University.

Zwislocki, J. (1975). The role of the external and middle ear in sound transmission. In: Tower, D.B. (Ed.). *The Nervous System. Volume 3: Human Communication and Its Disorders*. New York: Raven Press.

Part Four

Perception, Cognition and Performance

The sensory systems provide information that is processed by higher brain areas to promote perceptual understanding of the world. An HMD “tricks” these perceptual processes by providing more information about the world than the sensory system normally has or by presenting the information in a different format. It is thus important to understand how the perceptual systems respond to these new kinds of sensory environments. Perceptual experiences also are interpreted and analyzed by higher-level systems that carry out a variety of cognitive tasks such as: attention, memory, recognition, language, decision-making and problem solving. It is important to understand how these cognitive systems operate, because they often depend on details of the perceptual and sensory information.

10 VISUAL PERCEPTION AND COGNITIVE PERFORMANCE

José E. Capó-Aponte
Leonard A. Temme
H. Lee Task
Alan R. Pinkus
Melvyn E. Kalich
Allan J. Pantle
Clarence E. Rash

The Warfighter in the modern battlespace has a predetermined, but ever-changing, set of tasks that must be performed. Performance on these tasks is affected strongly by the amount and quality of the visual input, as well as by the resultant visual perception and cognitive performance. Visual perception is defined as the mental organization and interpretation of the visual sensory information with the intent of attaining awareness and understanding of the local environment, e.g., objects and events. Cognition refers to the faculty for the human-like processing of this information and application of previously acquired knowledge (i.e., memory) to build understanding and initiate responses. Cognition involves attention, expectation, learning, memory, language, and problem solving.

The direct physical stimuli for visual perception are the emitted or reflected quanta of light energy from objects in the visual environment that enters the eyes. It is important to understand that the resulting perception of the stimuli is not only a result of their physical properties (e.g., wavelength, intensity, and hue) but also of the changes induced by the transduction, filtering, and transformation of the physical input by the entire human visual system.

This chapter explores some of the more important visual processes that contribute to visual perception and cognitive performance. These include brightness perception, size constancy, visual acuity (VA), contrast sensitivity, color discrimination, motion perception, depth perception and stereopsis. An analogous discussion of input via the auditory sense is discussed in Chapter 11, *Auditory Perception and Cognitive Performance*.

Brightness Perception

In physics, the luminance of an object is exactingly defined as the luminous flux per unit of projected area per unit solid angle leaving a surface at a given point and in a given direction. A more useable definition is the amount of visible light that that reaches the eye from an object. But, when an observer describes how "bright" an object appears, he/she is describing his/her brightness perception of the object. This brightness is the perceptual correlate to luminance and depends on both the light from the object and from the object's background region.

Human visual perception of brightness and lightness involves both low-level and higher levels of processing that interact to determine the brightness and lightness of parts of a scene (Adelson, 1999).[1] If a scene was scanned by a photodetector, it would measure the amount of luminance energy at each point in the scene; the more light coming from a particular part of the scene the greater the measured value. The human eye's retinal receptors (cones) respond in a similar manner when a scene is imaged unto it. However the appearance (perception) of a region of the scene can be drastically altered without affecting the response of retinal receptors. The well-known simultaneous contrast effect demonstrates this phenomenon (Figure 10-1). In reality, the two center regions have

[1] *Brightness* is the perceptual correlate of luminance and may be thought of as perceived luminance; *Lightness* is the perceptual correlate of reflectance and may be thought of as perceived reflectance.

the same luminance, but their apparent greyness' (luminance) are different and depend upon spatial interactions with the surround. The grey region surrounded by a dark area looks (is perceived) brighter than the same grey region surrounded by a light region. Hering (1878) attributed this effect to adaptation and local interactions. This phenomenon is just one example of a number of illusions that illustrate problems that can arise when one visual element is viewed in the context of others. While the human visual system is very good at such complex tasks as edge detection and compensation for ambient lighting conditions, it sometimes can alter the appearance of the stimulus in unexpected ways before its message reaches the conscious part of the brain (Flinn, 2000).

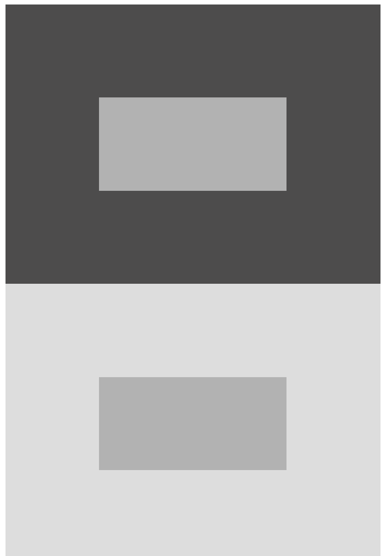


Figure 10-1. The simultaneous contrast effect.

The illusion associated with simultaneous contrast is not confined to grayshade images; it is equally applicable in the presence of color. Color perception has a strong dependency on two adjacent colors (Dahl, 2006). Figure 10-2 illustrates the different perception of the same blue color tone with two different backgrounds (Witt, 2007). While in Figure 10-2a blue is perceived as dark and opal, the same blue in Figure 10-2b is perceived as bright.

Two side-by-side colors interact with one another and change our perception according. Since colors rarely are encountered in isolation, simultaneous contrast will affect our perception of the color that we see. Consider a realistic example involving red and blue flowerbeds adjacent to one another in a garden; their perceived colors will be modified where they border each other. The blue will appear green, and the red will appear orange. The real colors are not altered; only our perception of them changes. Simultaneous contrast affects every pair of adjacent colors. This illusion is strongest when the two colors are complementary colors. Complementary colors are pairs of colors, diametrically opposite on a color circle (wheel) (Figure 10-3). Yellow complements purple; if yellow and purple lights are mixed, white light results. In the example of the red and blue flowerbeds, the red bed makes the blue bed seem green because it induces its complementary color, green, in the blue bed. The blue bed makes the red bed seem orange because it induces its complementary color, yellow, in the red bed.

When presenting information on helmet-mounted displays (HMDs) and other displays, this phenomenon of simultaneous contras is an important user interface design consideration The surroundings of a area of color will not only affect color brightness perception but also hue. This important property of adjacent colors should be considered in user interface designs and particularly where colors could be best used in structuring simple interfaces (Witt, 2007).

Figure 10-2. Simultaneous color contrast effect (adapted from Witt, 2007).

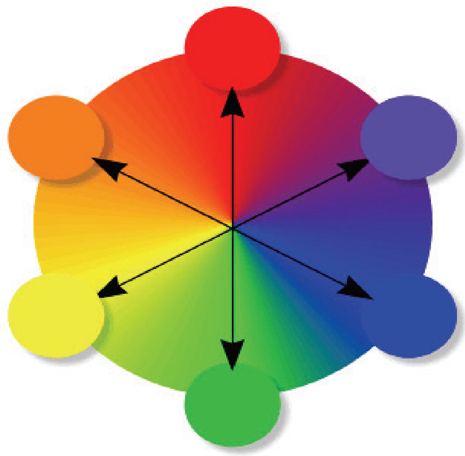


Figure 10-3. Complementary colors on color circle (wheel).

Size Constancy

Size constancy is the recognition that the same object viewed at different distances and orientations is interpreted and can appear to be the same size and shape, regardless of image changes at the retina due to distance, visual angle and perspective. This is usually combined with the easy, routine human ability to respond to the object appropriately. Size constancy labels a large percentage of the perceptual and cognitive processes that provide a stable view of the world. It has been the subject of investigation since the ancient Greeks with many seminal papers that provide excellent discussions on the issues associated with perceptual constancy (e.g., Blake and Sekuler, 2005; Cutting and Vishton, 1995; Epstein, Park and Casey, 1961; Graham, 1966; Luo, 2007; Roscoe, 1984; Stevens, 1951; Wagner, 2006a, 2006b; Woodworth and Schlosberg, 1954; Zalevski, Meehan and Hughes, 2001). A consensus of these papers and their historical reviews is that "There is no such thing as an impression of size apart from an impression of distance" (Gibson, 1950).

There are very practical reasons for understanding how we reliably relate our representative perceptions to objective space when there is a less-than-"transparent" device like a HMD in-between. A prominent individual once asked what the value was of studying vision in aviation. The simple answer is, try flying without it. A more nuanced response is that we don't always see things as they are and we need to know how to deal with that. Humans survive because of our ability to figure out what is in the environment and respond in suitable ways.

As elegantly described by Cutting and Vishton (1995), we understand the layout of objects in space, their size and distance, by using multiple sources of information weighted in a hierarchical fashion that is based largely on information availability, task, and logarithm of distance. We actively work to assemble a functionally accurate representation of objective space and the layout of objects it contains.

Our ability to assemble the information sources necessary to provide a useful perceptual representation of layout under continuously changing conditions depends on redundancy to guard against failure of information sources and on the ability to correct errors (Cutting and Vishton, 1995). HMDs usually constrain or degrade this active assembly process (e.g., reducing field-of-view (FOV), reducing contrast, reducing resolution). These degradations have an impact on our ability to create an accurate picture of what is out there and where it's located, thereby allowing appropriate behavior (Zalevski, Meehan and Hughs, 2001). Redundancy allows flexibility and the ability to adapt to an amazing variety of situations by assembling reliable sets of information sources. Witness a pilot's ability to adapt when using the Integrated Helmet and Display Sighting System (IHADSS), a monocular display used on Apache helicopters. It displays visually degraded imagery and symbology with a narrow FOV that requires active suppression of the image in one eye to avoid binocular rivalry.

Hyperstereopsis provides another example of how HMDs can impact the perception of objects. It is created when image intensifier (I^2) tubes are mounted temporally on the sides of a helmet (with a separation distance greater than normal) and their images frontally displayed on a combiner. Figure 10-4 shows how such a design can paradoxically make an object appear closer and smaller. Normally when an object is closer it forms a larger image on the retina.

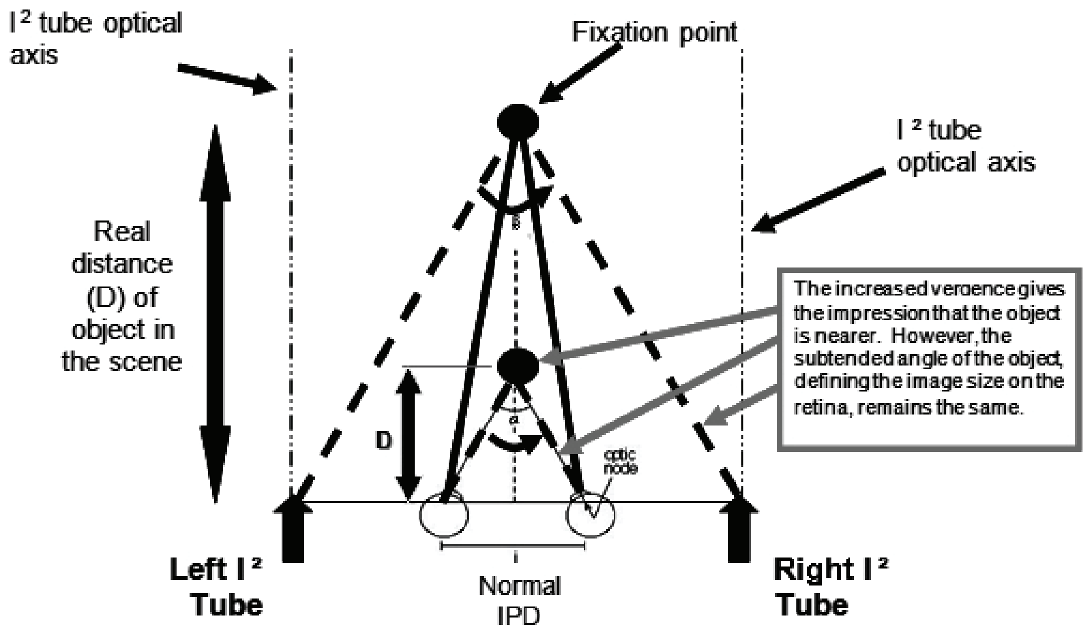


Figure 10-4. Size constancy is affected by hyperstereopsis when image intensifier (I^2) sensors are mounted on the sides of a helmet. Due to the apparent increase in interpupillary distance, near objects can paradoxically appear closer and smaller.

Another consequence of hyperstereopsis is diagramed in Figure 10-5. The near ground appears to rise up to the observer, while the ground farther away looks normal. This is because retinal disparity and convergence are reduced when viewing objects a few meters out and absent for greater distances.

It should be noted that there is considerable evidence that the distortions of object in visual space begin to wane with experience, as a pilot adapts to the impact of hyperstereopsis (Kalich et al., 2007; Priot et al., 2006). A recent study evaluating pilot debriefings from 3 pilots wearing a hyperstereo-producing HMD seemed to confirm this impression (Kalich et al., 2009).

Figure 10-5. Increased separation of I^2 tubes mounted on a HMD exaggerates horizontal, but not vertical perspective. The increased horizontal perspective makes near objects appear closer, as represented by the grid lines, creating a 'crater' illusion. The distant ground appears to level off due to reduced effects of convergence and retinal disparity.

Zalevski, Meehan and Hughes (2001) reviewed the effect of using binocular NVGs on size estimates. NVGs use electro-optical image intensification to amplify visible light and near infrared energy. The images created are monochromatic and have less resolution and contrast than we are used to during the day, consequently reducing the use of retinal disparity as a source of distance information. In addition, the images have a 'softer' appearance, and there is a random scintillation produced by electronic noise. The FOV of most modern binocular NVGs is 40^o. Combined with the degraded image, this increases the potential for spatial disorientation.

In general, as ambient light declines and images from NVGs deteriorate, the estimate of object distance increasingly relies on the visual angle of objects (Zalevski, Meehan and Hughes, 2001). Depth perception diminishes. As size and shape constancy depend on the availability of depth information, the perception of size constancy diminishes. Size constancy works best in an environment rich with depth cues,

The concept of retinal image size, combined with distance, provides a basis for size constancy (Figure 10-6). Epstein, Park, and Casey (1961) point out that this relationship manifests itself in two distinct ways. First is "an object of known physical size uniquely determines the relation of the subtended visual angle to apparent distance." Second, often called Emmert's Law, is that "the apparent size of an object will be proportional to distance when retinal size is constant."

Note that the issue of distance is central to both of these statements. In the first case, if we don't know the distance, due to reduced visual information, as when using NVGs with very low ambient light (starlight and or clouded night), we have to use visual angle subtended by objects. A large object objectively some distance away may be judged as smaller, and a smaller, near object may be judged to be farther away than it actually is. This could make an estimate of closing velocity problematical. Emmert's law is particularly important when using see-through HMDs for near work like surgery. The information on the display forms an image on the retina that is of constant size. This can interact with surfaces seen through a display. As a surface appears closer, the displayed information can appear smaller; as the surface appears farther away the image can appear larger.

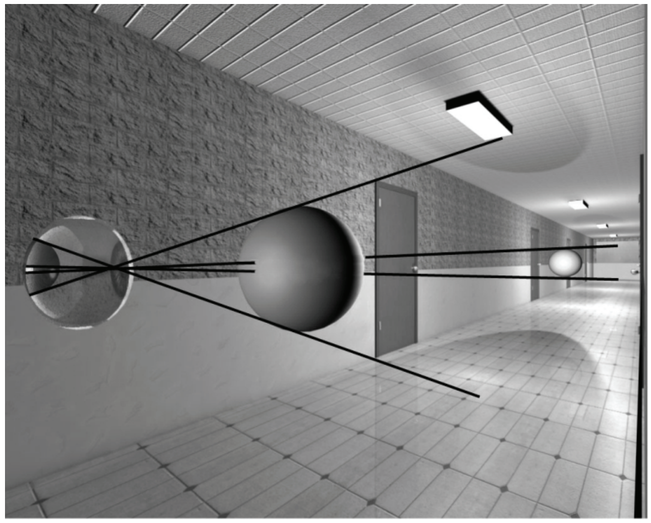


Figure 10-6. When an object, a sphere in this case, is viewed at different distances, the angle subtended at the eye, and correspondingly at the retina, is changed. Distant objects subtend a smaller visual angle and produce a smaller retinal image. Near objects subtend a larger angle and the retinal image appears larger.

Context also interacts with how we see and interact with objects. Context can make one distant object that subtends the same angle at the retina as another appear larger (Figure 10-7). By using movement and additional sources of information, we are usually able to arrive at a correct interpretation of the size of objects and their layout. However, when movement is restricted, as is the case with a pilot, it may be very difficult to obtain a correct interpretation of object size and distance.

Figure 10-7. A hallway rich with distance cues provides a context that makes the two identical black discs, ones that subtend the same visual angle, appear to be of different size. In most natural situation this can be corrected by changing position or using additional information, such as knowledge of their actual size.

Another aspect of size constancy is the use of information and memory (cognitive factors) to evaluate size (Blake and Sekuler, 2005). It is clear in Figure 10-8 that the people and cars that we identify that form a smaller angular subtense are behind the people who appear larger; and we behave accordingly. Environments rich in information sources provide many cognitive cues for distance and size. These are important for determining how we respond. Any contrivance placed between objective and representational visual spaces can reduce the number of sources of information about layout, decrease our ability to compensate for errors, and decrease chances for appropriate behavior as we try to navigate the real world.

Figure 10-8. Images of individuals and cars in this photograph that subtend smaller angles are normally treated as about the same size as the individuals and cars that are actually seen as larger. This is in large part due to our knowledge and memory. We also place the identified smaller object images in the background and the larger in the foreground, a depth interpretation. This ability to treat objects at different distances as the same actual size is critical when a pilot is on approach for landing.

Cutting and Vishton (1995) segment surrounding egocentric space into personal space (within 2 meters [m] [6 feet]), action space (within about 30 m [98 feet]), and vista space (beyond 30 m). The way we handle information, manipulate and deal with objects, the time frame of events, and the sources of motion differ in each of these egocentric regions In general, the order of relative dominance or efficacy of information about layout is occlusion, retinal disparity, relative size, convergence and accommodation for personal space; occlusion, height in the visual field, binocular disparity, motion perspective, and relative size for action space; and occlusion, height in the visual field, relative size, and aerial perspective for vista space. Each of these sources of information about layout can be divided into sources that are invariant with the logarithm of distance, sources that dissipate with the logarithm distance, and aerial perspective, increasing in effectiveness with logarithm of distance.

For example, occlusion, which is invariant with distance, almost always dominates, regardless of the egocentric region of operation. On the other hand, accommodation and convergence dissipate with distance and have little impact beyond personal space on assembling an accurate perception of layout. Similarly, the efficacy of retinal disparity, operating well into action space, also has reducing impact on how we assemble the sources of information to form a perception of the layout of perceptual space. A similar argument can be made regarding textural gradients. Even under conditions of hyperstereopsis, the impact of retinal disparity is significantly reduced beyond 30 m (98 feet) (Kalich et al, 2007).

The efficacy of ocular cues like convergence is significantly reduced beyond 6 m (20 ft), and beyond 30 m (98 ft). As ones' attention moves into action space and beyond, monocular sources of information such as interposition/occlusion, linear perspective, and motion parallax, increasingly dominate (Blake and Sekuler, 2005; Wagner, 2006a). In discussing this issue Zalevski, Meehan and Hughes (2001) state that motion parallax cues

> "…are most useful in visually complex environments such as open woodland and urban environments, and possibly less so over expanses of water or flat desert. Motion perspective, a cue resulting from the change in angular size of objects as they are approached (Braunstein, 1976), will be affected by the visibility and contrast of objects which, in the case of NVGs, is determined by illumination and reflectivity of objects. Another general source of spatial information is object familiarity, and cultural objects and structures such as vehicles and buildings on the ground can serve as "anchoring" cues for spatial perception, particularly object size."

Hermans (1937) very convincingly showed that convergence directly impacts the apparent size of objects. In general, objects requiring greater convergence appear smaller than the same objects viewed monocularly. As paired objects that are different distances from an observer, but angularly near to one another, move farther away, differences in their respective convergences are reduced. Consequently apparent size differences are also reduced. However, when using see-through binocular HMDs, as in a helicopter, convergence issues primarily apply in personal space within the cockpit. Leibowitz (1966) showed that the *greatest* effects of accommodation and convergence on apparent size operate at distances of one meter or less. When see-through HMDs are used in surgery there can consequently be very noticeable effects.

One factor that is of particular importance with HMDs and their use is the dominance of particular cues for distance. In general, accommodation and convergence have marginal or low dominance. This makes adaptation to a HMD much easier when the normal relation between accommodation and convergence is interrupted. Although this uncoupling can cause considerable discomfort, depending on the particular HMD under consideration, users visually adapt to use fairly well (Mon-Williams and Wann, 1998; Peli, 1995)

An issue that has been the source of much debate is whether visual space is best described as Euclidean or non-Euclidean (Wagner, 2006a). Euclidean geometry describes the local objective space we operate in and comes close, with considerable variability, to describing the visual space used in distance estimations. However, Wagner (2006b) concludes, after extensive review of the experimental literature, that our visual space and physical space are simply not the same. It may well be that Euclidean geometry best describes the space we constantly strive to approximate in our efforts to correctly constrain behavior.

The relationship between behavior and perception is not simple. Perception does not define behavior and is not the only thing that constrains it. A good example is the piloting of a helicopter. The relation between visual inputs, our perception of the world, our memory, our learned patterns of behavior, and the cognitive framework we are using all combine to help us perform some very subtle and indirect movements necessary to accurately guide the flight of a helicopter (Zalevski, Meehan and Hughes, 2001).

So, what is size constancy, and how is it important to the use of HMDs? It is a category of visual perceptions arrived at through multiple sources of information that are opportunistically assembled from moment to moment. We use our senses, our cognitive abilities, our memories, and our information to determine whether an object viewed at varying distances, from various perspectives and from various orientations is the same unvarying object in size and shape. This is important for our navigation through the environment, the identification of objects, the avoidance of harm, and the precise applications of our behavior. It is important that we be accurate and flexible enough to adapt to continuously changing environments, and it is fair to say that we have been.

HMDs affect our ability to assemble sources of information and thereby evaluate the layout of objects in our environment. The challenge to design engineers is to make the process as "transparent" (as easy and reliable) as possible.

Visual Acuity

Webster's Ninth New Collegiate Dictionary defines normal VA as the relative ability of the visual organs (eyes) to resolve detail that is usually expressed as the reciprocal of the minimum angular separation in minutes (of arc) of two lines just resolvable as separate and that forms in the average human eye an angle of one minute (of arc). The words in parenthesis were added for clarity. There are two important points that should be noted in this definition: 1) VA is a characteristic of the human eye and 2) the average (normal) human eye can resolve detail to about one minute of arc. The first point will be explored further in this section and the second point is addressed in a later section. (See Chapter 7, *Visual Function*, for addition reading on VA.)

It is apparent from this dictionary definition of VA that this parameter is a characteristic of the human eye that relates to the ability of the human eye to see detail. There is no mention of night vision goggles NVGs), HMDs, or other intervening viewing devices. In fact, implicit in the definition is the assumption that the only significant factor that affects the resolvability of the two lines is the human eye’s ability. How, then, can this parameter be used to describe a quality characteristic of a viewing device of HMDs such as the IHADSS and NVGs?

It is not uncommon to see reference to the “VA” of an HMD as a way to describe how good the display (and sensor) system performs. Usually, some viewing conditions are included within the VA statement such as: “This NVG has a VA of 20/25 under optimum light conditions and 20/50 under starlight conditions.” Strictly speaking, NVGs and other HMDs do not have, and cannot have, a VA, since they are nothing more than an image transducer or a viewing device. What is really meant when one refers to the VA of an HMD is that this is the expected VA of a normal observer when viewing through the HMD under the conditions described, since the concerns of interest usually revolve around the human-NVG system capability as a whole. This may seem like an unimportant, subtle difference, but it can have a real impact if one does not understand this difference. The implications of this difference will be addressed further in the section on measuring VA through NVGs.

The characterization of image quality of most displays, including HMDs (other than NVGs) usually includes some parameter that relates to the display’s capability to produce detail. Such parameters as resolution, number of pixels, pixel pitch or modulation transfer function (MTF) are used to convey information regarding the level of detail that one can expect the display to produce. Although NVGs contain image intensifier (I^2) tubes that are often characterized by their resolution or MTF, the NVG itself is almost always characterized by stating the VA. Even though this is something of a misnomer, if properly accomplished and reported, the “visual acuity of the NVG” (VA that can be achieved when viewing through the NVG) can be a useful parameter when comparing NVGs or determining what visual tasks can be accomplished using the NVGs.

Regardless of the potential usefulness or potential for error associated with the concept of VA of NVGs, it is a fact of life that it is a parameter that is often used and reported in the NVG community as a means of conveying information regarding the quality of the NVG, and it is not likely to disappear from usage any time soon. It can be a useful tool for comparing two NVGs and it can be a misleading factor if not properly understood. It is therefore important to understand what is meant by “visual acuity of the NVGs,” how it is measured, what units are used and how to convert between them, what affects it and how accurate it is. These are explored in the following sections.

Converting between visual acuity units used for HMDs

The definition cited above states that VA is the reciprocal of the separation of two lines, expressed in minutes of arc that can just be resolved by the eye. So, if two lines are separated by just one minute of arc when they are resolved then the VA would be 1 (no units) and if the separation were two minutes of arc, the VA is 0.5 and so forth. The reason for defining VA in terms of the reciprocal is to make larger numbers correspond to better capability (i.e., finer detail can be resolved). Although visual scientists tend to use this specific measure of VA, it is rarely used within the NVG and HMD communities. There are many different vision test charts and measurement units that are commonly used in assessing the VA of NVGs. This section describes the three most common measurement units and how to convert between them. Later sections describe the different vision charts and measurement procedures that are, or have been, used.

Three common units for specifying VA (through NVGs) are Snellen acuity (20/xx),[2] cycles per milliradian, and cycles per degree. Snellen acuity was primarily developed for fitting eye glasses and is normally associated with a vision chart composed of rows of letters that get smaller as one looks farther down the chart (Figure 10-9). Snellen acuity is always stated as the ratio of two numbers such as 20/20 (read as "twenty-twenty") or 20/40 (read as "twenty-forty"). The first number is the distance in feet that a test subject can read a particular chart line and the second number is the distance in feet that a "normal" person could see that same line. So, for example, if an individual can only see 20/40, this means he/she has to be 20 feet away from something that a normal person could see at 40 feet (twice as far away). In Europe the two numbers are based on the observation distance in meters instead of feet, and the first number is 6 (corresponding to 6 meters). Snellen acuity of 20/20 (normal vision) corresponds to Snellen acuity of 6/6 in European format.

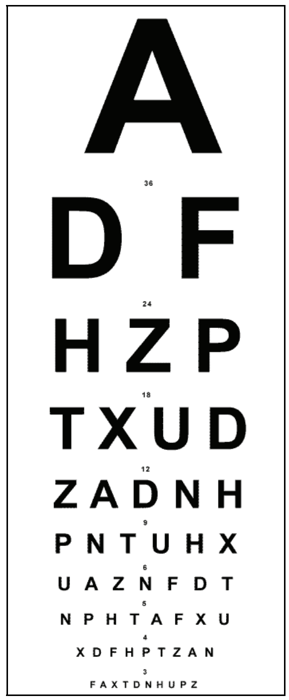


Figure 10-9. Snellen vision chart.

Snellen acuity is based on the assumption that a normal person can resolve high contrast detail that subtends one minute of arc (there are 60 arc minutes in 1°). This way of referring to VA is particularly popular with the users of NVGs, since they typically have a comfortable familiarity with Snellen acuity from their eye exams. Note that for Snellen units, the larger the denominator, the poorer the VA.

A second common unit of VA that is typically used by engineers in specifying and characterizing the NVGs is *cycles per milliradian*. This type of measure normally relates to a periodic type of vision chart such as a square-wave pattern or sine-wave pattern (Figure 10-10). A cycle refers to one dark and one bright bar of the pattern. So if the periodic vision chart were viewed from a distance such that the width of one dark bar plus the width of one light bar of the pattern subtends 1 milliradian, then the pattern would correspond to 1 cycle per milliradian.

[2] A number of tests have been developed for measuring visual acuity, but *Snellen acuity* has remained the standard. It does however have limitations. It also is important to note that many individuals can have better than 20/20 (6/6) "normal" vision.

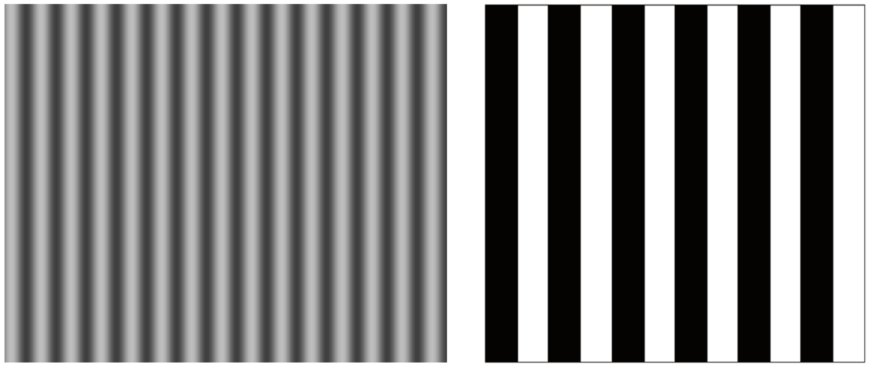


Figure 10-10. Sine-wave (left) and square-wave gratings (right).

The third unit that is occasionally used to characterize VA is *cycles per degree*. This unit is most commonly used by individuals that have a visual science background. Like the cycles per milliradian unit discussed above, it is also normally related to a periodic type of vision test chart. One cycle per degree means that one dark bar plus one light bar of the pattern subtends 1°.

While these different measures of VA were originally based on different types of vision test charts, it is possible to convert from one type of measure to another using certain widely-accepted assumptions. The basic assumption is that the minimum resolvable detail for normal vision is one minute of arc. The additional assumptions are that it takes two minutes of arc to resolve one cycle of a periodic pattern type vision chart, and that it takes five minutes of arc to resolve a Snellen letter. Using these assumptions, it is possible to derive equations that allow useful conversions between the different VA units. A convenient table for convert from one of these VA units to another is available in Barfield and Furness (1995).

Measuring visual acuity through NVGs

The term "resolution" is defined (the definition of interest for this topic) by Webster's Ninth New Collegiate Dictionary as "the process or capability of making distinguishable the individual parts of an object, closely adjacent optical images, or sources of light." As noted earlier, the same dictionary defines "visual acuity" as "the relative ability of the visual organ to resolve detail that is usually expressed as the reciprocal of the minimum angular separation in minutes of two lines just resolvable as separate and that forms in the average human eye an angle of one minute." It is apparent from these two definitions that "resolution" and "visual acuity" are connected but are not quite the same thing. This is, in effect, the difference between the VA "of" the NVGs (actually, the *resolution* of the NVGs) and VA "viewing through" the NVGs.

There is a subtle, but very real, difference between "NVG resolution" and "visual acuity through NVGs." This can be demonstrated by the following example. Suppose that some day advanced technology produces a "super" NVG capable of producing details down to a tenth of a minute of arc (well beyond normal human vision). If unaided (no magnification) vision is used to assess these "super" NVGs, we would get a reading of about 1 minute of arc (20/20 Snellen), since that is the limit of visual capability; even though the NVGs were producing details down to one tenth of this size (20/2). Thus, in this case, what is being measured is actually VA "through" NVGs and not the actual NVG resolution. As long as NVG capability is worse than human visual capability, there is not a significant difference between the two. However, even with today's NVGs, the difference between NVG resolution and NVG VA can be significant at low light levels. There are many combinations of vision test charts and assessment procedures that are used to determine NVG VA.

The Snellen chart displays rows of high contrast letters starting with a very large size (e.g. 20/200) and stepping down to the smallest (e.g. 20/10). Miller et al., (1984) used the Snellen eye chart to measure VA through NVGs.

The tumbling E (used by Wiley, 1989; Levine and Rash, 1989) chart has also been used to measure VA through NVGs. Some researchers (Kotulak and Rash, 1992) prefer to use the Bailey and Lovie (1976) eye chart, which has logarithmically spaced letter sizes.

One of the most frequently used resolution test standards is the 1951 Air Force tri-bar target (see Figure 12-11), which was originally developed as a tool to evaluate the optical performance of airborne reconnaissance systems (Military Handbook 141, MIL-HDBK-141, Defense Supply Agency [1962]). A conversion factor must be used to convert from the *Group* and *Element* number of the tri-bar chart to NVG VA.

NVG VA is determined by having a visually qualified, trained observer view the tri-bar pattern under specified illumination conditions (which may be between overcast starlight up to full moon illumination equivalent) and then state which Group and Element number he/she can "resolve." This is then converted to a Snellen acuity equivalent. When doing NVG evaluations, agencies may have 3 trained observers whose responses to this test are averaged to determine the "visual acuity" of the night vision goggles. Although the 1951 tri-bar target pattern has proved to be very useful over the years in comparing lens systems, it still has a certain amount of variance due to differences in observer criteria as to when the tri-bars are "resolved" (Farrell and Booth, 1984). Studies using the tri-bar pattern have shown observer response discrepancies of as much as 60% (Farrell and Booth, 1984).

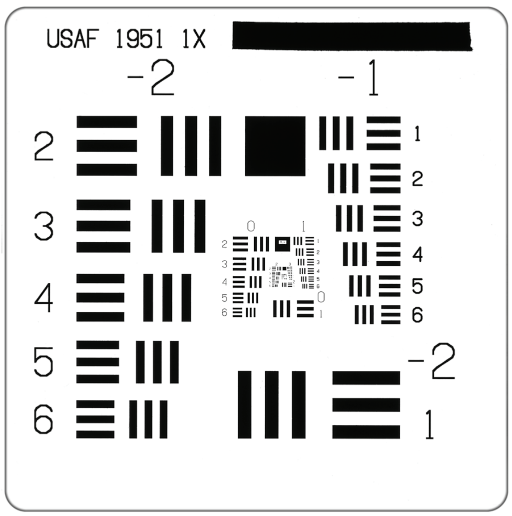


Figure 10-11. Air Force 1951 tri-bar resolution chart.

The 3x3 square-wave target array (Task and Genco, 1986) was developed as a means for pilots to do a quick verification that their NVGs were operating correctly and were capable of resolving detail to a specified level. The chart has nine square-wave patterns, arranged in a 3x3 array as shown in Figure 10-12 its standardized viewing distance of 20 ft., each pattern was sized to equal specific Snellen values of 20/20 through 20/60 in increments of five. To increase the number of randomized grating orientations for a repeated measurements test, the chart is simply rotated to any one of its four orientations, which has the effect of quickly changing grating locations and orientations within the 3x3 array. Charts having different levels of contrast were also constructed.

It should be noted that the step sizes between patterns are relatively large making this pattern unsuitable for comparing the capability of different NVGs that are somewhat close in their resolving power (i.e., VA).

An array of square-wave gratings to assess VA is also used in the Hoffman 20/20™ device. This device was designed for aircrew members to adjust their NVGs and verify that they have the minimum VA through the NVGs prior to flight (Angel, 2002). Figure 10-13 shows the device and the square-wave grating patterns that it displays. The gratings correspond to Snellen visual acuities of 20/20 through 20/70 with step sizes as shown. This is a subjective assessment method that is often used to determine the VA of the NVGs.

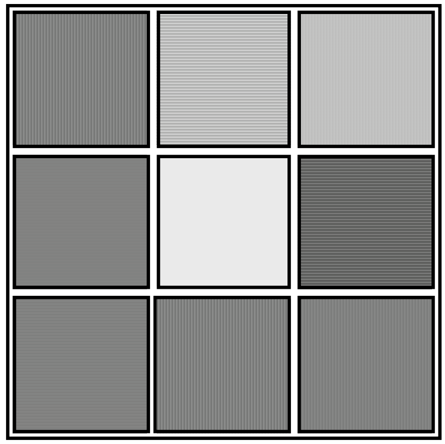


Figure 10-12. The 3x3 NVG chart (Task and Genco, 1986, US Patent 4,607,923).

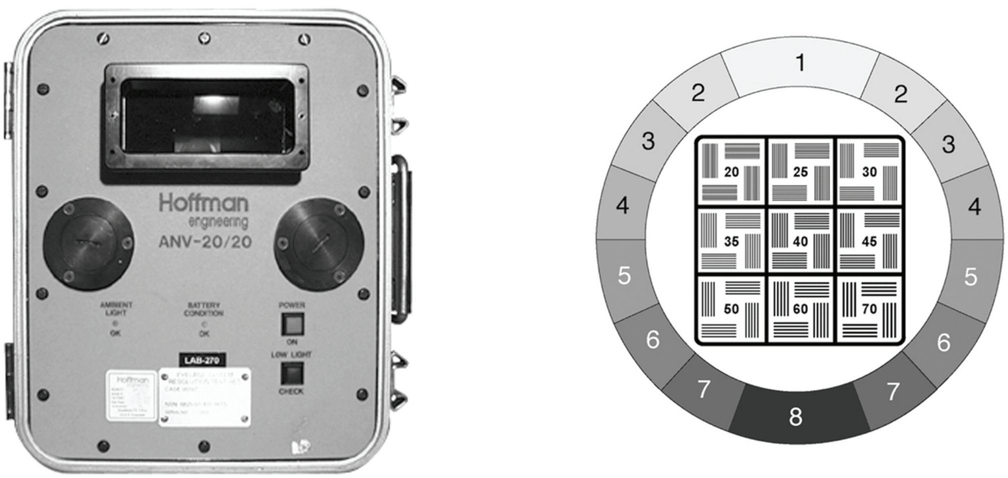


Figure 10-13. Hoffman Engineering ANV-20/20™ device (left) used to pre-flight NVGs. Pattern on the right is the array of square-wave gratings that is seen through the NVGs when the NVG objective lenses are positioned in front of the large, rectangular viewing port visible at the top of the picture on the left.

Another assessment method uses Landolt C stimuli (National Academy of Sciences, 1980). The Landolt C is a perfectly circular C (no serifs) that has a specified contrast and gap size. The gap size is varied as is the orientation. The observer's task is to detect the orientation of the gap. Pinkus and Task (1997) used closely sized Landolt C stimuli in a two-alternative, forced-choice (2AFC) method to determine VA through NVGs as a function of nighttime ambient illumination levels. A computer executed the 2AFC (gap seen up or down), using a Step Program adapted from Simpson (1989). Based on the observer's last response, the program selected the specific gap size (smaller or larger) of the next Landolt C to be presented, according to *a priori* rules inherent in the algorithm. This method allowed relatively efficient convergence to threshold acuity usually within 10 to 35 trials. The step method yielded reasonable results, but informal repeatability tests found that the observer's scores varied from day to day. These variations could be due to a number of variables: working at threshold levels, NVG drift, good guessing in the 2AFC method, fatigue, eye strain, sinus headaches and so on.

In summary, there have been numerous test charts and targets to assess VA including the Snellen chart, square-wave gratings, sine-wave gratings, tumbling E, 1951 USAF Tri-Bar chart and Landolt C. These have been used with several assessment procedures including both objective procedures and subjective procedures. The quasi-objective procedures, such as the two-alternative forced-choice method described above, require the subject to provide information about the target type that would only be reliably available if the subject could actually "resolve" the critical characteristic of the target type used. For example, which way the gap is oriented in a Landolt C or which way the arms of the E are pointed in a tumbling E target. Subjective measures involve the subject making a judgment that they can or cannot resolve the critical detail of the target. An example of a subjective assessment procedure is when a subject reports which group and element number of a USAF 1951 Tri-Bar chart he/she can just barely resolve. In general, objective tests should provide more accurate data but take much longer to accomplish. Both subjective and objective assessment results can depend heavily on the specific subjects that participate in the assessments. In general, better results are obtained if more subjects are used in the assessment (ideally at least 3, if possible) and the subjects are trained or have substantial experience in the assessment procedure.

Measuring visual acuity through HMDs connected to remote sensors

A person seldom sees explicit references to the VA of a HMD that is connected to a remote sensor, such as the IHADSS HMD on the AH-64 Apache helicopter. However, providing an acuity value for thermal forward-looking infrared (FLIR) sensor-based systems (e.g., the AH-64's Pilot's Night Vision System [PNVS]) is difficult since the parameter of target angular subtense is confounded by the emission characteristics of the target being viewed. This is not unlike the difficulty of determining the VA through NVGs for different ambient lighting conditions (see following section on conditions affecting NVG VA results). For comparison purposes, Snellen VA with the AH-64 PNVS/IHADSS is cited as being 20/60 (Greene, 1988).

Whether the sensor is a FLIR or a low light level TV or a short-wave infrared (SWIR) device the primary determinant of what one can expect in the way of VA (ability to see detail) is typically a combination of the capability of the HMD optics and image source with the sensor optics and detector array. If the FOV of the sensor is identical to the FOV of the HMD (which it should be for piloting-type tasks) then the VA expected through the system is determined by the angular subtense of the smallest detail that can be resolved through the entire system compared to one minute of arc. In the case of the AH-64 PNVS/IHADSS (HMD and sensor have the same FOV), which had a Snellen acuity of 20/60 (noted above), the observer was presumably able to resolve details to approximately three minutes of arc.

In the unusual situation where the sensor FOV is not the same as the HMD FOV (such as systems that produce magnification by making the sensor FOV narrower than the HMD FOV), there is can be an ambiguity in determining the effective VA. The basic issue is whether to use one minute of arc in the HMD FOV as a reference or one minute of arc in the actual, real world geometry as a basis. For example, if the sensor FOV was 1/5th of the HMD FOV (producing a magnification of 5X) and the sensor could resolve objects that were one arc minute in size as measured from the sensor then this would subtend 5 minutes of arc in the HMD. So, should the "visual acuity" be stated as 20/100 (HMD FOV referenced) or as 20/20 (real world geometry referenced)? There are arguments for each way that are beyond the scope of this discussion. Suffice it to say that if the system provides magnification with respect to the real world, then it is necessary to always state which reference (HMD FOV or real world geometry) was used to quote the "visual acuity" of the HMD-sensor system.

Conditions affecting NVG visual acuity results

The primary reason for measuring NVG VA is to obtain information regarding the image quality capability of the NVG. However, because the assessment procedure involves not only the NVGs but also a human observer and is accomplished under some ambient or artificial environmental conditions, the results are due to the combination of these three factors. There are several parameters contained within each of these factors that can affect the NVG VA results obtained, as noted in the following sections.

NVG parameters that can affect NVG visual acuity

Gain, maximum luminance, signal to noise ratio (SNR), objective lens quality (e.g., MTF), objective lens focus setting, I^2 tube micro-channel plate pitch, fiber optics twister (if any) quality, eyepiece lens MTF, eyepiece focus setting (diopter adjustment), and eye motion box size and quality can all affect the results obtained when assessing VA through NVGs (Figures 10-14). While all of these parameters are fundamental characteristics of the NVG, only a few of them have an effect on the NVG VA assessment that is totally independent of the human observer. Most of them involve an interaction with the way in which the human eye operates.

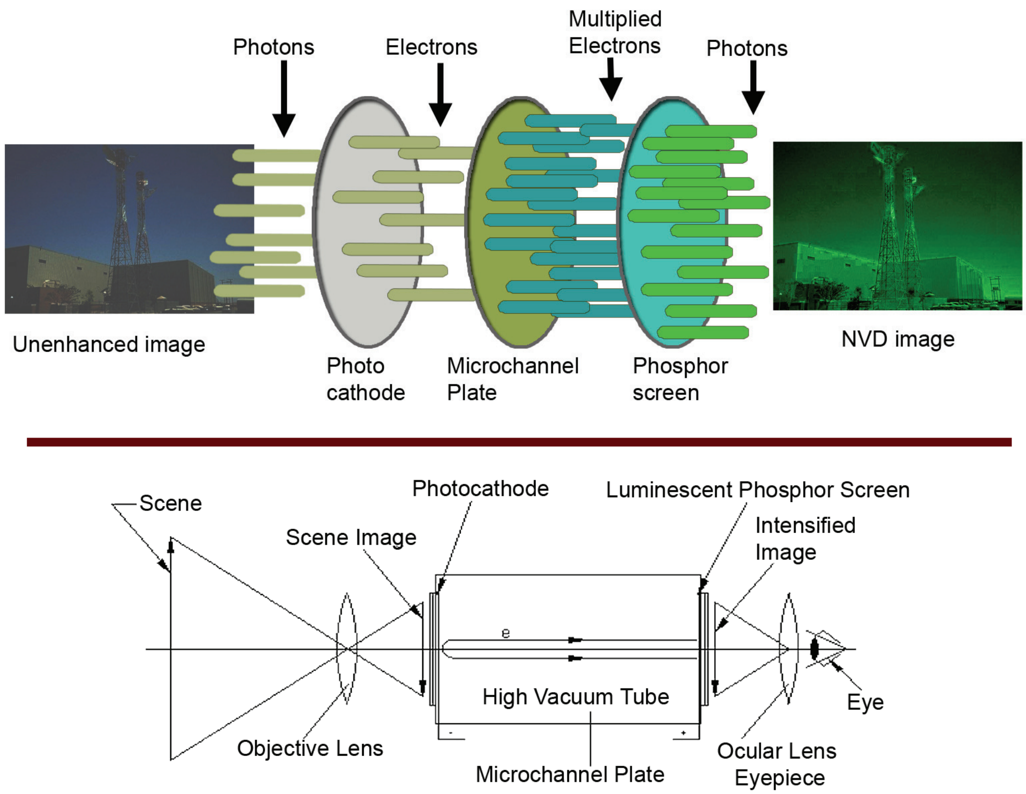


Figure 10-14. Operation of an image intensifier tube.

The gain of an NVG is the ratio of the input luminance to the output luminance for a light source that has a spectral distribution equivalent to a 2856K° blackbody emitter. This is actually an oversimplification of NVG gain, but the main point here is that, in general, the output luminance (what the eye is going to see) is higher for NVGs that have higher gain values for the same input (ambient scene) radiance. This assumes that the ambient radiance conditions are low enough that the I^2 tube within the NVG is operating at maximum gain (the automatic gain control circuitry is not activated). Under these conditions, NVGs with higher gain will have a higher output luminance. Since at these low NVG output luminance levels (on the order of a few thousandths to a few tenths of a foot-Lambert [fL]), the VA of the human eye is improved as luminance is increased, it is apparent that VA is better with higher NVG gain.

The maximum output luminance of the NVG is typically determined by circuitry within the I^2 tube power supply system, which limits total current to some maximum value. If there is sufficient ambient radiance that this circuitry is activated, then NVGs with a higher maximum output luminance should result in better VA for the reason stated above.

The SNR of the NVG is a result of several factors. In general, the higher the SNR the better VA one will obtain (Riegler et. al., 1991) since the masking effect of the noise is reduced.

The imaging quality of the objective lens of the NVG oculars can also affect the resultant VA. The objective lens (the lenses on the front of the NVGs) produces an image of the outside world scene onto the photo-cathode of the image intensifier tube. The "sharpness" of this image depends chiefly on MTF[3] of the objective lens, and in general, the better the MTF, the better the VA (up to a point). It should also be noted that the MTF is typically different for different parts of the image. In general, the MTF is better at the center of the image and becomes worse as one looks further out from the center of the image towards the edges. This is often the main reason that the VA obtained through NVGs is better in the center of the image than at the edges (other factors typically don't vary across the image as much as the MTF does).

Another factor that can have a significant effect on the VA through the NVGs is the objective lens focus setting (Pinkus and Task, 2000). Because of the very low f-numbers (ratio of focal length of lens to the diameter of the lens), the "sharpness" of the image produced by the objective lens can suffer significantly if the focus adjustment isn't set correctly. Note that this is not the same as the MTF (which is determined under the assumption that the focus setting is correct). However, the focus adjustment effect on the VA is similar; namely, it produces a blurry image on the photo-cathode of the I^2 tube for which nothing else in the imaging chain can compensate.

At the heart of the image intensifier of present day NVGs is a micro-channel plate (MCP) that is the workhorse in amplifying the image signal. The MCP is a thin disc that has many thousands of tiny holes each of which acts like a miniature photo-multiplier tube. These individual holes are essentially the *pixels* (picture elements) of the I^2 tube. Although there is an interaction with the eyepiece lens focal length, in general, the more holes the MCP has and/or the closer together these holes are, then the better the VA obtained when viewing through the NVGs.

Most NVGs produced today require a fiber optics *twister* to produce an image that appears upright to the viewer. As its name implies, this twister rotates the output image 180° (±) with respect to the input image. It does this by means of thousands of tiny fibers each one of which could be considered a pixel similar to the MCP holes. In general, the closer these fibers are to each other (achieved through smaller fiber diameters) the better VA one should obtain. It should be noted that typically the quality and size of the fiber optics twisters currently produced result in a much better pixel count and pixel pitch (basically the distance between individual pixels) than the MCP. This means that typically the fiber optics twister is not a significant factor in limiting VA through NVGs, although it theoretically could be.

The eyepiece lens is the final lens in the NVG optical train and is the lens the eye looks through to see the output image from the I^2 tube. Just like the objective lens, the eyepiece lens has an MTF that can influence VA. Because of the limiting effects of the human eye's entrance pupil, the impact of the eyepiece lens MTF on VA is usually not significant. However, if the eye's pupil is not positioned along the center of the optical axis of the eyepiece lens, one can experience a rapid deterioration of the MTF. This is related to the concept of the *eye motion box,* which is the zone within which the eye pupil should be positioned in order to have an acceptable level of image quality. Outside of this zone the MTF can drop off rapidly resulting in poor or blurry image quality corresponding to worse VA. In general, better VA is obtained for eyepieces with better MTFs and with larger eye motion boxes.

Many NVGs currently produced permit the operator to adjust the eyepiece focus. This is also frequently called the diopter adjustment or diopter setting. The eyepiece lens produces a virtual image of the output of the I^2 tube. The apparent distance of this image from the viewer is determined by the eyepiece diopter setting. The apparent distance in meters is calculated by taking the reciprocal of the diopter setting value. For example, if the diopter setting is one diopter, the image will appear to be one meter away. Similarly, if the diopter setting is two diopters

[3] The modulation transfer function (MTF) is defined in this context as the sine-wave spatial-frequency amplitude response used as a measure of the resolution and contrast transfer of an imaging component, device or system.

the image will appear to be only 0.5 meter away (i.e., the reciprocal of two). This parameter of the NVG interacts with the viewer's ability to focus at the apparent distance associated with the diopter setting. There also may be some minor interaction with the MTF of the lens since this typically varies a small amount depending on the diopter setting. In general, VA improves as the diopter setting is adjusted correctly for a particular user's eyes (Angel and Baldwin, 2004; Angel, 2003).

Although all of the parameters covered in this section relate directly to the characteristics of the NVGs, it is also apparent that many of them interact with the characteristics of vision. In general, better VA is obtained for NVGs with higher gain, higher SNR, better objective lens/eyepiece lens MTF, higher density holes in the MCP, higher density fibers in the fiber optics twister, better adjusted objective (focus) and eyepiece (image distance) settings, and optimized eye position within the eye motion box.

Human vision parameters (of the observer) that affect NVG visual acuity

Since the human visual system is an obvious integral part of any VA assessment through NVGs, it should be apparent that the visual capability of the specific user(s) is critical. Ideally, users should have excellent VA at the relatively low NVG output light levels (luminance of a few fL at most), since the objective of the test is to assess the NVGs, not the subject's vision. Other factors besides the user's innate VA (without NVGs) can also affect the test results. These include the user's dark adaptation state at the time of the test and whether or not the test is conducted binocularly (both eyes and NVG channels test simultaneously) or monocularly (testing one NVG channel at a time).[4]

A significant factor that can affect the VA obtained for an individual is the adaptation state. It takes the human eye a certain amount of time to recover (bio-chemically) when switching from a higher light level environment to a lower light level environment. For example, if one enters a movie theater on a bright day the movie screen appears to be very dim until the eyes have had a chance to adapt to the lower light level. The same effect can occur when assessing VA through NVGs if the observers go directly from a lighted room to viewing through the NVGs. Typically, this adaptation issue is resolved by requiring the subject to dark adapt for 10 to 20 minutes.

In addition to the relatively short adaptation state effect discussed above one can also encounter a longer term adaptation effect. If an individual spends a large amount of time during the day exposed to very high light levels, such as spending the day at the beach or snow skiing, then it may take more than just a few minutes to achieve full adaption; it could take several hours. (See Chapter 7, *Visual Function*, for addition reading on visual adaptation.)

There has been some evidence that the effects of smoking, which decreases the oxygen content in the bloodstream and therefore the oxygen getting to the retina, may result in poorer low-light VA compared to non-smokers (see Chapter 16, *Performance Effects Due to Adverse Operational Factors*).

Another significant impact on low-light VA can occur depending on whether or not the VA is being achieved (or measured) binocularly (both eyes at the same time) or monocularly (one eye at a time).[5] This interacts with the NVG characteristics in that if the two channels of an NVG are different because of some physical parameter (such as objective lens focus or MTF) the resultant VA obtained binocularly is typically governed by the image quality of the best NVG channel. In other words, if one conducts a binocular VA on an NVG it is possible to overlook a poor NVG ocular if the other ocular has produces good image quality. There are, therefore, advantages for conducting both binocular and monocular VA assessments of NVGs.

In general, one obtains improved VA values if the individual has good VA capability, is properly dark adapted, is a non-smoker, and the test is conducted binocularly (although, as noted, monocular testing has its own advantages).

---

[4] NVGs have a luminance output (brightness) that falls in the range associated with human *mesopic vision.* Therefore, wearers of NVGs are not fully dark-adapted.

[5] Standard NVGs are binocular, but several I^2-based HMDs have proposed a single tube design.

Environmental parameters that affect NVG visual acuity

Environmental parameters that are independent of both the NVGs and the user can affect the achieved VA. It has already been noted that with NVGs, human VA is better if the light level is higher. This is a fundamental characteristic of the human eye and does not really relate to the NVG's capability to produce a high-resolution image but rather the NVG's capacity to produce luminance. Environmental parameters that can affect the VA achieved with NVGs include NVG radiance level of the vision target and surrounding area, the type of vision target used (Landolt "C," Tri-Bar Chart, square-wave grating, etc.), the apparent contrast of the target (through the NVGs), degradation effects (e.g. glare off of the vision test chart or reflections from a windscreen or canopy), and the distance from the test chart to the NVGs.

The NVG-weighted radiance (Task and Marasco, 2003; 2004) of the vision chart and the gain of the NVGs determine the output luminance level, which in turn can affect the VA obtained (at least for lower radiance levels). Two typical NVG radiance values that are often used for NVG VA evaluation correspond to high moonlight level (full-moon or ¼-moon) and clear starlight. The higher radiance level is sufficiently high so that the NVG is in automatic gain mode and the output luminance is limited to the maximum luminance allowed by the circuitry. At these higher radiance levels, the NVG is providing its maximum output luminance, which is typically in the 2 to 4 fL range depending on the specific image intensifier tube used. At the lower radiance level the output luminance is dependent primarily on the gain of the NVG and is typically on the order of a few tenths of a fL for currently fielded NVGs. Lower input radiance levels that correspond to overcast starlight are also sometimes used resulting in output luminances that can be in the hundredths of a fL range. At these very low output luminance levels, the VA obtained can depend heavily on the low light VA capability of the subject.

The contrast of the vision test chart, and anything that degrades that contrast (glare and reflections), can significantly affect the NVG VA value obtained (Pinkus et. al., 2003). Test procedures for conducting a VA assessment through NVGs typically call for "high" (Department of Defense, 2001) or "medium" contrast charts.

In summary, all of the vision test charts and assessment procedures discussed in this section are useful and can provide some insight into the quality of NVGs or HMD systems. However, it cannot be stressed enough that these multiple VA test charts and procedures can produce different VA values for the same NVG or HMD. Therefore, while any of these procedures can be useful to compare NVGs or HMD systems, care must be taken when comparing VA values for different NVGs and HMDs if they were determined using different procedures and charts (and observers!).

Contrast Sensitivity

Exceptional vision is necessary to achieve high levels of performance under a wide range of viewing conditions. While all human senses are important to the Warfighter, vision is the only sensory system that is used to its fullest capacity during flight tasks (Swamy, 2002). Advances in HMDs allow Warfighters continuous 24-hour, all-weather operation (e.g., night and foul-weather) by using imaging sensor systems on aircraft, mounted vehicles, as well as on individual Warfighters. However, the amount of visual information that can be conveyed by the HMDs is essentially limited by the capacity of the human visual system to perceive contrast (i.e., difference in luminance). While wearing a HMD, optimum viewing conditions are achieved when the luminance of the display is matched to the capacity of the visual system (i.e., maximally sensitive). Optical devices can improve vision by decreasing the spatial frequency of an image or correcting the optical blur (e.g., glasses, contact lenses, refractive surgery), which results in better contrast at high spatial frequencies. Even though, visual enhancement HMDs provide Warfighters with tactical advantage during extended military operations, they can reduce contrast sensitivity and have the potential to decrease performance.

Although VA is often used to describe the quality of vision (i.e., level of spatial vision), contrast sensitivity appears to be a better indicator of visual performance under both, photopic (i.e., day) and scotopic (i.e., night) conditions; this is especially true for aviators (Rabin, 1993; van de Pol, 2007). The visual system depends on a series of visual channels that gather information regarding the object's size, shape, and contrast. The statistical distribution of these channels matches in general the distribution of important visual objects that humans need to navigate around and manipulate, i.e. it is peaked at about 4 cycles/degree, a factor of 5 below the visual system's highest resolution (i.e., around 20 cycles/degree). The collected information is relayed to the brain to create a complete picture. Unlike VA, that tests only one type of these visual channels, a contrast sensitivity test assesses multiple channels that are required to achieve exceptional functional vision. Thus, the visual function is not just acuity (resolution), but includes a combination of complex optical and neural aspects of our visual system. For example, an observer who has low contrast sensitivity may be able to read the small print on an eye chart but may still experience trouble seeing objects at night or in dim tactical or operational conditions. Accordingly, as a metric for spatial vision performance, contrast sensitivity can provide a more comprehensive index of visual function than VA, mainly because most "real world" visual scenes comprise a complex combination of *contrasts* and *spatial frequencies*, instead of isolated high-contrast/high-spatial frequency stimuli that are displayed in a VA test.

Contrast

In real situations, objects and their surroundings are of varying contrast. The ability of an observer to perceive the details of a scene is limited by the capacity of the visual system to discern contrast. As described in Chapter 7, *Visual Function*, a high contrast grating is always easier to see than low contrast gratings. The visual system achieved this level of perception by discriminating between luminosities of different levels in an image. The minimum contrast required to reliably detect the object from its background is known as the spatial contrast threshold. Contrast threshold is affected by several factors such as target size, background luminance, and viewing duration. Contrast threshold is the reciprocal of the contrast sensitivity, therefore the lower the contrast threshold the higher the contrast sensitivity and visual performance.

Optimum contrast and luminance of the imagery is required to optimize visual performance and prevent perceptual problems when wearing an HMD. In order for the symbology to be viewed in a see-through HMD or head-up display (HUD), the luminance of the symbology must be sufficient to discriminate it from the see-through real world scene (Harding, 2007). In addition, to prevent perceptual problems, both the virtual image projected on the see-through combiner lens of the HMD (e.g., Integrated Helmet and Display Sighting System used on the AH-64 Apache helicopter) and the real world scene must be clearly visible at the same time. In order to see both views clearly, they must be within the pilot's depth of field. The depth of field is the range of distances within which the different objects appear in sharp focus (Patterson, 2006) and this in turn will be affected by the focal distance at which the HMD has been set. As long as the optics of the HMD are collimated so that the images appear to lie at or near optical infinity, similar to the real world scene, both the virtual image and the real world scene will fall within the observer's depth of field and perceived to be in focus. When this is achieved, the virtual image will appear as being on the same plane as the real world scene (i.e., overlapping). The level of luminance also affects the depth of field. A decreased luminance level of the HMD induces a larger pupil diameter, which in turn results in a smaller depth of field (Ogle and Schwartz, 1959).

According to the Michelson definition of contrast, a minimum contrast (i.e., luminance ratio) level of 0.10 is required to discriminate the object from its background. Accordingly, if the monochrome imagery displayed on the HMD is viewed against the real world scene under scotopic conditions, the luminance of the image source must exceed 5,000 foot-Lamberts in order for the symbology to be discerned from its background created by the real world scene (Velger, 1998). In addition, the complexity of the real world scene in terms of contrast must be taken in consideration when determining the luminance specifications for HMDs (Harding, 2007). It has been suggested that the use of color symbology in HMDs has the potential to provide the Warfighter with a substantial operational advantage compared to the monochrome symbology (Martinsen and Havig, 2002). Although the development of color symbology is still ongoing, this technology is more complex and may require a tradeoff in resolution and luminance contrast in order to allow recognition of color symbology (Havig et al., 2001).

Spatial frequency

Contrast sensitivity is also dependent upon the size or spatial frequency of the features in the image. The visual system is more sensitive to contrast at certain spatial frequencies. The highest spatial frequency humans can see at any contrast is limited by the optical process. The concept of an optical transfer from the imaging system to the neural processing system has led to the development of the contrast sensitivity function (CSF). The CSF measures relative sensitivity versus spatial frequency and is accepted as a measure of assessing visual performance. Generally, high spatial frequencies gradients are harder to visualize than low spatial frequencies. However, this is not a direct relationship, as in some cases larger objects (lower spatial frequencies) are not always easier to see than smaller objects, as illustrated in Figure 10-15. This is also demonstrated by the CSF (Figure 7-11, Chapter 7, *Visual Function*) in which the sensitivity of the visual system to detect contrast decreases for lower and higher spatial frequencies. In those cases where the size of the object is not optimum—spatial frequency below two and above six cycles per degree (cpd) – the object's contrast needs to be increased in order to be discerned from the background. However, under photopic conditions, frequencies higher than 40 cpd are undetectable even at maximum contrast.

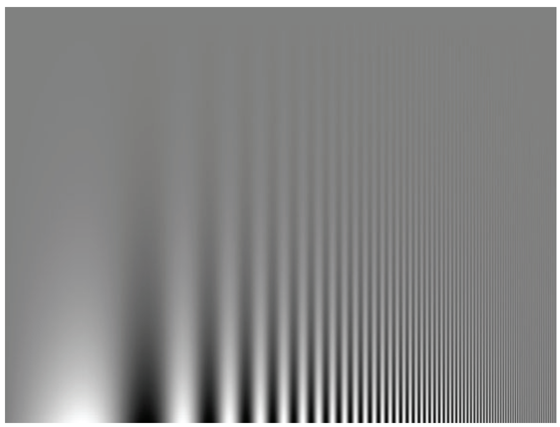


Figure 10-15. The human visual system is more sensitive to middle spatial frequencies. This illustration depicts a sine-wave grating in which spatial frequency increases exponentially from left to right, and the contrast increases logarithmically from 100% at the bottom to 0.5% at the top. At the top, the contrast is too low to see the grating to the point that only homogeneous grey is seen. Very wide (low spatial frequency) and very thin (high spatial frequency) gratings are harder to see than the middle bars, even with high contrast. (Courtesy of Dr. Izumi Ohzawa, University of California, School of Optometry). This figure was originally produced by F.W. Campbell and J.G. Robson, Applications of Fourier Analysis to the visible of gratings, *Journal of Physiology* (Campbell and Robson, 1968).

Scotopic contrast sensitivity

There is a marked difference between spatial contrast sensitivity under photopic and scotopic conditions. For instance, under scotopic conditions, frequencies higher than 8 cpd are undetectable even at maximum contrast. The contrast sensitivity of an aviator while wearing its night vision imaging systems (i.e., ANVIS) is decreased further by a factor of two over a range of spatial frequencies even under optimal ambient levels of illumination. Contrast sensitivity also is decreased considerably with decreasing night sky illumination. The sensitivity loss resulting from decreased ambient illumination is observed across all spatial frequencies; however, this effect is slightly greater for higher spatial frequencies (Rabin, 1993; Wiley and Holly, 1976). This reduction in contrast sensitivity with decreased night sky illumination was found to be a combined effect of lower display luminance and increased electro-optical noise. Rabin (1993) suggested that the development of image intensifiers will improve visual performance by providing greater display luminance and lower noise at starlight and overcast level of illumination. Measures of contrast sensitivity are useful in assessing the potential degradation of visual capability from visual enhancement and visual protection devices used by the Warfighters.

Aging and contrast sensitivity

Contrast sensitivity can become an issue as the Warfighter ages. Contrast sensitivity varies between individuals, reaching maximum at approximately 20 years of age and at spatial frequencies of about 2-5 cpd (Figure 10-16). Aging affects the visual system, which in turn affects the way the visual system and the brain process the collected information. Changes in both the optics and neurons of the eye are the primary causes of reduction of contrast sensitivity with age. With aging, the pupil decreases in size, and the intraocular crystalline lens becomes less transparent. These changes act to reduce the amount of light reaching the retina. Higher-order aberrations also have been associated with age-related cataract development and decreased CSF. Neural changes, such as a reduction of the number of retinal ganglion cells, also can have substantial impact on the observers contrast sensitivity. Accordingly, measures of contrast sensitivity are valuable predictors of the physiological and pathological status of the visual system. In particular, the shape and the height of the CSF can predict if an individual is prone to having difficulties seeing visual targets. Owsley and Sloane (1987) showed that the best predictors of thresholds for real world targets are age and visual function in the middle to low spatial frequencies. Therefore, an understanding of the anatomical and physiological limitations of the visual system is imperative to maximize the contrast required for optimum performance while wearing an HMD.

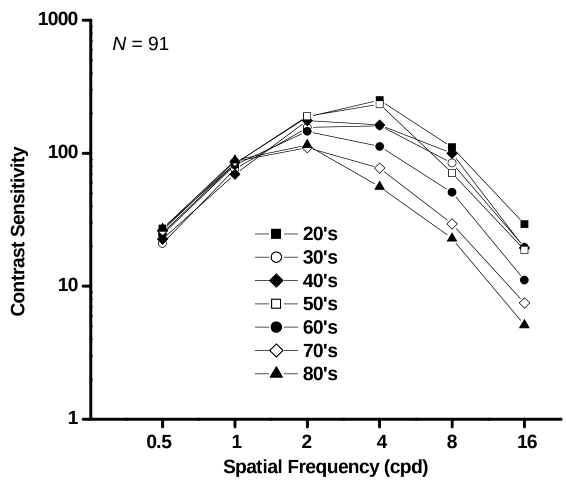


Figure 10-16**.** The contrast sensitivity function (CSF) demonstrates decreased contrast sensitivity as a function of age at middle and high spatial frequencies in cycles per degree (cpd) (adapted from data published [Owsley, 1983] with permission of Dr. Cynthia Owsley).

Effect of refractive surgery on contrast sensitivity

Vision correction by refractive surgery, similar to the use of contact lenses, help to overcome most of the interface problems—e.g., comfort, restricted FOV, lens reflections and glare—usually introduced by spectacles while wearing HMDs. Vision correction by refractive surgery further solves the problems induced by contact lenses wear such as contact lens intolerance, tearing, lens dislodging, lower VA that with spectacles, difficulty of lens hygiene and professional care in the field environment as well as the increased risk for corneal infections (Rash, 2002). Contrast improvement at high spatial frequencies by surgical correction of the optical blur has a positive effect on vision and flight performance under low contrast and low luminance conditions typically encountered in flying conditions. Among the most common surgical procedures undergone by U.S. Army aviators to correct their refractive error are photorefractive keratectomy (PRK) and laser in situ keratomileusis (LASIK). Conventional PRK and LASIK correct first and second lower-order aberrations—such as myopia, hyperopia, and astigmatism. However, they induce higher-order optical aberrations that positively correlate with the amount of myopia correction (Mrochen, 2001). In particular, coma-like aberrations have been shown to influence the contrast sensitivity function. An increase in the aberrations of the eye following refractive surgery also is associated with difficulties with night vision, halos, and glare (Bailey, 2003; Fan-Paul, 2002).

There are conflicting reports regarding the effect of refractive surgery on contrast sensitivity. Some studies have demonstrated that the CSF is compromised by refractive surgery, to include PRK and LASIK, and that increases in higher-order aberrations correlate with deterioration of the CSF. A decline in contrast sensitivity and visual performance under glare conditions after PRK (Dennis, 2004) and reduction on contrast sensitivity across a wide range of spatial frequencies after conventional LASIK have argued against the benefit of conventional refractive surgery to improve optical blur over spectacle correction (Yamane, 2004). Conversely, a more recent study evaluating flight performance of pilots after PRK and LASIK under day as well as unaided and aided night (i.e., NVG) conditions, indicates that there is not a significant baseline performance difference between subjects that underwent these procedures (van de Pol, 2007). In addition, the same study shows there is not significant difference in contrast sensitivity between conventional PRK and LASIK subjects one month after surgery. The advent of wavefront- and topography-guided LASIK that corrects both lower- and higher-order aberrations has resulted in significant improvement in contrast sensitivity and visual performance compared with conventional LASIK (Kaiserman, 2004).

Importance of contrast sensitivity of target detection

Pioneer work by Ginsburg (1983) demonstrated the usefulness of contrast sensitivity as a metric of reduced visual performance—compared to VA—when viewing through aircraft transparencies. This work determined that reduction in the CSF due to HUDs was correlated to diminished target detection ranges. In a subsequent study, Ginsburg and Easterly (1983) demonstrated that pilots with increased contrast sensitivity were capable of acquiring targets further away than less sensitive observers under similar scotopic conditions. The study also showed that increasing the contrast by a factor of only 1.5 to 2 is required for going from chance detection to definite detection. Therefore, while a highly sensitive pilot is able to see the target definitely, a less sensitive one still may be unsure of its presence. These variations in contrast sensitivity and target detection are critically important, as survival in today's combat environment can depend on making split second decisions (Swamy, 2002).

Color Discrimination

Color is a characteristic of display elements often used to encode information. While early display technologies generally were monochromatic (having no variation in hue),[6] multicolor displays have recently become the norm for virtually all display technologies.

Normal color vision and the ability to discriminate between colors is essential to the Warfighter who must identify the colors of targets, smoke, flags, signal and navigation lights, and terrain differences (Tredici and Ivan,

[6] Monochromatic displays should not be interpreted as black and white, as many of these displays were green on black, red on black, yellow on black, etc.

2008). (A thorough discussion of color vision is presented in Chapter 7, *Visual Function*.) All military services, as well as civil aviation agencies, have color vision requirements, but these requirements have been under scrutiny in recent years. Color vision testing generally has relied on the use of pseudoisochromatic plates and, more recently on the Farnsworth Dichotomous test (an aviation standard). However, color contrast and resulting color discrimination capability under real-world conditions can be affected by environmental conditions (e.g., ambient lighting and the presence of fog and haze) and by physiological conditions (e.g., hypoxia and fatigue).

The ability to discern small color differences is easier when the areas to be discriminated are large, contiguous (share an edge near the viewed point), and are viewed simultaneously (National Aeronautics and Space Administration, 2004). As the viewed areas decrease in size or are separated from each other, discrimination becomes more difficult if not impossible. Color discrimination is greatest when a sharp edge separates the colors to be discriminated, e.g., between a symbol and a uniform background color. When a smooth gradient separates two color areas, the smallest detectable difference in color is larger (National Aeronautics and Space Administration, 2004).

Color discrimination and identification is more difficult when the color areas are small and narrow such as would be the situation for symbols and alphanumeric characters used in displays.

The NASA Color Usage Research Lab[7] has provided the following guidelines for the use of color where discrimination and identification are critical:

- *Use no more than six colors to label graphic elements* – How many can be reliably identified depends on several characteristics of the application. In cockpit and automotive applications the user can afford only a glance at the display as part of a rotation among items that must be monitored, and errors can have severe consequences. Fewer and highly distinct colors must be used in this type of application. On planning displays (e.g., maps, scientific visualizations) the user typically has time to more carefully scrutinize elements and refer to a legend. The consequences of errors are less immediate and more likely to be noticed before there are problems. Often more colors can be used in these cases.
- *Use colors in conformity with cultural conventions* – Some hues have become associated with particular meanings through widespread use or tradition. Red, yellow, and green are associated with safety status. Other uses of these colors can lead to unintended interpretations. In applications where only six-to-eight colors are identifiable this severely restricts the options for color coding of non-safety variables.
- *Use color coding consistently across displays and pages* – Users should not be required to associate different meanings with the same hue in various parts of their work environment. Remembering different interpretations in different contexts increases cognitive effort and opens opportunities for error.
- *Use color coding redundantly with other graphic dimensions* – When user populations may include users with anomalous color vision (8-10% of the population), important information must be identifiable on some basis other than color discrimination. Even for individuals with normal color vision, this can be a valuable design goal.
- *Don't use color coding on small graphic elements* – Color discrimination is better for large areas than for small (e.g., small fonts and symbols). This is more of a concern for *at-a-glance* applications than for those where careful examination is possible. Even in the latter it can slow the user down.
- *Use neutral gray surrounds where color judgments are critical* – Simultaneous and successive color contrast can interfere with accurate color identification.

[7] NASA Ames Research Center, Moffett Field, CA.

In military aviation the two longest-fielded HMDs are monochromatic systems: the NVG and the IHADSS. Both present imagery as green on black. Color HMDs have been late in development due mostly to their high cost and weight; color displays also require resolution and luminance tradeoffs. Also, the use of color image sources increases the complexity of the relay optics design, since a polychromatic design must be used. However, these factors have not decreased their desirability to the user. This desirability lies in the fact that color is a very conspicuous attribute of objects. Color can facilitate three functions: Serve as the actual work object, support cognitive functions, and to assist in spatial orientation (Spenkelink and Besuijen, 1996). Overall, color has the potential to reduce workload and improve visual performance.

The color of monochrome cathode-ray-tubes (CRT) and I^2 displays is defined primarily by the choice of phosphor.[8] And, the choice of phosphor is defined primarily by luminous efficiency. Approaches to achieving color in liquid crystal displays (LCDs) are numerous and increasing every day. One approach is similar to the additive color method employed in modern CRT displays. In this approach, pixels are composed of three or more color subpixels. By activating combinations of these subpixels and controlling the transmission through each, a relatively large color gamut can be achieved. The most promising near-term LCD color technology is subtractive-color. Another display technology, Active Matrix Electroluminescent (AMEL), can provide limited or full color, achieved either by classic filtering techniques of color-by-white or by patterned phosphors similar to those used in conventional CRTs. See Chapter 4, *Visual Helmet-Mounted Displays*, for a discussion of the various display technologies.

A number of studies have expounded on the positive impact of color on performance. In one of the more comprehensive studies, DeMars (1975) concluded that, for certain applications, color enhanced accuracy, decision time, and workload capability. However, Davidoff (1991) and Dudfield (1991) found that the actual significance of color far outweighed its perceived importance. An investigation (Spenkelink and Besuijen, 1996) of whether the use of color, and the resulting available chromatic contrast, could help improve performance in the presence of low luminance contrast concluded that only under special conditions was there an additive effect, and, in general, chromatic contrast cannot be substituted for luminance contrast. Rabin (1996) compared Snellen and vernier acuity, contrast sensitivity, peripheral target detection, and flicker detection for simulated green ($x = 0.331$, $y = 0.618$) and orange ($x = 0.531$, $y = 0.468$) phosphors. For central visual tasks, no differences were found. However, peripheral target detection was found to be enhanced for the green phosphor.

Efforts to develop color HMDs date back at least to the 1970s (Post et al., 1994) at which time Hughes Aircraft under the direction of the U.S. Air Force Armstrong Laboratory, Wright-Patterson AFB, Ohio, produced a monocular display around a miniature, 1-inch, P45 CRT which used a rotating filter to provide field-sequential color. Since this effort, a number of other attempts based on multiple image source technologies and methods have been made with only limited success. However, the most promising approach to providing full color in an HMD is based still on field-sequential color, with its potential field breakup problem.[9] Post, Monnier, and Calhoun (1997) have looked at this problem and developed a model for predicting whether this breakup will be visible for a given set of viewing conditions.

It has been suggested that full color HMDs may not be necessary in some applications, and that, through the use of limited color displays, the cost and complexity of color HMDs may be reduced while maintaining the advantages of color. Reinhart and Post (1996) conducted a study looking at the merits and human factors of two-primary color active matrix liquid crystal displays (AMLCDs) in helmet sighting systems. One of their conclusions was that such a design could prove beneficial in an aviation HMD application.

---

[8] A phosphor is a substance that emits light when struck by electrons or ultra-violet energy. Cathode-ray-tubes (CRTs) are a typical example of display devices that use phosphors.

[9] For sequential color displays, when the observer's eyes move rapidly relative to the display, the R,G, and B images will not fall on the same location on the retina. This can result in color breakup, or perceived spatial separation of the R,G,B components (Zhang and Farrell, 2003).

Besides cost, weight, and complexity drawbacks to the implementation of color HMDs, additional issues are present. The luminous efficiency of the eye is a function of wavelength and adaptation state. For example, at photopic levels of illumination, the eye is most efficient at 555 nm, requiring at other wavelengths more energy to perceive the same brightness. Therefore, it is recommended by some researchers that care must be taken in multiple color display designs to ensure isoluminance (Laycock and Chorley, 1980). Also, it has been found that larger size symbols are required to ensure that both detail and color can be perceived when color is selected over black and white (DeMars, 1975).

The monochromatic displays have produced some problems, with chromatic aftereffects reported with I^2 devices. This problem first was raised in the early 1970s (Glick and Moser, 1974). This afterimage phenomenon was reported by U.S. Army aviators using NVG for night flights. It was initially, and incorrectly, called *brown eye syndrome*. The reported visual problem was that aviators experienced only brown and white color vision for a few minutes following NVG flight. Glick and Moser (1974) investigated this report and concluded that the aviator's eyes were adapting to the monochromatic green output of the NVGs. When such adaptation occurs, two phenomena may be experienced. The first is a *positive* afterimage seen when looking at a dark background; this afterimage will be the same color as the adapting color. The second is a *negative* afterimage seen when a lighter background is viewed. In this case, the afterimage will take on the compliment color, which is brown for the NVG green. The final conclusion was that this phenomenon was a normal physiological response and was not a concern. A later investigation (Moffitt, Rogers, and Cicinelli, 1988) looked at the possible confounding which might occur when aviators must view color cockpit displays intermittently during prolonged NVG use. Their findings suggested degraded identification of green and white colors on such displays, requiring increased luminance levels. Another chromatic issue with display imagery and symbology in see-through HMDs is the effects of the real world background color(s) adding to the display color, resulting in an unintended perceived display color (Wood and Howells, 2007).

Havig et al. (2001) raised an issue with see-through color HMDs in aviation (although the issue will also apply to any see-through HMD application), that of symbol colors summing with the outside scene. They argued that, as a result, the colors may not be sufficiently recognizable due to color mixing, i.e., colors on the display will sum with the colors from outside the cockpit. They further argue that the bright ambient light present during daytime viewing could desaturate colors, e.g., pilots would have trouble discriminating between green and yellow.

Attention Capture

The primary goal of an HMD is to make information available to the user essentially at any time, regardless of the orientation of the user's head. In order to achieve this in a see-through system the display information is superimposed optically on the user's FOV. The user looks through the HMD to view the distal world, which is the physical environment in which the user is functioning and in most instances, interacting. If the user is a pilot controlling an aircraft, the distal world is the airspace and/or terrain through which the vehicle is moving. An important issue to clarify is consequences of superimposing the informational display elements of an HMD on the pilot's view of the world. A first step toward this clarification is to differentiate between the optically superimposed image of the HMD symbols and the distal world whose visual image exists independently of the HMD. One helpful distinction is to refer to the visual elements that are on the HMD as the near-domain (ND) and to refer to the visual elements of the distal world that are independent of the HMD as the far-domain (FD). One might also view through the HMD other displays mounted on a nearby instrument panel inside the cockpit or other objects within arm's reach inside the cockpit.

The motivation behind the strategy of optically superimposing the ND information on the FD is to alter the user's visual search and scanning requirements in order to minimize the amount of time the user needs to look away from the FD to look down and acquire information from inside the cockpit. The superposition of the HMD symbols on the FD enables the user to look through the ND in order to see the FD. Thus, the ND and FD are simultaneously available to the user without a head movement or even an eye movement. This does reduce the requirements for visually scanning between the ND instruments and the FD, but likely will incur some costs vs. performance in each domain separately.

Ocular accommodation

To begin assessing the possible costs, consider accommodation, which is the change of focus of the eye's lens (see Chapter 7, *Visual Function*). The changing focus of the eye is accomplished by the balance of the opposing tensions between the eye's ciliary body and the elastic properties of the lens and its capsule. It is well established in the literature that it takes time for the optical power of the lens to change in order to focus between far and near objects; near objects being those closer than twenty feet. Of course, the magnitude of these accommodation changes is age dependent; but even in a person thirty years old or younger, these changes in accommodation can take a substantial amount of time, as much as a quarter of a second. In order to eliminate these time requirements of accommodation, HMDs are designed to ensure that the ND is at essentially the same optical distance as the FD. This optical technique eliminates the time required to change the focus of the eye between ND to the FD. However, even though the eye need not change its focus when shifting between the ND and FD, the shift in attention between them may not be instantaneous.

Attention switching

Simply because the HMD superimposes the ND on the FD, co-locating them in the user's FOV at the same apparent visual depth, does not guarantee that the user is capable of attending to both the ND and FD at the same time. In fact, just as it takes time for the power of the lens to change, it takes time for attention to change, even though objective or physical measurements of these changes in attention are not as straight forward as the measures of optics of the eye. Furthermore, as discussed below, research shows that the shift of attention is important. For the most part, this research has been conducted with HUDs, e.g., display systems that are not attached to the user's head. Nevertheless, since they superimpose the ND on the FD, it is clearly appropriate to extrapolate from the HUD to the HMD (Yeh et al., 2003; Yeh, Wickens and Seagull, 1998).

These issues were addressed systematically as far back as 25 years ago. The findings of one of the early studies are particularly relevant to the present discussion (Fisher, Haines and Price, 1980). Eight subject pilots flew a fixed-based simulator configured to simulate a Boeing 727-type aircraft. These subjects were all highly trained commercial pilots who flew the Boeing 727-type aircraft for one of two commercial airlines, with thousands of hours of experience. Since, at the time of the study these pilots had little or no previous experience with HUDS, they all received a number hours in HUD training. Almost all of the displayed HUD information was presented graphically in a conformal fashion, e.g., the display "… moved in a one-to-one manner with the real world both in pitch and roll, and that certain elements, such as the runway symbol and the horizon line, were designed to overlay their real-world counterparts" (Fisher, Haines and Price, 1980). The HUD provided an extensive suite of symbols that included pitch, heading, altitude, airspeed, glide slope, flight path, speed error, aircraft reference, localizer, as well as flare information. The HUD instrumentation was designed to be sufficient for a zero-zero[10] landing.

While the study evaluated several flight conditions, one condition is most important for the current discussion; it involved landing with a cloud ceiling of 180 feet (55 meters) and a runway visual range of 2000 feet (610 meters). There was light turbulence, but no cross wind; and, a 150- foot (46-meter) decision height was used. Each simulated test flight began at 1500 feet (457 meters) and 8 miles (13 kilometers) from the runway and lasted approximately 4 minutes. The pilots performed the maneuver with and without a HUD. In order to control for experience effects, half the pilots first flew the maneuver with the HUD, and the other half first flew the maneuver

---

[10] *Zero-zero* is an aviation term used to describe no ceiling (altitude of lowest clouds) and no visibility.

without it. A number of flight parameters were recorded, including whether the pilot landed or executed a missed approach. Video and audio recordings were also made of the pilots.

An additional and important point is that each pilot was exposed to a completely unanticipated event, a runway incursion. As the pilot was coming into the runway, another Boeing-727 was presented halfway onto the runway at a 45° angle, as if it was turning from an adjoining taxiway near the runway threshold. This incursion was completely unannounced and unanticipated. Four of the pilots encountered it for the first time with the HUD; the remaining four encountered it without the HUD. The four pilots who encountered this event with the HUD eventually encountered the same event during a subsequent flight that did not involve the HUD; and, those four pilots who encountered the incursion first without the HUD eventually encountered it during a subsequent flight with the HUD. Although the pilots were not warned that runway incursion would occur again, when it occurred the second time, the pilots were probably not nearly as surprised as they were when it occurred the first time. Of interest is how long it took for the pilots to see the incursion, and when the pilot initiated a missed approach.

Since the incursion was a complete surprise to the pilots only the first time it occurs, there was only one first time for each pilot. So the important results of this study, for our purposes, rests on only eight observations, one per pilot, which was far too few for a statistical analysis. Nonetheless, the results are interesting. Of the four pilots encountering the surprise with the HUD, two of them never saw it. They were landing, looking straight at the runway, and the Boeing-727 sitting there, totally undetected. One pilot said, during the debriefing after viewing the tape of the flight; “If I didn’t see it (the tape), I wouldn’t believe it. I honestly didn’t see anything on that runway” (Fisher, Haines and Price, 1980). The other two pilots did see the incursion and initiated the appropriate missed approach; but these pilots reacted several seconds slower than did the pilots without the HUD.

For the second incursion, the pilots were aware of the possibility of unexpected events. However, each of the four pilots without the HUD initiated the appropriate missed approach more quickly than did the four pilots with the HUD (2, 2, 1, 1 vs. 2, 3, 3, 3 sec.).

About fifteen years later, in a study that partially replicated Fisher, Haines and Price, Wickens and Long (1995) found essentially the same pattern of results. They studied thirty-two pilots landing a flight simulator. The subjects were provided conformal or non-conformal flight instrument suite in either a HUD or head-down display (HDD) configuration. During the last flight of each subject, “… a wide-body jetliner taxied into takeoff position on the runway on which the participant was about to land. … the latency between the time the participants broke out of the clouds and the time at which they initiated a go-around…” was the dependent measure. Again, the subjects were not warned about possible runway incursions; so it was a completely unanticipated event. The results were unambiguous: The participants flying with the HDD responded more quickly to the incursion than did those flying with the HUD; about 6 seconds compared to about 8 seconds, a difference that was statistically significant. Furthermore, there was an interaction effect; the delay was significantly longer with the non-conformal HUD (about 9 sec), that with the conformal one (about 7 sec).

These results should not be taken to suggest that HMD or HUDs are bad by any means. These deficits or negative effects seem to be specific for the detection of unexpected events (Yeh et al., 2003). As far as expected events go, even very low frequency events that the user has been prepared to expect, HUDs, and HMDs seem to support performance as good as if not better than conventional HDD displays. However, the superimposed ND of the HMD and HUDS seem to make detecting the truly unexpected event in the FD more problematic.

It seems fairly obvious that cluttering the FD by superimposing the ND on it should make the FD harder to see simply because there are more things to look at. This general effect of clutter means that the user has more things through which to search for the important specific information (Gish and Staplin, 1995). It also means that more things have to be ignored. There seems to be another more specific crowding effect of clutter, that is, items close to each other interfere with their mutual visibility (Ericksen and Ericksen, 1974). This crowding effect may result from crosstalk among retinal neurons, can extend over substantial regions of the visual field (Westheimer, 2004) and can be exacerbated by increasing stress and/or workload (Larish and Wickens, 1991).

The ND is not a just a scattering of random visual elements cluttering the view of the FD, it is a man-made system of regular geometric shapes and alphanumeric characters designed to convey information. The ND is planned and organized to convey information important for the user. When a pilot uses the ND, rather than merely turning it off, and at least to some extent attends to it and the information it provides, it is obviously not being ignored. This observation introduces another factor that maybe more important than the visual clutter. The symbols and icons of the ND interact with the user's attention in a way that is more compelling that if the symbols were random clutter. This particular factor, that HUDs and HMDs seem to capture a user's attention, emerges from the interaction among the symbols, their informational content, and the characteristics of human attention.

In order to address this second issue, *attention capture*, a few introductory words about human attention may seem appropriate. Attention is often likened to a spotlight that can be directed to specific items of interest. Attention is considered to be a limited cognitive resource that can be allocated in specific ways. The HMD literature has described attention being focused, selective, or divided (Prinzel and Risser, 2004). Focused attention refers to the fact that attention seems to illuminate specific elements in the environment, much the same way that vision is directed to specific elements in the environment. The selective nature of attention refers to the fact that attention, again like vision, seems to go from one element to another in a serial fashion, rather than attending to everything all at once. But even though specific items can be selected for special scrutiny, it is also possible to maintain awareness of more than one thing at a time, thereby dividing attention. Furthermore, it seems that visual attention may be allocated to objects as well as to locations in the visual world. In other words, one attends to an object and to some extent, the space around the object, where the object is located. Usually eye movements play a role in this.[11] But if the HUD and HMD are functioning as designed by optically collocating the ND and FD, the need to make eye movements may be reduced or at least minimized. It is even possible that the user may be able to allocate some fractional attention simultaneously between the ND and FD so that an explicit eye movement may not be necessary. But the shifting of attention between the ND and FD may be more effortful without an eye movement than with one. In other words, the absence of an associated eye movement may even make it more difficult for an individual to shift attention.

Ververs and Wickens (1998) have provided a more formal definition of the phenomenon of attention capture as a "… involuntary (and generally undesirable) fixation of mental resources on an information source, for some length of time, at the expense of other elements. This phenomenon is characterized by the inability to effectively switch (sic) cognitive capacities between sources of information. In the aviation domain, a pilot's attention might become locked on a particular instrument resulting in the failure to scan the rest of the environment. When pilots are flying with a HUD where the instrumentation is superimposed on the far domain scene, pilots may fixate on the centrally located near symbology and ignore important information beyond it in the environment."

They point out that attention capture is a misleading term for several reasons. The word capture implies that it is a one time, all or nothing event, like a trapping or locking up of attention. But it need not be; it may be more like a stumbling or stuttering than an actual capture. Furthermore, ascribing the phenomenon to attention is to ignore the fact that many additional cognitive components such as reasoning, remembering, processing, recognition, response strategy selection and preparation may be involved with the phenomenon. Each of these different cognitive functions may be differentially involved depending on the specifics of the situation. For example, some may involve eye movements and a breakdown of instrument scan patterns while others may not involve eye movements at all. Furthermore, as Ververs and Wickens point out, attention capture is a term that was originally used to describe a different phenomenon that may only be tangentially related to 'attention caption' by the HMD/HUD (Jonides and Ynatis, 1988). In general, the abrupt appearance of an object in a visual display has the capacity to draw attention to itself reliably under a wide variety of stimulus conditions. The compelling nature

[11] These eye movements involve the muscles outside the eye that move it to look from place to place and are different from those involved in accommodation, which involve the muscles inside the eye and that control the focusing. [See Chapter 7, *Visual Function*.]

of the transient nature of the stimulus is due to specific processing characteristics of the visual system (Franconeri and Simons, 2005).

Yet the phrase ‘attention capture’ appears intuitively correct since, according to Fisher, Haines and Price (1980), “… several pilots admitted that from time to time they caught themselves totally fixating on the (HUD) symbology, oblivious of anything else, and had to consciously force their attention to the outside scene.” But, both HUD and HMD instruments are designed to be redundant with the FD information. When pilots simultaneously have available both the ND information from the HMD and in the FD, they may simply prefer to use the HMD information. After all, it allows them to control aircraft heading, airspeed, and altitude more precisely than using the FD. Since ND instrumentation provides the pilots with sufficient information, the pilot eventually may become complacent, having little reason to reference the FD. This complacent reliance on the ND contributes to the vulnerability to totally unexpected events. In such situations, it may be reasonable to question how often a pilot does intentionally direct attention from the ND to the FD, and how successful such attempts to shift attention really are. After all, the frequency of such shifts is on a pilot’s own internal schedule that is maintained with no other time-keeping device for self checking. Furthermore, there are such questions as how does the pilot know that the switch of attention from the ND to the FD was successful and is the shift of attention under the pilot’s control.[12] These are purely self monitoring phenomena for which there are no external checks and it has been well established in the ‘attention blindness’ literature that people invariably over estimate their ability to detect changes in their environment. Thus, they are blind to their blindness, which may make them all the more vulnerable (Levin et al., 2000). According to Fisher, Haines and Price (1998), “It is interesting to note that the six pilots who did see the obstacle through the HUD believed (falsely) that they detected it sooner with the HUD than without it. The typical explanation was that ‘The airplane was easier to see with the HUD because I was head-up.”

Foyle, McCann and their colleagues have conducted a series of psychophysical/human performance laboratory studies to examine the ability of individuals to monitor simultaneously the information presented in the ND and in the FD; as well as the time required to shift attention between the two domains (Foyle et al., 1993; McCann et al., 1993; McCann, Foyle and Johnson, 1993; Sanford et al., 1993; Shelden, Foyle and McCann, 1997). In some of these studies individuals also performed a flying-type tracking task that required the individuals to control the heading and altitude of a low-fidelity simulation. Many of these studies used a common overall experimental approach and strategy, with similar equipment, design, and procedures. The ND mimicked the HUD while the FD mimicked the airspace; and both of them were computer-generated graphics presented on an unidentified and unspecified CRT display, presumably a generic desk top unit common at the time.

In a typical study, for example, the HUD image consisted of four small squares; each of which was 1.9 cm (0.75 inch) wide by 1.1 cm wide (0.4 inch). These were arranged in a 2 X 2 pattern, with a horizontal separation of 5.4 cm (2 inches) and a vertical separation of 0.6 cm (0.2 inch). All the HUD information was presented in these four boxes. The HUD also contained a pair of pitch ladders that provided the individual with no task relevant information. The ladders were merely graphical elements whose only purpose seemed to be to define the HUD as a single perceptual object. Other than a passing mention, the pitch ladders were not described in the reports but appeared in the illustration of the stimulus display. Each of the pitch ladders in the illustration consisted of seven horizontal lines arranged in a column that appeared to be about 5 cm (2 inches) high. The two pitch ladders were mirror images of each other, positioned between the boxes, and extending approximately an equal amount above and below the boxes. The HUD was horizontally centered on the CRT, remained stationary throughout each trial, and was blue against the black background.

[12] This question is similar to the one raised in the literature on ocular accommodation, which showed that people are notoriously poor at knowing and controlling where their eyes are focusing. Without something to look at, focus goes to a resting point that is remarkably resistant to volitional control.

The FD mimicked an out-the-window view of a runway outlined from an approach perspective. The runway, like the HUD, was a computer generated graphical image comprised of straight lines. In order to create the illusion of depth on the flat screen of the CRT, the runway icon was a trapezoid. The two horizontal lines, conjuring the near and far ends of the runway, respectively, were 1 cm (0.4 inch) and 23 cm (9 inches) at the start of a trial. These horizontal lines were connected by two oblique lines conjuring the sides of the runway, and a third line down the center of the runway icon to conjure the runway centerline. This runway icon was outlined in yellow against the black background of the screen. There was also a dotted horizon line that seemed to be midway on the CRT, extending its full width.

During a trial, the dimensions of the runway icon changed "… making it appear as if the subject was on final approach. In addition, small vertical and lateral displacements were superimposed on the descent (flight path), simulating changes in the aircraft's pitch and yaw. … It took approximately 5 seconds to make contact with the surface of the runway, considerably longer that subjects typically required making their response (McCann, Foyle and Johnson, 1993)." Consequently, in this particular study the subject was not controlling the simulated aircraft, but merely observed a 5-second long computer animation in which the yellow runway icon of yellow straight lines moved against the stationary HUD icon of blue straight lines, both icons against the common black background.

It is worth pointing out that presenting both the ND and FD at the same optical distance on the CRT ensures that the subjects do not need to change to accommodate when sifting vision between the ND and FD, thus eliminating accommodation as a potentially confounding variable.

The task of the individual participating in the experiment was to press one of two keys on a keyboard, selecting one or the other depending on information presented during the trial. The individual's response accuracy and reaction time were recorded. The specific experimental manipulations of this study were the patterns of stimuli presented in the HUD and runway icons. There were three types of stimuli: one type was a cueing stimulus, the second was a discriminative stimulus and the third was a distracting stimulus.

The cueing stimulus could be either the alphanumeric group for visual flight rules (VFR) or the group for instrument flight rules (IFR). At the start of a trial, one of these cues was presented in one of the two lower HUD boxes or on the runway just below but proximal to these lower pair of HUD boxes. The cueing stimulus told the individual whether the next stimulus, which was the discriminative one and which was presented 125 milliseconds (ms) after the cue, would be presented in the HUD or in the runway icon. IFR meant that the discriminative stimulus would be presented on the HUD whereas VFR means that the discriminative stimulus would be presented on the runway. Consequently, if the cue was IFR and appeared on the HUD, then the discriminative stimulus would also appear on the HUD and the subject would not have to shift attention from the HUD to the runway in order to respond to the discriminative stimulus. Similarly, if the cue was VFR and appeared on the runway, then the discriminative stimulus would also appear on the runway and the subject would not have to shift attention from the runway to the HUD in order to respond to the discriminative stimulus. In these two situations, the cue and discriminative stimuli were both presented in the same domains, either in the ND or in the FD. Conversely, if the cue was IFR and appeared on the runway, then the discriminative stimulus would appear on the HUD and the subject would have to shift attention from the runway to the HUD in order to respond to the discriminative stimulus. Similarly, if the cue was VFR and appeared on the HUD, then the discriminative stimulus would appear on the runway and the subject would have to shift attention from the HUD to the runway in order to respond to the discriminative stimulus. In these two situations, the cue and discriminative stimuli were presented in different domains, and the subject had to shift attention between the ND and FD.

The discriminative stimulus was either a stop sign or a diamond and the subject pressed one or the other key depending on whether the stop sign or diamond was the discriminative stimulus. The subjects were told that the stop sign meant that the runway was closed and that the key press initiated a missed approach, whereas the diamond meant that the runway was open and the key press signaled the continuation of the landing. The discriminative stimulus was presented on the HUD or on the runway, in a location unoccupied by the cue.

Simultaneous with the onset of the discriminative stimulus (250 ms after the cue onset) distracting stimuli were presented in the remaining unoccupied boxes on the HUD and the unoccupied locations on the runway. These distracting stimuli were squares and triangles.

The results of this study showed unequivocally that it took longer to shift attention between the HUD and runway than when the cue and discriminative stimuli were both in the HUD or both in the runway. Subsequent experiments suggested that the difference in shifting attention between the HUD and runway depended upon the extent to which these two graphically created icons were distinguished as separate perceptual objects. For example, one of the differences between the HUD and runway was that the runway appeared to move whereas the HUD was stationary. When the study was conducted with a runway that did not appear to be moving, then the difference in shifting attention between the ND and FD was reduced; however, the results contained an important hint. There was little difference in reaction time when both the cue and the discriminative stimuli were both on the (stationary – nonmoving) runway or when the cue was on the runway and the discriminative stimulus was on the HUD. In other words, the subject could just as easily shift attention within the runway or from the runway to the HUD. But; sifting attention from the HUD to the (stationary) runway, took significantly longer than shirting attention within the HUD. Somehow, the HUD icon still seemed to hold attention more strongly than did the runway iconography.

Subsequent elaborations of the basic experimental paradigm required the subjects to fly the low-fidelity simulator. The performance measures were the accuracy (root mean square error) with which the subjects were able to hold assigned altitudes and headings. The experiments manipulated the configurations of the HUD and out-the-window, i.e., the ND and FD views, to identify further the characteristics of attending to these two domains either simultaneously or in succession. The results of these studies agreed with the previous findings. The display of information in the ND interfered with the components of flight performance that were dependent on information from the FD. But, most important, the extent to which the ND affected the subjects' ability to attend to the FD, depended critically on the configuration of the ND. These results suggested to Foyle and his colleagues a strategy that promised to mitigate the perceptual tunneling effects of the HUD, and by extension, the HMD.

This strategy is sometimes referred to as scene-linking and at its core is the notion of reducing as much as possible the perceptual differences between the ND and FD. The ND display components are designed to appear to be part of the FD. For example, the differential motion between the FD and the components of the ND is reduced. The ND components should move with the FD. Sheldon et al. (1997) identified several forms of potentially scene-linking ND symbols "*Scene enhancements* are the graphical outlines of existing objects in the external world, such as a graphic runway that overlays an actual runway, or a virtual horizon. *Scene augmentations* are the addition of virtual, three-dimensional (3-D) objects that are otherwise non-existent in the real worlds, such as 'virtual traffic lights' that may operation on taxiways to separate aircraft. *Virtual instruments* are the depiction of *ownship* flight instrumentation and data such as a glideslope readout on 'virtual billboards' that appear to the side of the aim point of a cleared runway at landing (Sheldon et al., 1997)."

Researchers have realized some of these ideas in the Taxiway-Navigation and Situation Awareness (T-NASA) Cockpit Displays, a system that integrates information from the Differential Global Positioning Satellite system (DGPS), surface radar, and data line to provide graphically on the HUD final approach and cleared taxi route information augmented with a moving map display (Hooey et al., 2000). The T-NASA is one of several cockpit display systems designed to overcome the limitations of the conventional steam gauge-type instruments of the head down instrument panel while meeting the challenges of the HUD and HMD.

Motion Perception

The physical world comprises an ongoing series of spatio-temporal events. The human visual system is sensitive to a limited range of these events. Spatially, some of those events take place among elements that are too small to be resolved by the human visual system (e.g., atoms), and some are too large to be encompassed within the FOV (e.g., galaxies). Temporally, some take place so quickly that they escape our notice (e.g., the flight of a bullet), and some, so slowly, that they appear static during a given observation interval (e.g., a plant growing). Real-world events, which involve continuous changes of position over time and which fall within certain spatio-temporal boundaries, give rise to perceptual experiences that are called *real* motion (Goldstein, 2007). Real motion percepts belong to the more general class of *visual motion* percepts that include *apparent* motion (Anstis, 1978), *induced* motion or motion contrast (Nawrot and Sekuler, 1990), and motion *aftereffects* (MSEs) (Mather, Verstraten and Anstis, 1998). The additional classes of motion percepts are demonstrations that continuous motion is not necessary for the experience of visual motion.

The goal of this section is to describe the basic phenomena of visual motion and their underlying mechanisms, with limited references to (implications for) the design of visual displays. After an initial review of the spatio-temporal characteristics of the overall visual system, the section proceeds sequentially from the most basic building block of visual motion, the directionally selective cell (modeled as a local, first-order, motion-energy detector for luminance-defined inputs), to more complex processing of motion events (various forms of apparent motion, induced motion, MAEs, temporal motion priming, structure-from-motion, biological motion, optic flow, ego motion) mediated by the spatial and temporal integration of local motion signals and their inputs to higher motion processing stages. The section on visual motion with luminance-defined inputs is followed by a major discussion of the variety of non-luminance stimulus dimensions that support motion percepts, along with the second-order motion mechanisms that underlie them. The section is written at a higher level than an introductory text, so it presumes some familiarity with concepts like the retina, receptive fields, psychophysics, frequency analysis, and filter concepts. The visual phenomena and mechanisms included in the section are chosen primarily for their ability to contribute to an organized understanding of visual motion in general and only secondarily for their contributions to display/HMD design. Certainly, not all display/HMD implications are discussed explicitly (nor could they be in limited space), but a few are included in the text in the appropriate locations (e.g., refresh rates for displays, breaking of camouflage by motion, ego motion, and input saliency). The section is *not* intended to be a comprehensive review of existing applied research on display/HMD design based upon vision/cognition principles. Moreover, the section does not address issues related to optic flow and ego motion when they involve the processing of non-visual motion information. Its scope would have to be expanded significantly to include cues from other sensory systems (e.g., tactual, proprioceptive, vestibular) and even elements of cognitive interpretation of multi-modality information. The limitations engendered by not addressing non-visual cues in motion perception is illustrated by a study (Schulte-Pelkum, Riecke and von der Hyde, 2003) that obtained differences in the degree of ego motion (perception of self motion) generated by a visual stimulus displayed on a projection screen and on an HMD. Ego motion was significantly less with the HMD, with which, when compared to a projection screen, an observer is in tactual contact and which moves when an observer moves. Non-visual cues on motion perception notwithstanding, the analysis of the relationship between visual information and motion perception and the mechanisms underlying the relationship cannot be over-estimated.

Historically, one could easily argue that the modern, scientific approach to understanding visual motion began with the study of *apparent* motion. If two stationary stimuli are presented a short distance apart in rapid succession, humans report an *apparent* motion of a single stimulus between the two positions of the stimuli, even though no physical motion actually occurs between them. Perhaps because the discrete display can be considered the minimum for specifying motion physically, it has been exploited as one way to analyze and characterize visual motion sensitivity in general (Anstis, 1978). Exner (1875) used electrical sparks as stimuli and found that, when two sparks were too close to be resolved spatially, they could nonetheless give rise to a perception of motion when presented sequentially. Exner concluded that apparent motion could not be inferred from a change of position over time, but must be a primary perception on its own.

Spurred on by Exner’s observations, other researchers have pursued apparent motion for both theoretical and practical reasons. For the Gestalt psychologist Wertheimer (1912; cited in Palmer, 1999), apparent motion constituted an example of an emergent property whose nature was explored by varying the timing of and spacing between discretely displayed elements. Korte (1915; cited in Palmer, 1999) extended the analysis further and developed a set of descriptive laws relating the perception of motion to three parameters of apparent motion displays: stimulus timing, spacing and intensity. One major limitation of the early studies was their reliance on subjective reports of the presence or absence of apparent motion, or reports of its quality. A second limitation was the implicit assumption that the empirical relationships they discovered were a description of one, more or less, homogeneous motion system. With more advanced psychophysical techniques, recent studies of apparent motion have provided results which contribute substantially to theories of multiple motion processing mechanisms (more detail below) and to data valuable for the practical design of imaging systems.

Spatio-temporal range of the overall visual motion processing system

Happ and Pantle (1987) used *d'* (Green and Swets, 1966) as an objective measure of directional motion and required observers to discriminate the temporal order of onset (stimulus onset asynchrony [SOA]) of two side-by-side light-emitting diodes. They found that *d'* was an approximately linear and increasing function of log(SOA). For foveal vision, the SOA's were smallest for spatially abutting diodes (0° separation), and sensitivity to differences of onset order as small as 1.6 msec were discriminable at above-chance levels. For peripheral vision, SOA's were smallest with a spatial separation of approximately 1° of visual angle. Again, SOA's in the neighborhood of 1 to 2 msec were sufficient for directional judgments at above-chance levels. The results demonstrate that directional judgments are possible at presentation rates that are an order of magnitude faster than refresh rates commonly used for television and computer displays (~16 msec).

Other researchers have used frequency analysis to characterize the overall spatio-temporal performance of the visual system with luminance-defined stimuli. Using flickering gratings produced by spatial and temporal modulations of the luminance of a display, Robson (1966) and van Nes et al. (1967) measured the minimum contrast required for an observer to detect a grating as a function of its spatial and temporal frequencies. In both studies an interaction between spatial and temporal frequency was obtained. Spatial (temporal) contrast sensitivity behaved like a low-pass filter for high temporal (spatial) frequencies, but like a band-pass filter at low temporal (spatial) frequencies. More importantly here, it was demonstrated that the high spatial and temporal frequency cutoffs of the contrast sensitivity functions were relatively independent of one another. The cutoffs describe frequency limits above which contrast variations are not visible, no matter how high their contrast.

It is possible to use the high spatio-temporal frequency cutoffs to construct a window of visibility for contrast variations (Watson, Ahumada and Farrell, 1986), with spatial frequency along one (vertical) side of the rectangular window and temporal frequency along the other (horizontal) side. Visible spatial and temporal frequencies of luminance modulation would be represented by points within the window, with invisible ones falling outside. With such a window, it would be predicted that the perception of a time-sampled display of a continuously moving stimulus would not be changed as long as the sampling introduced frequency components that lie only outside the window of visibility. Measurements of the ability of human observers to discriminate between apparent (time-sampled) and real motion displays confirmed predictions derived from the window of visibility. In general, critical temporal sampling frequencies below which apparent and real motion appeared identical increased with stimulus velocity as predicted. Temporal sampling frequencies in the 200 to 300 Hz range were required for a line moving at 15°/sec to appear identical to a continuously moving line. These results, like those on temporal order judgments, indicate that modern display devices with refresh rates of 60 to 120 Hz may act as temporal filters of environmental information potentially useful to human observers.

Motion processing with luminance-defined stimuli: First-order motion mechanisms

Part of the basis for concluding that visual motion is a primary sensation in its own right can be found in the directionally selective (DS) mechanisms of the visual system. DS elements compare the changing distributions of luminance within local neighboring regions of the retina. Their ability to respond selectively to direction of motion derives from two anti-symmetric inputs (from sub-units) with different time courses. Direction-sensitive neurons have been found in many species, and their operation has been described extensively for the fly (Reichardt, 1961), the rabbit retina (Barlow and Hill, 1963), and the visual cortex of cats and monkeys (Hubel and Wiesel, 1962, 1968; Rodman and Albright, 1987). The existence of DS mechanisms in humans was first demonstrated by Sekuler and Ganz (1963) in psychophysical experiments. After prolonged adaptation to a grating moving in one direction, the threshold contrast required to detect a grating moving in the same direction was higher than that for a grating moving in the opposite direction.

Besides the direction-specific threshold elevations found by Sekuler and Ganz (1963), other psychophysical results have been interpreted as support for the existence of DS elements which are selectively sensitive to the direction of motion of luminance-defined stimuli. The contrast threshold for a sine-wave grating moving in one direction is not changed when it is superimposed upon a sine-wave grating moving in the opposite direction (Sekuler, Pantle and Levinson, 1978). When added together physically, the contrasts of the two gratings do not sum visually to make the result (a flickering counter-phase grating) any more visible than either directional component viewed alone.

After fixating a pattern moving in a uniform direction for a period of time, a stationary pattern will appear to move in the opposite direction, the so-called MAEs. According to Sekuler and Pantle (1967), the moving pattern is hypothesized to selectively adapt DS elements for one direction of motion and leave elements sensitive to the opposite direction unaffected. The resulting imbalance provides a signal for the stationary pattern to move in the opposite direction. Because the population of DS elements is assumed to comprise units with different spatio-temporal response characteristics, adaptation to a moving pattern would be predicted to be velocity-specific, as well as direction-specific. It is not surprising then that adaptation to a moving grating has been found to elevate the contrast threshold for a test grating moving at a similar velocity, but not those for test gratings moving appreciably slower or faster (Pantle and Sekuler, 1968).

Computational models, based upon the physiological properties of DS neurons, have been developed by a number of researchers (Adelson and Bergen, 1985; Marr and Ullman, 1981; van Santen and Sperling, 1984, 1985; Watson and Ahumada, 1985) to simulate local motion detectors in humans. While the algorithms employed in the different models differ in detail, in each case the inputs to a DS unit are modeled with a pair of sub-units (filters) with spatial weighting functions (receptive fields) in an approximate quadrature phase. In addition, the inputs of the sub-units to the DS element are temporally offset or filtered to produce appropriate time courses of action on the DS element. An array of DS units with different spatio-temporal characteristics is assumed to service each local region of the retina and to produce a crude, local Fourier analysis of a given input stimulus. As a class, the models are called motion-energy models, and the spatio-temporal luminance distribution in a local region of the retina is defined as their input. For this reason, they are also said to generate first-order motion signals in contrast to motion mechanisms (second-order) which take contrast, texture, depth or motion differences as their input [presented in more detail later (Smith, 1994)]. The hypothesis which links outputs of the motion-energy class of models with the perception of motion by human observers is the *motion-from-Fourier-components principle* (Chubb and Sperling, 1988). The motion percept elicited by a complex stimulus will be in the direction of the spatio-temporal frequency components with the greatest expected power. If the expected power in any one direction is matched by the expected power in the opposite direction, the stimulus is said to be *drift-balanced*, and no motion will be perceived. The first-order motion models have been used to explain, simulate and predict the results of human psychophysical experiments with simple and complex luminance-defined stimulus patterns. A few empirical results obtained with specially constructed stimuli demonstrate the usefulness of the motion-energy model.

Observers report that a square-wave grating which jumps ¼-cycle to the right will appear to move rightward. However, the same grating with its fundamental spatial frequency (Fourier) component removed will appear to move leftward (Adelson and Bergen, 1985). This result is explained by the motion-energy model in the following way. A square-wave grating is made up of a fundamental sine-wave component along with odd harmonics of the fundamental whose amplitude decreases in proportion to their frequency. For a square-wave grating, the fundamental and every other spatial frequency component (1f, 5f, 9f, etc.) shift ¼-cycle to the right with each jump and contain more average rightward power than the average leftward power of the remaining components (3f, 7f, 11f, etc.) which shift ¾-cycle to the right (¼-cycle to the left) with each jump. For a missing fundamental grating, there is more average leftward power than rightward power. For each rightward shifting component (5f, 9f, 13f, etc.) there is a leftward shifting component with greater power (3f, 7f, 11f, etc.).

If two identical pictures are presented sequentially in overlapping but slightly displaced positions, motion will be perceived in the direction of the physical displacement as expected in normal apparent motion. If, however, the second picture is a contrast-reversed (negative) version of the first picture, then surprisingly motion will be perceived in a direction opposite the physical displacement (Adelson and Bergen, 1985; Anstis, 1970; Anstis and Rogers, 1975). The reversal of apparent motion is consistent with the motion-from-Fourier-components principle of the motion-energy model. The control exercised by a number of other variables on forward and reversed motion in two-frame, apparent motion displays are simulated with computational models based upon motion-energy detectors (Pantle and Turano, 1992; Strout, Pantle and Mills, 1994).

Lastly, consider a compound stimulus which results from the linear superposition of a drifting sine-wave grating (motion stimulus) and a stationary sine-wave grating of the same spatial frequency (called a pedestal) (van Santen and Sperling, 1984). The compound stimulus contains luminance peaks which merely oscillate back and forth and do not provide any non-equivocal information about direction of motion to a system designed to track features. On the one hand then, it is somewhat surprising that human observers' reports are not only directional, but also virtually identical when the moving sine-wave grating is shown alone or superimposed on the stationary pedestal. On the other hand, a first-order motion-energy system possesses the property of *pseudo-linearity* whereby its response to the compound stimulus is simply the sum of its responses to the individual sine-wave components. As a corollary, the addition of the stationary pedestal grating with a temporal frequency of zero would produce a zero output from a motion-energy system and would not disturb its non-zero response to the moving component grating.

Comparisons of the putative motion-energy detectors in humans with their physiological correlates in other mammals and primates makes it likely that they are located at early stages of visual processing (V1 of the striate cortex) (Emerson, Bergen and Adelson, 1992; Movshon and Newsome, 1996). Hypothetical interactions between the motion-energy detectors and further processing of their outputs by higher-level mechanisms have been offered as the basis of other visual motion phenomena (Simoncelli and Heeger, 1998). A few examples are described in more detail here -- motion priming, structure-from-motion, motion contrast and assimilation, biological motion, and self-motion.

The perceived motion of a vertical sine-wave grating which undergoes an abrupt 180°-phase shift (motion step) is ambiguous. The grating sometimes appears to move rightward; sometimes, leftward. In a system of motion-energy detectors the output of rightward and leftward detectors would be expected to be balanced, but in any one instance "internal noise" would favor one or the other direction. When the ambiguous, 180°-step follows closely upon an unambiguous step (e.g., 90°) which would activate only motion-energy detectors for one direction, the perceived direction of the ambiguous step is biased in the direction of the unambiguous step (Pinkus and Pantle, 1997). The bias is termed visual motion priming and lasts approximately a second. The biasing can be explained by a persistence of the directional response of motion-energy detectors to the priming motion and its temporal integration with the balanced response of motion-energy detectors to a 180°-step. Variations on the priming paradigm support the temporal integration explanation. Visual motion priming demonstrates the benefits of multi-frame representations of directional motion over the minimum two-frame representation (Snowden and Braddick, 1989).

Perhaps the simplest example of spatial interactions generated by local motion-energy detectors is the formation of a *structure-from-motion* with random-dot kinematograms (RDKs). RDKs are motion displays typically consisting of two frames of random black and white dots presented in alternation. In one version of an

RDK, one rectangular region in both frames contained identical elements and is shifted slightly from one frame to the next. The remaining portion of each frame contains independently generated black and white dots; they are therefore uncorrelated across frames. Viewed alone each frame looks only like a pattern of random dots. When animated however, the coherently displaced subset of random dots emerges as an organized structure, a rectangular figure against a noisy background (Braddick, 1974). A simple pooling of local motion-energy signals generated by the coherent global displacement of the dots in the rectangle in the absence of any consistent directional signal in the surrounding area could be the physiological process underlying the perceived structure. If indeed local motion signals are necessary for the emergence of the perceived structure, then spatial displacements of the rectangular region which are large and fail to activate the local motion-energy detectors will cause the motion-generated structure to disappear. Similarly, if the time between the two frames is made too long, no structure-from-motion will be seen. Initially, the spatial limit (D_{max}) obtained by experimentation was approximately 1/4° of visual angle, and the temporal limit (T_{max}) was approximately 80 ms (Braddick, 1974). The underlying substrate responsible for the emergence of the structure-from-motion was termed the short-range process by Braddick (1974). Since Braddick's early research, new experiments (for a review, see McKee and Watamaniuk, 1994) have found that D_{max} and T_{max} are not absolute limits, but can vary with stimulus conditions and stimulus filtering. In the real world, structure-from-motion mediated by first-order motion detectors is one of the most potent factors in the breaking of camouflage and the attraction of visual attention to an otherwise hidden object.

Spatial interactions among first-order motion signals have been shown to be more complex than excitatory summation or facilitatory pooling across common motions in time or space. Nawrot and Sekuler (1990) used RDKs in which dots in alternating (spatial) strips tended to move uniformly in one direction or in random directions (dynamic noise). When the alternating strips were narrow, the strips with uniform motion induced a common motion in the noise strips (motion assimilation); when they were wide, the strips with uniform motion induced a motion of the opposite direction in the noise strips (motion contrast). Motion contrast has been explained by inhibitory interactions between motion-energy units (Murakami and Shimojo, 1996). The motion-energy units activated by a directional stimulus are assumed to upset the balance (a zero net response) of motion-energy units that response equally (or not at all) to a stationary stimulus.

Even more complex are the point-light displays which give rise to biological motion. Johansson (1973) filmed an actor in the dark with small lights attached to his joints (shoulders, elbows, wrists, hips, knees and ankles) so that nothing was visible except the lights. When the actor was stationary, observers perceived only a meaningless pattern of lights. When the actor moved, observers reported that they saw a person moving within fractions of a second. The biological motion percept requires the integration of signals for motions in different directions and velocities. Like the motion of a single point, biological motion appears to be a primary sensation in its own right, and single neurons have been found in higher stages of the visual system (superior temporal sulcus, STS) which respond selectively to biological motion (Oram and Perrett, 1994).

The instantaneous motion of elements (optic flow pattern) portrayed on the retina of an observer as (s)he moves about can be represented by a vector field. For example, when a person moves forward toward an object, the vector field would consist of vectors of different directions and lengths pointing outward (optical expansion); when moving backward, a pattern pointing inward (optical contraction). It has been suggested that a mechanism which combines the motion vectors would provide information about the direction in which an observer is headed (Blake and Sekuler, 2006). Regan and Beverley (1978) have shown that it is possible to selectively adapt the human visual system to optical expansion and contraction providing evidence for the existence of cells which explicitly encode expansion and contraction patterns. The existence of such cells has been confirmed in studies of single neurons of area of the medial superior temporal area pars dorsalis (MSTd) in the primate cortex by Tanaka and Saito (1989). Interestingly, those MSTd cells have extremely large receptive fields, likely a reflection of each neuron's input from many motion-energy detectors at earlier stages of the primate visual system. When patterns of optical expansion and contraction are displayed in a virtual environment, an observer experiences self-motion even though they are stationary.

In conclusion, luminance-defined stimuli are thought to generate elementary, low-level, motion signals in so-called first-order, motion-energy detectors. The elementary sensations are elaborated into more complex motion experiences through the interaction and combination of the elementary signals at later stages in the visual system. Both elementary and some complex motion experiences appear to be primary sensations in their own right.

Motion processing with non-luminance defined stimuli: Second-order motion mechanisms

The spatial (D_{max}) and temporal limits (T_{max}) for the perception of motion in RDKs (discussed earlier) are markedly shorter than what has been found with classical studies of apparent motion (large objects on a uniform background). Assuming that the RDK limits are properties of an early-stage, low-level system of motion-energy units, some other system was assumed to be responsible for the apparent motion in the classical studies. This second mechanism was called the long-range motion system by Braddick (1974), but see also Petersik (1989) and Cavanagh and Mather (1989) for further viewpoints on the nature of the short- and long-range motion systems.

Visual bistable figures are stimuli that produce perceptions which oscillate over time. One classical static example is the Necker cube. The element-group movement display is another example of a dynamic bistable stimulus (Pantle and Picciano, 1976). The motion display contains two frames with three equally spaced dots (elements) in each frame on a homogeneous background. The dots in one frame are displaced back and forth between frames by the distance between the dots, such that the center and rightmost dots in one frame overlap the leftmost and center dots of the second frame. When the time between frames is of the order of 10's of milliseconds, the animation is bistable. Observers alternately report a perception in which all three dots appear to shift together by the same amount (group motion) and a perception in which the overlapping dots remain stationary and the remaining dot appears to flicker or shift from one end of the display to the other (element motion). Attneave (1971) explained bistable phenomena in general by proposing that they were analogous to an astable multi-vibrator electronic circuit which alternated between two states and was the result of two interacting semiconductors. Borrowing upon the multi-vibrator model, Pantle and Picciano (1976) explained element-group movement bistability in terms of two competing motion mechanisms. Further research (Petersik and Pantle, 1979) demonstrated that one or the other of the competing movement perceptions could be favored by the manipulation of stimulus conditions. However, those conditions which favored group movement were not like those of first-order, motion-energy detectors.

As it became clear that not all motion percepts were mediated directly by first-order motion-energy detectors, researchers sought to specifically develop displays which would elicit motion percepts, but which were not based upon luminance-defined stimuli. Pantle (1973) reported that human observers experienced apparent motion with a stimulus not defined by luminance. Each frame of a two-frame apparent motion sequence contained a rectangular area with randomly positioned line segments, all with the same orientation, on a background of randomly positioned line segments whose orientation differed from that in the rectangular area by 90°. The position of the rectangular area was shifted laterally across frames. When the two frames were temporally alternated, observers saw the line segments in the rectangular area move back and forth as a group across the line segments in the background (texture motion). The global movement of the rectangular group of elements was seen despite the fact that the rectangular area was not defined by luminance; the rectangular area had the same average luminance as that of the background. What is most significant about this finding is the fact that the perceived texture motion could not have been mediated by first-order motion-energy units which require luminance-defined inputs. Besides orientation differences, other non-luminance differences have been studied extensively to determine whether or not they have the ability to define stimuli (second-order stimuli) which support motion percepts. The goal of the research has been (1) to investigate the variety of non-luminance defined stimuli that support motion processing, (2) to determine what type of non-linear transformations of stimulus luminance might make a second-order stimulus amenable to motion-energy computations, and (3) to study and compare the response characteristics of first- and second-order motion processing.

The variety of non-luminance defined stimuli which support motion perception is large. They include both non-periodic and periodic stimuli. An amplitude-modulated (contrast-modulated) grating is one whose spatial contrast varies periodically across the pattern. It is the product of a high spatial frequency sine-wave (carrier) and a lower spatial frequency modulating waveform. If the modulating waveform is itself a sine wave, then the resulting complex wave can be analyzed as the sum of a fundamental frequency and two sideband frequencies. If the modulating waveform moves, and the carrier is stationary, the two sideband components move in opposite directions. First-order motion-energy detectors would signal no motion and would not support a motion percept because the net directional energy would be zero according to the motion-from-Fourier components principle. Yet, human observers do see the motion of the contrast variations of the amplitude-modulated grating (Pantle and Turano, 1992). The motion would be revealed to motion-energy detectors, if a point-wise transformation like rectification were first applied to the grating stimulus. The second-order contrast variations would be transformed to intensity variations which would be visible by motion-energy detectors.

A slightly more complicated stimulus transformation prior to motion processing could reveal the motion of the orientation-defined figure in the example described earlier. The application of a spatially oriented filter followed by the application of a grossly non-linear point-wise transform would produce an intensity-defined output capable of activating motion-energy detectors. Even more stringent principles can be followed to guarantee more strongly that the motion of any second-order stimulus is not due to activation of first-order motion detectors. Chubb and Sperling (1988) created second-order stimuli, which they defined as drift-balanced. The expected energy of any Fourier component of a drift-balanced stimulus is equal to the expected energy of the component of the same spatial frequency drifting at the same rate in the opposite direction. Following this maxim guarantees that the response of all first-order motion detectors, no matter what their spatio-temporal frequency tuning, would be balanced for opposite directions of motion, not just the response expected across all detectors as a group. One example of a drift-balanced stimulus is a flicker grating, which is the result of the modulation of the flicker frequency of spatial noise (a random array of black and white pixels) with a drifting sinusoid. The motion of the flicker-defined grating is invisible to first-order motion-energy detectors, but nonetheless observers perceive its motion. The motion can be revealed by second-order motion-energy computations applied to the rectified output from an earlier temporal filtering stage.

In conclusion, it is clear on the one hand, that human motion perception is not mediated solely by first-order motion-energy detectors which operate directly on the raw spatio-temporal luminance distribution of an image, as is demonstrated by the sheer number and variety of non-luminance defined stimuli which induce some motion percepts. On the other hand, computational findings demonstrate that motion-energy detectors are capable of signaling motion with second-order stimuli provided only that the stimuli are first subjected to suitable filtering followed by a non-linear transformation. Moreover, more analytical experiments with specially constructed second-order stimuli show that visual phenomena analogous to reverse motion and pedestal immunity which are signatures of first-order, motion-energy processing also obtain for second-order motion processing (Chubb and Sperling, 1988). Findings of the immobility of second-order motion in the periphery notwithstanding (McCarthy, Pantle and Pinkus, 1994; Pantle, A., 1992), properties of first- and second-order motion processing have been found to be remarkably similar (Lu and Sperling, 2001). Despite the demonstrated explanatory power of motion-energy computations for first- and second-order stimuli, there are some remaining visual motion phenomena which cannot be explained by such mechanisms. For example, animated apparent motion sequences in which frames are alternately presented to the right and left eye are capable of creating vivid impressions of motion, yet it is known that motion-energy computations are strictly monocular. Interocular motion provides hints of a third human visual motion system (Lu and Sperling, 2001). The search for physiological substrates of motion processing, no matter what the final outcome of psychophysical research and computational modeling, shows that motion processing takes place in channels or pathways that are segregated from form (object) processing.

Motion processing: Physiological substrates

It is generally accepted that the primate visual system comprises two partially independent, parallel pathways defined by the input attributes (dimensions) which they are optimized to analyze. The division is based upon physiological research on primates, and neurological and psychophysical studies on humans (Lennie, 1980; Livingstone and Hubel, 1988; Merigan and Maunsell, 1993). Alternative names (what/where, dorsal/ventral streams) have been used to refer to the two pathways (subsystems), but here, we will follow the lead of those who have named them the parvocellular (P) and magnocellular (M) pathways, based on the dichotomy of the cell body sizes predominant in each system. The P-pathway extends from P-cells in the retina to structures in the temporal lobe; the M-pathway, from M-cells in the retina to structures in the parietal lobe [MT (V5) and MST]. Single-cell recording of P- and M-cell activity show that P-cells code color differences whereas M-cells do not. P-cells have a greater spatial acuity (higher spatial frequency cutoff) than M-cells. P-cells respond less well to temporal fluctuations of stimulus intensity (have a lower temporal frequency cutoff) than M-cells. Finally, transmission of signals is slower in P-cells than in M-cells. Given the functional differences between P- and M-cells, it is not surprising that lesions in the P-pathway produce deficits in color vision, texture/form perception, and spatial acuity, whereas lesions in the M-pathway produce deficits in flicker and motion perception (Merigan and Maunsell, 1993). The difference of behavioral functions ascribed to the P- and M-pathways can be exploited in display/HMD design. On the one hand, for a dynamic display primarily intended to portray motion, there would be no advantage to color coding or maximizing spatial resolution. Fast refresh rates as outlined earlier in the section would be desirable. On the other hand, for a static display primarily intended for detailed object recognition, fast refresh rates would be superfluous, whereas color coding and high spatial resolution would be beneficial.

More detailed analyses of the MT-pathway with lesions, single-cell recordings, cell micro-stimulation, and functional magnetic resonance imaging *(*fMRIs) have provided data that demonstrate even more strongly the connection between the M-pathway and the results of psychophysical and computational studies of visual motion. They also show a correlation between M-pathway response characteristics and saliency/eye fixations. A number of researchers have noted the similarities between motion-energy detectors in computational models used to explain first-order motion phenomena and single DS cells in cortical V1. Emerson, Bergen and Adelson (1992) made extensive measurements of 1- and 2-bar test responses of DS complex cells of V1 in the cat. The single-bar responses and 2-bar interactions yield highly distinctive patterns, and they matched the predicted responses of first-order, motion-energy detectors quite well.

The input of DS cells to MT (V5) and MSTd single cells at higher stages in the M-pathway allows for the combination of the outputs of first-order motion-energy detectors needed to explain various grouping phenomena observed behaviorally in monkeys and humans. One particularly useful stimulus contains a set of randomly positioned dots, a fraction of which are made to move in a common direction (percent motion coherence). Across trials, the percent coherence is varied. Using the coherence stimulus, Newsome, Britten and Movshon (1989) found that, as the dots' coherence increased, an MT neuron's firing rate increased, and a monkey judged the direction of movement more accurately. At a coherence value in the neighborhood of 12.8%, the MT neuron fired significantly greater than baseline, and motion was judged correctly on virtually all trials. Lesions of MT cortex reduce the number of correct judgments of dot direction (Newsome and Pare, 1988), and micro-stimulation of a column of *DS* MT cells during an experimental trial leads a monkey to shift its judgment in the direction of the stimulated cells (Movshon and Newsome, 1992). Single-cell responses to optic flow patterns (expansion/contraction or rotation) which produce induced self-motion in humans have been found in the MSTd area of the monkey cortex. Tanaka, Fukada and Saito (1989) proposed a scheme to explain the obtained preferences of MSTd cells for specific patterns of optic flow. Each MSTd cell was hypothesized to receive inputs from a number of MT cells with appropriate direction tuning and receptive field location.

Neurological studies of brain lesion deficits in humans reinforce psychophysical and computational studies which propose separate, specialized detectors for second-order motion. In clinical studies, one patient suffered brain damage which impaired perception of motion with first-order stimuli, but not second-order stimuli; a second patient with different brain damage had impaired second-order motion, but not first-order motion (Vaina, Cowey and Kennedy, 1999). In a thorough fMRI study Smith et al. (1998) examined activity levels produced by first-order motion and three types of second-order motion in seven different areas of the human visual cortex. Area V5 was found to be strongly activated by second-order as well as by first-order motion. Activity in Area V3 and VP was significantly greater for second-order motion than for first-order motion. The results are consistent with the hypotheses that first-order motion sensitivity arises in V1, that second-order motion is first represented explicitly in V3 and VP, and that V5 is involved in further processing of motion information, including the integration of motion signals of the two types. It should be noted that the hypotheses are in agreement with the findings from single-cell and neurological studies cited above on the M-pathway, but the conclusions about the exact physiological substrates of first- and second-order motion in humans should still be regarded as tentative.

The relationship between the M-pathway and attention is an important one for guiding behavior. The search for a target in a complex natural scene is generally a serial one in which saccadic eye movements and attention are directed successively to different salient areas (Parkhurst, Law and Niebur, 2002). Salient target areas are processed more completely and quickly than non-salient areas. Among other variables, first-order stimulus cues such as intensity or luminance contrast have been shown to contribute significantly to saliency. Second-order stimulus features, like orientation or texture contrast, are less effective in demarcating salient areas. Furthermore, a number of studies suggest that eye movements and the deployment of visual attention to salient areas defined by first-order stimulus cues are mediated by the M-pathway (Cheng, Eysel and Vidyasagar, 2004; Parkhurst, Law and Niebur, 2002; Steinman, Steinman and Lehmkuhle, 1997). Static second-order or isoluminant color cues, which activate the P-pathway alone, are less effective in signaling salient areas. It is not surprising then that stimuli which are designed to activate the M-pathway dominate visual processing when put in competition with stimuli which activate the P-pathway alone (Steinman, Steinman and Lehmkuhle, 1997) or that they produce faster response times in a search task (Cheng, Eysel and Vidyasagar, 2004). As a consequence, displays/MHD’s that highlight potential targets with flickering or moving markers would be more effective than those which employ markers based upon other visual dimensions (e.g., color) (Pinkus, Poteet and Pantle, 2008).

A review and thoughtful analysis of the many types of visual motion phenomena makes it clear that visual motion is not a simple perception mediated by a single, unitary mechanism or process. It is a complex perceptual dimension elaborated in a specialized pathway, which itself contains sub-pathways and multiple stages of analysis.

Monocular vs. Binocular Vision

The use of HMD systems is more prevalent in today’s complex operational environment to increase Warfighters’ situational awareness, command and control, survivability, and mobility. The dismounted Warfighter must maintain situation awareness—both globally and locally—during operational tasks such as land navigation, target identification and location and usually must do all this while moving within a complex operational environment of coarse terrain and adverse climates. Hence, HMDs provide Warfighters with visual enhancement in conditions where the unaided eye would be less than an optimal tool. HMDs display symbology or imagery to either one eye (i.e., monocular HMDs) or both eyes (i.e., binocular/biocular HMDs) by the way of imaging sensor systems—e.g., image intensification (I^2) and forward-looking infrared (FLIR)—that have been incorporated into military aircraft and mounted vehicles.

Despite the potential visual and operational advantages of HMDs, there can be problems with their use. For instance, a number of studies have documented complications such as eye and oculomotor strain, dizziness, nausea, headache, disorientation, visual illusion and visual distortion (Kooi, 1986; Rash and Hiatt, 2005; Rash et al., 2001; Wenzel, 2002). These problems are likely to be induced by the unnatural viewing conditions of HMDs.

Large differences exist between naturally perceived vision (e.g., cues of depth and true stereopsis) and the monocular or binocular/biocular vision obtained through HMDs. These problems may account for some reduction in visual performance while wearing HMDs such as decline of distance judgment, response time delay and target identification (Arditi, 1986; Conticelli and Fujiwara, 1964; Ginsburg and Easterly, 1983). Consequently, there are a number of visual perception trade-offs that must be considered during a 'human-centered' approach toward HMD selection (i.e., monocular vs. binocular/biocular) and design process (Leger, 1994).

Monocular viewing

Monocular HMDs have the advantage of being smaller, lighter weight, and lower cost than binocular designs. Monocular presentation also allows one eye always to be available for viewing cockpit instrumentation or for dark adaptation. However, two major concerns are associated with monocular HMDs: *binocular rivalry* and *suppression*. When wearing a monocular HMD, the optical input to the two eyes differs greatly; thus creating potential interocular differences in color, contrast, brightness, shape, size, motion, and accommodation demand (Patterson, 2006; Velger, 1998). In fact, visual problems associated with monocular visual stimulation by the Apache IHADSS have been reported during both combat and non-combat missions (Crowley, 1992; Rash and Hiatt, 2005; Rash et al., 2001). Among the most common reported complaints are: degraded visual cues, visual illusions (static and dynamic), and visual discomfort.

Depending on the type of monocular HMD, one eye views the symbology of the HMD while both eyes view the real world scene. Alternately, with other monocular HMDs such as the IHADDS, one eye (i.e., right) views the displayed symbology while the other eye (i.e., left) views the external real world scene or the cockpit displays. This perceptual condition is referred to as dichoptic viewing, which can induce binocular rivalry—the alternation of perceived images that results when different visual images are presented to the two eyes and cannot be fused into a single percept. Binocular rivalry usually is resolved by suppressing the visual input unilaterally, and the attention may alternate spontaneously between the views received from each eye (Patterson, 2006). However, suppression can further reduce the visibility of the background or the monocular symbology. Furthermore, such dichoptic viewing, under sustained periods of monocular viewing and suppression, places great demands on the visual system and may be expected to result in high workload and stress levels. Although alternation and suppression of an image are largely unconscious or involuntary, some pilots can, to some extent, learn to selectively suppress an image or reach conscious control over alternating images (Malkin, 1987). Winterbottom (2006) showed that binocular fusion of a static background scene can partially mitigate the incidence of visual suppression when wearing a monocular semi-transparent (see-through) HMD. However, suppression was not prevented when a dynamic background scene was viewed. These results are consistent with the notion that moving stimuli are more dominant than stationary stimuli during the rivalry process (Fox and Check, 1972; Norman, 2000). To add to the complexity of monocular HMD-induced rivalry problem, several other factors such as exposure time, spatial frequency, size, luminance and contrast level can affect the strength of the stimulus during the rivalry process (Winterbottom, 2006). Binocular rivalry is further discussed in Chapter 12, *Visual Perceptual Conflicts and Illusions*.

Eye dominance is another important factor to consider when viewing imagery through a monocular HMD. Eye or sighting dominance refers to the tendency to prefer one eye over the other for monocular tasks. This consideration is more critical when the design of the monocular HMD does not allow the pilot to select his preference eye—i.e., IHADSS is always displayed to the right eye. The IHADSS monocular design forces the Apache aviator to switch his visual input between the two eyes depending on the required task. Winterbottom (2006) showed that the aviator's ability to intentionally switch dominance between the two visual stimuli can also affect the visibility and detection threshold of targets undergoing rivalry suppression. An ongoing study to determine if the intermittent use of the monocular HMD by British Apache aviators has any long-term effect on binocular visual performance has the potential to clarify the role of eye dominance on aviator's performance while wearing a monocular HMD (Rash and Hiatt, 2005).

Perhaps, one of the greatest disadvantages of monocular HMDs is their reduced FOV. In fact, most Apache pilots partially attribute their physical fatigue and headaches to the narrow FOV provided by the IHADSS (30° [V] by 40° [H]) (Rash and Hiatt, 2005). The extent of available FOV can also be affected by the size of the exit pupil. Light passing through the optical system form an image at the exit pupil, therefore the eye will not capture some of the light rays if the eye is not placed directly at the exit pupil but instead laced behind or in front of it. Issues related to a reduced exit pupil can be overcome by positioning the helmet display unit (HDU) as close as possible to the eye and by maintaining a very stable head-helmet interface. A stable fit of the helmet is paramount to maintain the optimum exit pupil size in the presence of the high-vibration environment of military helicopters (Rash, 1987). These modifications will also maximize the FOV of the monocular HMD system.

Binocular/Biocular viewing

Efficient binocular vision occurs when the retinal image of both eyes are in good focus and of similar size and shape. In particular, both eyes must be capable of aligning themselves in a way that the retinal images of a fixed scene are located at the foveae (i.e., small regions of highest VA) of the two eyes. Proper eye alignment (i.e., motor fusion) results in response to retinal disparity which serves as a cue to activate eyes movement toward one another (i.e., convergence) or away from one another (i.e., divergence). In turn, motor fusion is required to achieve sensory fusion of the images into a single percept. Appropriate levels of motor and sensory fusion will prevent perceptual problems such as diplopia (i.e., double vision), rivalry and suppression as well as visual discomfort and stress (Grosvenor, 1996). Similarly, proper alignment and adjustment of binocular or biocular HMDs, with relation to the Warfighter's eyes, is required to achieve functional vision and prevent visual perceptual problems and eye strain.

An HMD is classified as *binocular* if it presents an identical visual scene to the two eyes from slightly different perspectives via two sensors displaced in space allowing the Warfighter to perceive the image with stereoscopic depth perception or stereopsis. However, a binocular presentation can be achieved using a single sensor if the sensor is manipulated (e.g., temporal delay) to provide two slightly different perspectives of the same visual scene. In contrast, a *biocular* display presents the same image to both eyes from the same perspective so that the resulting view is a two-dimensional display. This is attained using a single sensor as it is the case of the HMD currently in development by Vision System International, San Jose, CA, for the Joint Strike Fighter F-35. Systems that allow binocular perception have substantial advantages over those that provide monocular presentation since binocular visualization is closer to the natural conditions of the human visual system. Unfortunately, from the design point of view, building binocular systems are technically more complex, heavier and of a relative higher cost compared to monocular HMDs. Consequently, their development can call for several design trade-offs.

Generally, binocular and biocular HMDs prevent rivalry and suppression problems usually encountered with monocular HMDs. Moreover, several studies support the notion that binocular vision enhances visual functions such as brightness perception, VA, and contrast sensitivity over the entire spectrum of spatial frequencies as well as the extent of the visual field (Arditi, 1981; Campbell and Robson, 1968; Thorn and Boynton, 1974). These visual improvements are ascribed to binocular summation. As the name implies, binocular summation means that the detection threshold for a stimulus is lower with two eyes than with one; therefore providing an enhanced single binocular percept.

Binocular and binocular HMDs can achieve a larger FOV by presenting a partially overlapped FOV. This is designed to present monocular images to both eyes at the same time with some overlap of the two monocular FOV. Basically, partial overlapped HMDs have a central field of binocular (or biocular in the case of biocular HMDs) overlap region and peripheral regions of monocular viewing (Velger, 1998) and mimics the field of view of the two eyes in unaided vision. Such a partial overlap can be presented by either a convergent or divergent design (Leger, 1994; Rash 2001). A divergent design allows both eyes of the observer to see the central overlap region as well as the right monocular and left monocular regions with the right and left eye, respectively (Figure 10-17). In contrast, a convergent design allows both eyes to see the central overlap region, but the right monocular and left monocular regions are seen only with the left and right eye, respectively (Figure 10-18). For binocular HMDs, optimal conditions for binocular vision are achieved with a convergence design as it resembles the natural mechanism of visual perception and facilitates the processing of binocular disparity cues required to achieve stereopsis (Klymenko, 1994; Leger, 1994; Melzer and Moffitt, 1991). This implies that binocular vision is an essential element to attain stereopsis. Although convergent or divergent partial overlap displays provides larger FOV and stereoscopic advantages, they can potentially create perceptual conflicts such as luning (Figure 10-19). Luning is a subjective darkening in the flanking monocular regions of the FOV near the binocular overlap borders. These regions of luning can interfere with common visual tasks performed by Warfighters such as target detection (Klymenko, 1994).

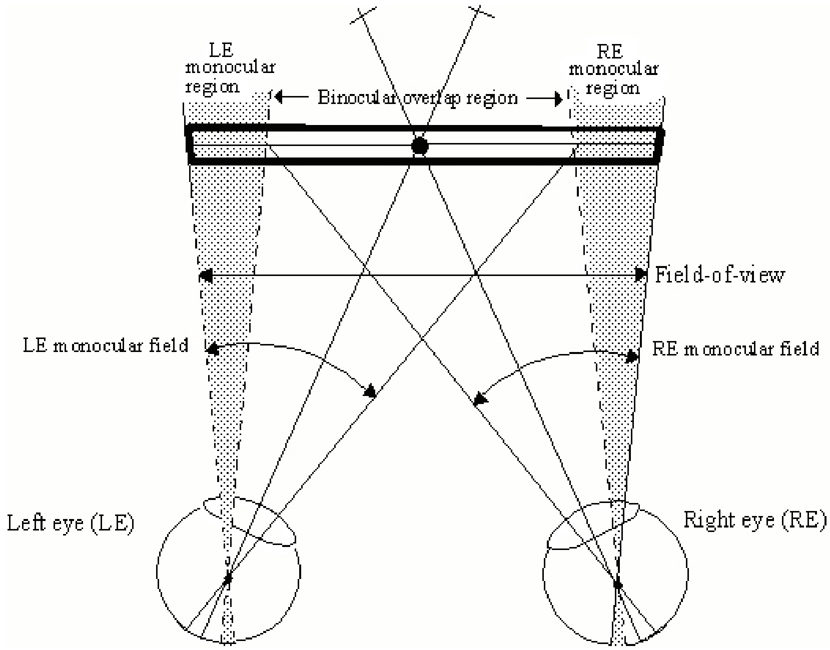


Figure 10-17. Visual interpretation of the divergent display mode of partially-overlapped HMD designs (Rash, 2001).

As discussed in the previous section, *Binocular vs. Monocular Vision*, binocular and biocular HMDs can use partial overlap of the monocular FOVs to achieve a larger FOV. They are designed to present monocular images to both eyes at the same time with some overlap of the two monocular FOV. But in order to provide stereopsis (i.e., binocular HMD) or enhanced monocular cues for depth (i.e., biocular HMDs), part of the available FOV from the two monocular fields must be sacrificed to gain the partial overlap region (Parrish and Williams, 1993). If the partial overlap is created by a binocular HMD system, the resulting central overlap region will provide the binocular disparity cues required to achieve stereopsis. In contrast, if the visual field is provided by a biocular design, the central overlap region of the FOV will only provide monocular cues for perception of depth; thus, cannot provide binocular disparity cues or stereopsis. Moreover, since both eyes of the Warfighter are viewing the same single image with a biocular HMD, the absence of cues for retinal disparity is a strong binocular cue to flatness. This cue to flatness can be in direct conflict with the monocular depth cues that are provided by a single image of the scene (CuQlock-Knoop, 1997). At expense of a reduced FOV, a complete overlap of the images can provide the retina with identical images (i.e., true biocular HMDs) or images with binocular disparity (i.e., binocular HMD) that provides the Warfighter with an extra depth cue.

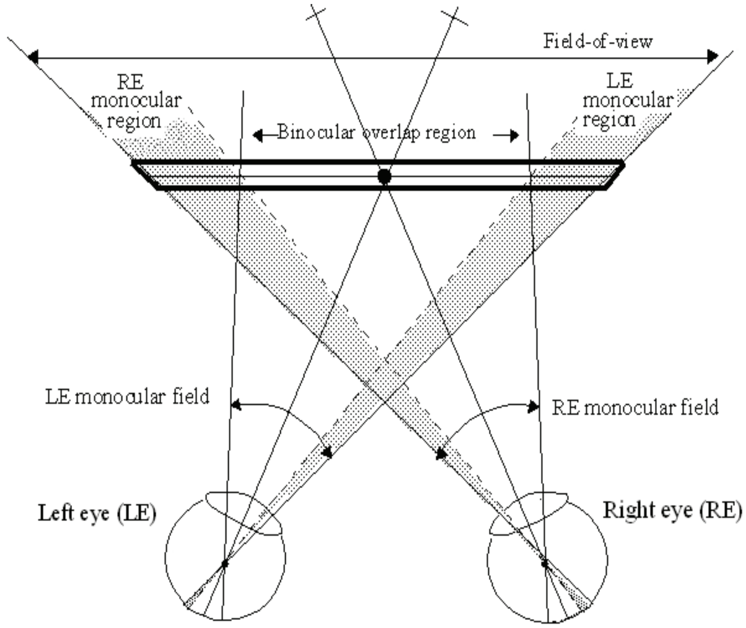


Figure 10-18. Visual interpretation of the convergent display mode of partially-overlapped HMD designs (Rash, 2001).

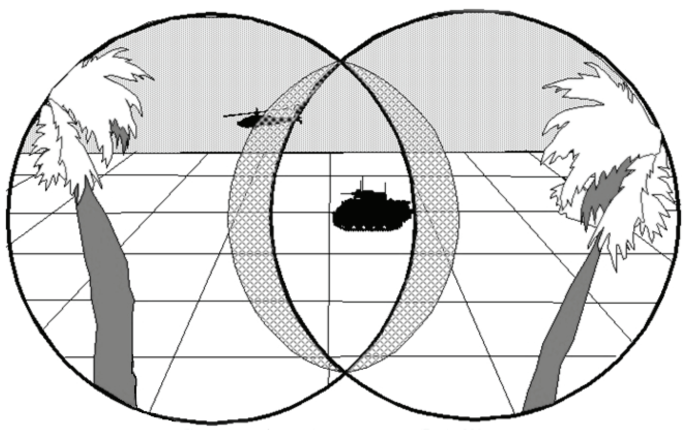


Figure 10-19. Luning in partial overlap displays (Rash, 2001).

The use of binocular or biocular HMDs introduces the possibility to have mismatches between the imagery presented to the two eyes. There are numerous reasons for this, some of which are induced by alignment errors and others by optical image differences. Self (1986) provided a summary of the optical tolerance limits of binocular HMDs in terms of vertical, convergence, and divergence misalignments, as well as rotational, magnification, and luminance differences. Also, proper alignment of the interpupillary distance of the NVG has been determined to be essential to prevent disruption of depth perception (Sheehy and Wilkerson, 1989). A more recent study by Kooi and Toet (2004) using static images demonstrated that in spite of the enhanced perception obtained with stereoscopic displays, a small amount of asymmetry between the two images (i.e., stereo imperfections) has the potential to reduce visual comfort. Stereo imperfections are induced by many factors such as optical errors (i.e., spatial distortions), imperfect filters (i.e., photometric asymmetries including luminance, color, and contrast), and stereoscopic disparities. This study also provides threshold values for the onset of visual discomfort induced by these factors of binocular image imperfections that should be taken into account during the HMD design and selection process.

Binocular cues for depth perception

Binocular cues for depth includes retinal disparity, convergence, and accommodation. Since these are innately determined, these cues come into play during the first few months of life as a consequence of the development and maturation of the visual pathways to the brain. Some neurons in the visual cortex are able to detect retinal disparity and act as depth detectors. Retinal disparity is the predominant cue for depth and results when a scene stimulates disparate (non-corresponding) retinal points in the two eyes. If the amount of retinal disparity is small, the observer will perceive stereopsis; otherwise the observer will experience diplopia. Empirical data by Boff and Lincoln (1988) demonstrated that retinal disparity can provide depth information from a distance up to 264 meters (866 feet). A subsequent study by Roumes et al. (2001) showed that binocular disparity can improve distance estimation using stereoscopic displays with stereo-near configuration – i.e., the point of zero disparity located is at the nearest point visible in the scene – for a range of distances up to 160 meters (525 feet).

Convergence and accommodation provide weak propioceptive (i.e., position sense) cues for depth. Convergence serves as a cue for depth because the convergence of the eyes depends on the distance of the fixating object. Therefore, it provides oculomotor propioceptive information arising from extraocular muscles and changes of the angle of inclination of the eyes. Accommodation also serves as a depth cue because the shape of the lens depends on the distance of the object an observer focuses on. Accommodation of the lens in response to blur provides information concerning position sense arising from the ciliary muscle. A study by Sheehy and Wilkinson (1989) with helicopter pilots that had failed a test of stereoscopic depth perception after a prolonged flight training employing night vision goggles suggested that loss of stereopsis might have been caused by a shift in lateral phorias. In this particular case, it would be expected that as additional fusional effort is required, the minimum resolvable disparity degrades due to increases in accommodation brought about through vergence accommodation.

Monocular cues for depth perception

Monocular cues for perception of depth are empirical cues that must be learned and therefore they are developed more slowly. Monocular cues for depth include relative size, overlay, geometrical perspective, aerial perspective, as well as light and shadow (Grosvenor, 1996) (Figure 10-20). The *relative size* of an image depends upon its distance from the observer. The size of the image is small when the object is far away, and becomes larger as the object approaches the observer. *Overlay* (i.e., interposition) refers on how an object that partially blocks another object is interpreted as being closer. *Geometrical* or *linear perspective* is perhaps the most common monocular cue of depth. The basis for the cue of linear perspective is given by the fact that distant objects necessarily produce a smaller retinal image than nearby objects of the same size. Consequently, the horizontal separation of the two sides of parallel lines (e.g., railroad track, road) converges toward the horizon – larger for the near portion of the parallel lines and smaller for the more distant portions. *Aerial perspective* or height as a monocular cue of depth is based on the perception that the further away an object is from the observer the higher in the visual field its image will be interpreted. The distribution of *light and shadow* on an object is also a dominant monocular cue for depth provided by the assumption that light comes from above. It also takes into account that objects do not usually allow light to pass through, therefore, they will cast a shadow. These monocular cues are of particular importance for Warfighters wearing monocular and biocular HMDs, but they also can offer enhanced details of the viewed scene while wearing a binocular HMD.

In summary, operational and occupational requirements for depth perception or stereopsis will strongly influence the final design of a particular HMD. While binocular HMDs provided the operator with stereopsis and perception of depth when monocular cues are absent, monocular or biocular HMDs can only provide perception of depth when monocular cues are present. Nevertheless, monocular cues enhance the operator's ability to perceived stereopsis while wearing a binocular HMD.

Figure 10-20. This picture of a complex scene demonstrates how monocular cues (relative size, overlay, geometrical perspective, aerial perspective, as well as light and shadow) are used by the human visual system to perceive depth or relative distance between objects in a two dimensional image in the absence of binocular cues.

References

Adelson, E.H. (1999). Lightness perception and lightness illusions. In: *The Cognitive Neurosciences*, Gazzaniga, M., (Ed.), Cambridge, MA: MIT Press.

Adelson, E., and Bergen, J. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2, 284-299.

Angel, S-D. (2003). IPNVG fixed-focus diopter study. *Proceedings of SPIE, Helmet- and Head-Mounted Displays VIII*, 5079, 154-163.

Angel, S-D., and Baldwin, J.B. (2004). Aircrew visual acuity viewing with different night vision goggle eyepiece diopter settings. *Proceedings of SPIE, Helmet- and Head-Mounted Displays IX*, 5442 1-12.

Anstis, S. (1970). Phi movement as a subtraction process. *Vision Research*, 10, 1411-1430.

Anstis, S. (1978). Apparent movement. In: *Handbook of Sensory Physiology*, Held, R., Leibowitz, H., and Teuber, H.L., (Eds.), *Perception* (Vol. VIII). New York: Springer-Verlag.

Anstis, S., and Rogers, B. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, 15, 957-961.

Arditi, A. (1986). Binocular vision. In: *Handbook of Perception and Human Performance,* Boff, K.R., and Thomas, J.P., (Eds.). New York: Wiley and Sons.
Arditi, A.R., Anderson, P.A., and Movshon, J.A. (1981). Monocular and binocular detection of moving sinusoidal gratings. *Vision Research*, 21, 329-36.
Attneave, F. (1971). Multistability in perception. *Scientific American*, 225, 63-71.
Bailey, I.L., and Lovie, J.E. (1976). New design principles for visual acuity letter charts. *American Journal of Optometry and Physiological Optics*, 53, 740-745.
Bailey, M.D., Mitchell, G.L., Dhaliwal, D.K., Boxer Wachler, B.S., and Zadnik, K. (2003). Patient satisfaction and visual symptoms after laser in situ keratomileusis. *Ophthalmology*, 110, 1371-8.
Barfield, W., and Furness, T.A. (Eds.) (1995). *Virtual Environments and Advanced Interface Designs*. Table 6-13, p. 226. New York: Oxford Press.
Barlow, H., and Hill, R. (1963). Evidence for a physiological explanation of the waterfall phenomenon and figural aftereffects. *Nature (London)*, 200, 1345-1347.
Blake, R., and Sekuler, R. (2005). *Perception* (5th Ed). New York: McGraw-Hill.
Boff, K.R., and Lincoln, J.E. (1988). Engineering data compendium: human perception and performance. Wright-Patterson Air Force Base: Armstrong Aerospace Medical Research Laboratory.
Braddick, O. (1974). A short-range process in apparent motion. *Vision Research*, 14, 519-527.
Braunstein, M.L. (1976). *Depth Perception Through Motion*. New York: Academic Press.
Campbell, F.W., and Robson, J.G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197, 551-66.
Cavanagh, P., and Mather, G. (1989). Motion: the long and short of it. *Spatial Vision*, 4, 103-129.
Cheng, A., Eysel, U., and Vidyasagar, T. (2004). The role of the magnocellular pathway in serial deployment of visual attention. *European Journal of Neuroscience*, 20, 2188-2192.
Chubb, C., and Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, 5, 1986-2007.
Conticelli, M., and Fujiwara, S. (1964). Visuo-motor reaction time under differential binocular adaptation. In: *Atti Della Foundizone*, Ronchi, G., (Ed.).
Crowley, J.S., Rash, C.E., and Stephens, R.L. (1992). Helmet-Mounted Displays III: Visual illusions and other effects of night vision devices. *Proceedings of SPIE*, 1695, 199-180.
CuQlock-Knopp, V.G., Myles, K., and Merritt, J.O. (1997). Perceptual conflicts between monocular and binocular depth of cues in biocular displays. *Digest*, 97, 955-958.
Cutting, J.E., and Vishton, P.M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In: Epstein, W., and Rogers, S. (Eds.) *Perceptual Space*. New York: Academic Press.
Dahm, M. (2006). *Grundlagen der Mensch-Computer-Interaction*. Munich, Germany: Pearson.
Davidoff, J. (1991). *Cognition through color*. Cambridge, MA: The MIT Press.
Department of Defense. (2001). Interface Standard. *Lighting, Aircraft, Night Vision Imaging System (NVIS) Compatible*. MIL-STD-3009.
Defense Supply Agency. (1962). *Military Standardization Handbook 141, Optical Design*. Washington, DC: Defense Supply Agency.
DeMars, S.A. (1975). Human factors considerations for the use of color in display systems. Cape Kennedy, FL: NASA. NASA TR-1329.
Dennis, R.J., Beer, J.M., Baldwin, J.B., Ivan, D.J., Lorusso, F.J., and Thompson, W.T. (2004). Using the Freiburg acuity and contrast test to measure visual performance in USAF personnel after PRK. *Optometry and Vision Science*, 81, 516-24.

Donohue-Perry, M.M., Task, H.L, and Dixon, S.A. (1994). Visual acuity vs. field of view and light level for night vision goggles. *Proceedings of SPIE, Helmet- and Head-Mounted Displays and Symbology Design Requirements*, 2218, 71-81.

Dudfield, H.J. (1991). Colour head-up displays: help or hindrance? *Proceedings of the Human Factors Society 35th Annual Meeting*, 146-150.

Emerson, R., Bergen, J., and Adelson, E. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, 32, 203-218.

Epstein, W., Park, J., and Casey, A. (1961). The current status of the size-distance hypothesis. *Psychological Bulletin*, 558(6), 491-514.

Ericksen, C.W., and Ericksen, B.A. (1974). Effects of noise letter upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, 16, 143-149.

Exner, S. (1875). Experimentelle Untersuchung der Einfachsten Psychischen Processe. *Archives of Gestalt Physiology*, 11, 403-432.

Fan-Paul, N.I., Li, J., Miller, J.S., and Florakis, G.J. (2002). Night vision disturbances after corneal refractive surgery. *Survey of Ophthalmology*, 47, 533-46.

Farrell, R.J., and Booth, J.M. (1984). *Design Handbook for Imagery Interpretation Equipment.* Seattle: Boeing Aerospace Company.

Fisher, E., Haines, R.F., and Price, T.A. (1980). Cognitive issues in head-up displays. Moffett Field, CA: NASA Ames Research Center. No 1711.

Flinn, S. (2000). Simultaneous contrast: luminance. Retrieved June 1, 2008 from University of British Columbia Web site: http://www.cs.ubc.ca/nest/imager/contributions/flinn/Illusions/ BW/ bw.html

Fox, R., and Check, R. (1972). Independence between binocular rivalry suppression duration and magnitude of suppression. *Journal of Experimental Psychology*, 93, 283-9.

Foyle, D.C., McCann, R.S. Sanford, B.D. and Schwirzke, M.F.J. (1993). *Proceedings of the 37th Annual meeting of the Human Factors and Ergonomics society*, Santa Monica, CA: Human Factors and Ergonomics Society. p. 1340-1344.

Franconeri, S.L., and Simons D.J., (2005). The dynamic events that capture attention: A reply to Abrams and Christ. *Perception and Psychophysics*, 67(6), 962-966.

Gibson, J.J. (1950). *The perception of the visual world*. Boston: Houghton-Mifflin.

Ginsburg, A.P. (1983). Direct performance of HUD systems using contrast sensitivity. In: *Optical and Human Performance Evaluation of HUD System Design*, Martin, W.L., (Ed.), Wright-Patterson AFB, OH: U.S. Air Force Aeromedical Research Laboratory.

Ginsburg, A.P., and Easterly, J. (1983). Contrast sensitivity predicts target detection field performance of pilots. *Proceeding of Human Factors Society*, 27th Annual Meeting. p. 269-273.

Gish, K.W., and Staplin, L. (1995). Human factors aspects of using head up displays in automobiles: A review of the literature. DOT HS 808 320 the Scientex Corp Landsdale PA, for the USDT National Highway Traffic Safety Administration.

Glick, D.D., and Moser, C.E. (1974). Afterimages associated with using the AN/PVS-5, night vision goggle. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL LR 75-1-7-1.

Goldstein, E.B. (2007). *Sensation and Perception* (7th Ed). Belmont, CA: Wadsworth.

Graham, C.H. (1966). Visual space perception. In: Graham, C.H. (Ed.), *Vision and Visual Perception*. New York: John Wiley and Sons, Inc.

Green, D., and Swets, J. (1966). *Signal detection theory and psychophysics*. Huntington, NY: Robert E. Krieger Publishing Company.

Greene, D.A. (1988). Night vision pilotage system field-of-view (FOV)/resolution tradeoff study flight experiment report. Fort Belvior, VA: U.S. Army Night Vision Laboratory. NV 1-26.

Grosvenor, T. (1996). Anomalies of binocular vision. In: *Primary care optometry: Anomalies of refraction and binocular vision*. Boston: Butterworth-Heinemann. p. 105-117.

Happ, A., and Pantle, A. (1987). Visual motion processing and display design. In: L. Mark, L., Warm, J., and Huston, R., (Eds.), E*rgonomics and Human factors: Recent Research*. New York: Springer-Verlag. p. 85-94.

Harding, T.H., Martin, J.S., and Rash, C.E. (2007). The legibility of HMD symbology as a function of background local contrast. *Proceedings of SPIE*, 6557, D1-D8.

Havig, P.R., Grigsby, S.S., Heft, E.L., LaCreta, J.M., and Post, D.L. (2001). Luminance contrast and color recognition in helmet-mounted displays. *Proceedings of SPIE*, 4361, 198-204.

Hering, E. (1878). *Zur Lehre vom Lichtsinne.* Vienna: Carl Gerolds Sohn.

Hermans, T.G. (1937). Visual size constancy as a function of convergence. *Journal of Experimental Psychology*, 21(2), 143-161.

Hooey, B.L., Foyle, D.C., Andre, A.D., and Parke, B. (2000). Integrating datalink and cockpit display technologies into current and future taxi operations. *Proceedings of the AIAA/IEEE 19th Digital Avionics System Conference*, 7.D.2-1 to 7.D.2-8.

Hubel, D., and Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's striate cortex. *Journal of Physiology*, 160, 106-154.

Hubel, D., and Wiesel, T. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215-243.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 23, 76-88.

Jonides, J., and Ynatis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception and Psychophysics*, 43(4) 346-354.

Kaiserman, I., Hazarbassanov, R., Varssano, D., and Grinbaum, A. (2004). Contrast sensitivity after wave front-guided LASIK. *Ophthalmology*, 111, 454-457.

Kalich, M.E., Rash, C.E., Thomas H. Harding, T.H., Jennings, S., Craig, G., and Stuart, G.W. (2009). Flight performance using a hyperstereo helmet-mounted display: post-flight debriefing questionnaire. *Proceedings of SPIE, Head-and Helmet-Mounted Displays XIV: Design and Applications*, 7326.

Kalich, M.E., Rash, C.E., McLean, W.E. and Ramiccio, J.G. (2007). A limited flight study for investigating hyperstereo vision. *Proceedings of SPIE, Head-and Helmet-Mounted Displays XII: Design and Applications*, 6557, 65570I1-65570I14.

Klymenko, V., Verona, R.W., Martin, J.S., Beasley, H.H., and McLean, W.E. (1994). The effects of binocular overlap mode on contrast thresholds across the field-of-view as a function of spatial and temporal frequency. U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL. USAARL Report No. 94-49.

Kooi, F.L. (1986). Visual strain: a comparison of monitors and helmet-mounted displays. In *Proceedings of SPIE: Imaging Sciences and Display Technologies*, 2949, 162-171.

Kooi, F.L., and Toet, A. (2004). Visual comfort of binocular and 3D displays. *Displays*, 25, 99-108.

Kotulak, J.C., and Rash, C.E. (1992). Visual acuity with second and third generation night vision goggles obtained from a new method of might sky simulation across a wide range of target contrasts. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 92-9.

Larish, I., and Wickens, C.D. (1991). Divided attention with superimposed and separated imagery: Implications for head-up displays (No 91-4/NASA HUD 91-1) Savoy, IL: University of Illinois, Aviation Research Laboratory.

Laycock, J., and Chorley, R.A. (1980). The electro-optical display/visual system interface: Human factors considerations. *Advancements on Visualization Techniques*, AGARD, p. 3/1- 3/15.

Leger, A., Roumes, C., Gardelle, C., Cursolle, J.P., and Kraus, J.M. (1994). Binocular HMD for fixed-wing aircraft: a trade-off approach. *Displays*, 15, 91-97.

Leibowitz, H. (1966). Role of changes in accommodation and convergence in the perception of size. *Journal of the Optical Society of America*, 56(8), 1120-1123.

Lennie, P. (1980). Parallel visual pathways: a review. *Vision Research*, 20, 561-594.

Levin, S.T., Momen, N., Drivdahl, S.B., and Simons, D.J., (2000). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition*, 7, 399-412.

Levine, R.R., and Rash, C.E. (1989). Visual acuity with AN/PVS-5A night vision goggles and simulated flashblindness protective lenses under varying levels of brightness and contrast. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 89-16.

Livingstone, M., and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, 240, 740-749.

Lu, Z., and Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update. *Journal of the Optical Society of America A*, 18, 2331-2370.

Luo, X., Kenyon, R., Kamper, D., Sandin, D. and DeFanti, T. (2007). The effects of scene complexity, stereovision, and motion parallax on size constancy in a virtual environment. *Proceedings of Virtual Reality Conference 2007, VR '07, IEEE*, Charlotte, North Carolina, 59-66.

Malkin, F.J. (1987). AH-64A Helmet-mounted display lessons learned. Discussion Paper No. 40. U.S. Army Human Engineering Laboratory. Aberdeen Proving Ground, Maryland.

Marasco, P.L., and Task, H.L. (2003). The impact of target luminance and radiance on night-vision device visual performance testing. *Proceedings of SPIE, Helmet- and Head-Mounted Displays VIII*, 5079, 174-183.

Marr, D., and Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society London B*, 211, 151-180.

Martinsen, G.L., and Havig, P.R. (2002). Luminance contrast requirements for colored symbols in helmet-mounted displays. Human Effectiveness Directorate, Crewsystem Interface Division. Dayton, OH; U.S. Air Force Research Laboratory.

Mather, G., Verstraten, F., and Anstis, S. (Eds). (1998). *The Motion Aftereffect: A Modern Perspective.* Cambridge, MA: MIT Press.

McCann, R.S., Foyle, D.C., and Johnson. (1993). Attentional limitations with head-up displays. *Proceedings of the7th International symposium of aviation psychology*. 26-29 April 1993. p. 70-75.

McCann, R.S., Lynch J., Foyle, D.C., and Johnson. (1993). Modeling Attentional effects with head-up displays. *Proceedings of the Human Factors and Ergonomics 37th annual meeting*, p. 1345-1349.

McCarthy, J., Pantle, A., and Pinkus, A. (1994). Detection and direction discrimination performance with flicker gratings in peripheral vision. *Vision Research*, 34, 763-773.

McKee, S., and Watamaniuk, S. (1994). The psychophysics of motion perception. In: Smith, A., and Snowden, R. (Eds.), *Visual detection of motion.* London: Academic Press.

Melzer, J.E., and Moffitt, K. (1991). Helmet-mounted displays. An ecological approach to partial binocular overlap. *Proceedings of SPIE*, 1456, 124-131.

McCarthy, J., Pantle, A., and Pinkus, A. (1994). Detection and direction discrimination performance with flicker gratings in peripheral vision. *Vision Research*, 34, 763-773.

McKee, S., and Watamaniuk, S. (1994). The psychophysics of motion perception. In: Smith, A., and Snowden, R. (Eds.), *Visual detection of motion.* London: Academic Press.

Merigan, W., and Maunsell, J. (1993). How parallel are the primate visual pathways. *Annual Review of Neuroscience*, 16, 369-402.

Miller, R., Provines, W., Block, M., Miller, J., and Tredici, T. (1984). Comparative visual performance with ANVIS and AN/PVS-5A night vision goggles under starlight conditions. Brookes AFB, TX: USAF School of Aerospace Medicine. Technical Report No. USAFSAM-TR-84-28.

Moffitt, K., Rogers, S.P., and Cicinelli, J. (1988). Chromatic aftereffects associated with a night vision goggle simulation. *Aviation, Space, and Environmental Medicine*, 59(2), 125-128.

Mon-Williams, M., and Wann, J.P. (1998). Binocular virtual reality displays: When problems do and don't occur. *Human Factors*, 40(1), 42-49.

Movshon, A., and Newsome, W. (1992). Neural foundations of visual motion perception. *Current directions in psychological science*, 1, 35-39.

Movshon, A., and Newsome, W. (1996). Visual response properties of striate cortical neurons projecting to Area MT in macaque monkeys. *Journal of Neuroscience,* 16, 7733-7741.

Mrochen, M., Kaemmerer, M., Mierdel, P., and Seiler, T. (2001). Increased higher-order optical aberrations after laser refractive surgery: a problem of subclinical decentration. *Journal of Cataract and Refractive Surgery*, 27, 362-369.

Murakami, I., and Shimojo, S. (1996). Assimilation-type and contrast-type bias of motion induced by the surround in a random-dot display: evidence for center-surround antagonism. *Vision Research*, 36, 3629-3639.

National Academy of Sciences. (1980). Recommended standard procedures for the clinical measurement and specification of visual acuity. *Report of Working Group 39, Committee on Vision, Advances in Ophthalmology,* 41, 103-148.

National Aeronautics and Space Administration. (2008). Color discrimination and identification. Retrieved 5 October 2008 from NASA Ames Research Center website: http://colorusage.arc.nasa.gov/discrim.php

Nawrot, M., and Sekuler, R. (1990). Assimilation and contrast in motion perception: explorations in cooperativity. *Vision Research*, 10, 1439-1451.

Newsome, W., Britten, K., and Movshon, J. (1989). Neuronal correlates of a perceptual decision. *Nature*, 341, 52-54.

Newsome, W., and Pare, E. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, 8, 2201-2210.

Norman, H.F., Norman, J.F., and Bilotta, J. (2000). The temporal course of suppression during binocular rivalry. *Perception*, 29, 831-41.

Ogle, K.N. (1950). *Research in binocular vision*. Philadelphia, PA: Saunders.

Ogle, K.N., and Schwartz, J.T. (1959). Depth of focus of the human eye. *Journal of the Optical Society of America*, 49, 273-80.

Oram, M., and Perrett, D. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to “biological motion” stimuli. *Journal of Cognitive Neuroscience*, 6, 99-116.

Owsley, C., Sekuler, R., and Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research*, 23, 689-99.

Owsley, C., and Sloane, M.E. (1987). Contrast sensitivity, acuity, and the perception of real-world targets. *British Journal of Ophthalmology*, 71, 791-6.

Palmer, S. (1999). *Vision Science*. Cambridge, MA: MIT Press.

Pantle, A. (1973). Stroboscopic movement based upon global information in successively presented visual patterns. *Journal of the Optical Society of America*, 63, 1280A.

Pantle, A., and Picciano, L. (1976). A multi-stable movement display: Evidence for two separate motion systems in humans. *Science*, 193, 500-502.

Pantle, A., and Sekuler, R. (1968). Velocity-sensitive elements in human vision: Initial psychophysical evidence. *Vision Research*, 8, 445-450.

Pantle, A. (1992). Immobility of some second-order stimuli in human peripheral vision. *Journal of the Optical Society of America A*, 9, 863-867.

Pantle, A., and Turano, K. (1992). Visual resolution of motion ambiguity with periodic luminance- and contrast-domain stimuli. *Vision Research*, 32, 2093-2106.

Parkhurst, D., Law, K., and Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42, 107-123.

Parrish, R.V., and Williams, S.P. (1993). Trade-offs arising from mixture color cueing and monocular, binocular and stereoscopic cueing information for simulated rotorcraft flight. CECOM Technical report 92-B-014. NASA Technical Paper 3268.

Patterson, R., Winterbottom, M.D., and Pierse, B.J. (2006). Perceptual issues in the use of head-mounted visual displays. *Human Factors*, 48, 555-573.

Peli, E. (1995). Real vision and virtual reality. *Optics and Photonics News*, July. p. 28-34.

Petersik, J. (1989). The two-process distinction in apparent motion. *Psychological Bulletin*, 106, 107-127.

Petersik, J., and Pantle, A. (1979). Factors controlling the competing sensations produced by a bistable stroboscopic motion display. *Vision Research*, 19, 143-154.

Pinkus, A.R., Task, H.L., Dixon, S.A., Barbato, M.H., and Hausmann, M.A. (2003). Twenty-Plus Years of Night Vision Technology: Publications and Patents From the Crew System Interface Division of the Air Force Research Laboratory at Wright-Patterson Air Force Base, Ohio. Wright-Patterson AFB, OH: Air Force Research Laboratory. Report No. AFRL-HE-WP-TR-2003-0048.

Pinkus, A.R., and Task, H.L. (2000). Night vision goggles (NVG) objective lens focusing methodology and effects of aircraft canopies on NVG focus. *Proceedings of the 2000 SAFE Association Symposium,* Reno, NV, October 9-11, 2000.

Pinkus, A.R., and Pantle, A. (1997). Probing visual motion signals with a priming paradigm. *Vision Research*, 37, 541-552.

Pinkus, A.R., Poteet, M., and Pantle, A. (2008). Dynamic stimulus enhancement with Gabor-based filtered images. *Proceedings of SPIE*, 6968, 69681R.

Pinkus, A.R., Task, H.L., Dixon, S., and Goodyear, C. (2000). Reproducibility limits of night vision goggle visual acuity measurements. *SAFE Journal*, 30(1), 31-139.

Pinkus, A.R., Task, H. L., Dixon, S., and Goodyear, C. (1999). Reproducibility of night vision goggle visual acuity measurements using Landolt C's. *Proceedings of the 1999 SAFE Association Symposium*, Atlanta, GA, December 6-8, 1999.

Pinkus, A.R., and Task, H.L. (1998). Measuring observers' visual acuity through night vision goggles. Wight-Patterson AFB, OH: U.S. Air Force Research Laboratory. Technical Report No. ASC98-1884.

Pinkus, A.R., and Task, H. L. (1997). The effects of aircraft transparencies on night vision goggle-mediated visual acuity. *Proceedings of the 1997 SAFE Association Symposium*, Phoenix, AZ, September 8, 1997.

Post, D.L., Monnier, P., and Calhoun, C.S. (1997). Predicting color breakup on field-sequential displays. *Head-Mounted Displays II, Proceedings of SPIE*, 3058, 57-65.

Post, D.L., Sarma, K.R., Trimmier, J.R., Heinze, W.C., Rogers, J.C., Ellis, R.K., Larson, B.D., and Franklin, H. (1994). A new color display for head-mounted use. *Journal of the Society of Information Display*, 2(4), 155-163.

Prinzel III, L.J., and Risser, M. (2004). Head-up displays and attention capture. NASA Langley Research Center NASA/TM-2004-213000 Hampton VA.

Priot, A., Houlier, S., Guillaume, G., Leger, A., and Roumes, C. (2006). Hyperstereopsis in night vision devices:Basic mechanisms and impact for training requirements. *Helmet- and Head-Mounted Displays X: Technologies and Applications. Proceedings of SPIE*, 6224, 6224ON-1 to 6224ON-11.

Rabin, J. (1993). Spatial contrast sensitivity through aviator's night vision imaging system. *Aviation, Space, and Environmental Medicine*, 64, 706-10.

Rabin, J. (1996). Comparison between green and orange visual displays. *Journal of the Society of Information Display*, 4(2), 107-110.

Rash, C.E. (Ed.) (2001). *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft*. Bellingham, WA: SPIE.

Rash, C.E., Verona, R.W., and Crowley, J.S. (1990). Human factors and safety considerations of night vision systems flight using thermal imaging systems. *Proceedings of SPIE, Helmet-Mounted Displays II*, 1290, 142-164.

Rash, C.E., and Hiatt, K.L. (2005). Apache flight experience in Operation Iraqi Freedom (OIF) shows reduced incidents of visual illusions and problems with the monocular helmet-mounted display. *American Helicopter Society 61st Annual Forum*, 51-66.

Rash, C.E., Kalich, M.E., and van de Pol, C. (2002). Solutions to helmet-mounted display visual correction compatibilities issues. *Proceedings of SPIE*, 4711, 270-289.

Rash, C.E., Martin, J.S., Gower, D.W., Licina, J.R., and Barson, J.V. (1987). Evaluation of the U.S. Army fitting program for the Integrated Helmet Unit of the Integrated Helmet and Display Sighting System. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 87-8.

Rash, C.E., Suggs, C.L., Mora, J.C., van de Pol, C., Reynolds, V.S., and Crowley, J.S. (2001). Visual issues survey of AH-64 Apache aviators (Year 2000). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2002-02.

Regan, D., and Beverley, I. (1978). Illusory motion in depth: aftereffect of adaptation to changing size. *Vision Research*, 18, 209-212.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In: Rosenblith, W.A., (Ed.), *Sensory Communication*. New York: Wiley.

Reinhart, W.F., and Post, D.L. (1996). Human factors of two-primary color AMLCDs. *Proceedings of SPIE, Head-mounted Displays*, 2735, 105-114.

Riegler, J.T., Whiteley, J.D., Task, H.L., and Schueren, J. (1991). The effect of signal-to-noise ratio on visual acuity through night vision goggles. Wright-Patterson AFB, OH: Armstrong Laboratory. Report No. AL-TR-1991-0011.

Robson, J. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America*, 56, 1141-1142.

Rodman, H., and Albright, T. (1987). Coding of visual stimulus velocity in area MT of the macaque. *Vision Research*, 27, 2035-2048.

Roscoe, S.N. (1984). Judgements of size and distance with imaging displays. *Human Factors*, 26(6), 617-629.

Roumes, C., Meehan, J. W., Plantier, J., and Menu, J. (2001). Distance estimation in a 3-D imaging display. *The International Journal of Aviation Psychology*, 11, 381-396.

Sanford, B.D., Foyle, D. C., McCann, R.S., and Jordan K. (1993). Head-up displays: effects of information location on the processing of superimposed symbology. In: *Proceedings of the 7th International Symposium on Aviation Psychology*, Jensen, R.S. (Ed.), p. 81- 87.

Schulte-Pelkum, J., Riecke, B. and von der Heyde, M. (2003). Influence of display parameters on perceiving visually simulated ego-rotations—a systematic investigation. Poster presented at TWK 2003, Tubingen, Germany.

Sekuler, R., and Ganz, L. (1963). Aftereffect of seen motion with a stabilized retinal image. *Science*, 139, 419-420.

Sekuler, R., and Pantle, A. (1967). A model for the aftereffects of seen movement. *Vision Research*, 7, 427-439.

Sekuler, R., Pantle, A., and Levinson, E. (1978). Physiological basis of motion perception. In: *Handbook of Sensory Physiology*, Held, R., Leibowitz, H., and Teuber, H-L., (Eds.), *Perception* (Vol. VIII). New York: Springer-Verlag.

Sheehy, J.B., and Wilkinson, M. (1989). Depth perception after prolonged usage of night vision goggles. *Aviation, Space, and Environmental Medicine*, 60, 573-9.

Shelden, S.G., Foyle, D. C., and McCann, R. S. (1997). Effects of scene-linked symbology on flight performance. *Proceedings of the 41st Annual meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA: Human Factors and Ergonomics Society. p. 294-298.

Self, H.C. (1986). Optical tolerances for alignment and image difference for binocular helmet-mounted displays. Dayton, OH: Armstrong Aerospace Medical Research Laboratories. AMRL-TR-86-109.

Simoncelli, E., and Heeger, D. (1998). A model of neuronal responses in visual Area MT. *Vision Research*, 38, 743-761.

Simpson, W.A. (1989). The Step method: A new adaptive psychophysical procedure. *Perception and Psychophysics*, 45(6), *572-576*.

Smith, A. (1994). The detection of second-order motion. In: Smith, A., and Snowden, R., (Eds.), *Visual Detection of Motion*. London: Academic Press.

Smith, A., Greenlee, M., Singh, K., Kraemer, F., and Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *Journal of Neuroscience*, 18, 3816-3830.

Snowden, R., and Braddick, O. (1989). Extension of displacement limits in multiple-exposure sequences of apparent motion. *Vision Research*, 29, 1777-1787.

Spenkelink, G.P.J., and Besuijen, J. (1996). Chromaticity contrast, luminance contrast, and legibility of text. *Journal of Society of Information Display*, 4(3), 135-144.

Steinman, B., Steinman, S., and Lehmkuhle, S. (1997). Transient visual attention is dominated by the magnocellular stream. *Vision Research,* 37, 17-23.

Stevens, S.S. (1951). Mathematics, measurement, and psychophysics. In: Stevens, S.S. (Ed.), *Handbook of Experimental Psychology*. New York: John Wiley and Sons, Inc. p.1-49.

Strout, J., Pantle, A., and Mills, S. (1994). An energy model of interframe interval effects in single-step apparent motion. *Vision Research*, 34, 3223-3240.

Swamy, S., Joseph, C., Aravind, A.S., and Vevail, R.J. (2002). Contrast sensitivity in IAF aircrew. *Indian Journal of Aerospace Medicine*, 46, 7-22.

Tanaka, H., Fukada, Y., and Saito, H. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, 62, 642-656.

Tanaka, H., and Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, 62, 626-641.

Task, H.L., and Marasco, P.L. (2004). The impact of changing night-vision goggle spectral response on night-vision imaging system lighting compatibility. Proceedings of SPIE, *Helmet- and Head-Mounted Displays IX*, 5442, 64-72.

Task, H.L. (2001). Night vision goggle visual acuity assessment: results of an interagency test. *Proceedings of SPIE, Helmet- and Head-Mounted Displays VI*, 4361, 130-137.

Task, H.L. (2000). Integrated panoramic night vision goggles fixed-focus eyepieces: selecting a diopter setting. *Proceedings of the 2000 SAFE Association Symposium*, Reno, NV, October 9-11, 2000.

Task, H.L., Riegler, J., and Goodyear, C. (1999). Effects of laser eye protection and aircraft windscreens on visual acuity through night vision goggles. *Proceedings of the 1999 SAFE Association Symposium*, Atlanta, GA, December 6-8.

Task, H.L., and Pinkus, A.R. (1987). Contrast sensitivity and target recognition performance: A lack of correlation. *1987 Society for Information Display International Symposium Digest of Technical Papers, 18,* 127-129.

Task, H.L., and Genco, L.V. (1986). *Contrast Sensitivity Function Measurement Chart and Method.* U.S. Patent No. 4,607,923.

Thorn, F., and Boynton, R.M. (1974). Human binocular summation at absolute threshold. *Vision Research*, 14, 445-458.

Tredici, T.J., and Ivan, D.J. (2008). Ophthalmology in aerospace medicine. In: *Fundamentals of Aerospace medicine*, Davis, Johnson, Tepanek and Fogarty, (Eds.). New York: Lippincott Williams and Wilkins.

Vaina, L., Cowey, A., and Kennedy, D. (1999). Perception of first- and second-order motion: separable physiological mechanisms. *Human Brain Mapping*, 7, 67-77.

van Nes, F., Koenderink, J., Nas, H. and Bouman, M. (1967). Spatiotemporal modulations transfer in the human eye. *Journal of the Optical Society of America*, 57, 1082-1088.

van Santen, J., and Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, 1, 451-473.

van Santen, J., and Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2, 300-321.

van de Pol, C., Greig, J.L., Estrada, A., Bissette, G.M., and Bower, K.S. (2007). Visual and flight performance recovery after PRK or LASIK in helicopter pilots. *Aviation, Space, and Environmental Medicine,* 78, 547-53.

Velger, M. (1998). *Helmet-Mounted Displays and Sights*. Boston, MA: Artech House.

Ververs, P.M., and Wickens, C.D., (1998). Conformal flight path symbology for head-up displays: Defining the distribution of visual attention in three-dimensional space. (University of Illinois Institute of Aviation Tech Report No. ARL-98-5/NASA-98-1) Savoy, Ill: Aviation Research Laboratory.

Wagner, M. (2006a). Effects of context on judgements of distance, area, volume, and angle. In: Wagner, M. (Ed.), *The Geometries of Visual Space*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Wagner, M. (2006b). Factors affecting size constancy. In: Wagner, M. (Ed.), *The Geometries of Visual Space*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Watson, A., and Ahumada, A. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 322-342.

Watson, A., Ahumada, A. and Farrell, J. (1986). Window of visibility: a psychophysical theory of the fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A*, 3, 300-307.

Wenzel, B., Castillo, A., and Baker, G. (2002). Assessment of the virtual environment safe-for-maintenance trainer (VEST). Air Force Research Laboratory, Human Effectiveness Directorate, Warfighter Training Research Division, Mesa, AZ.

Westheimer, G. (2004). Center-surround antagonism in spatial vision: retinal or cortical locus. *Vision Research,* 44(21) 2457-2465.

Wickens, C.D., and Long J. (1995). Object versus space-based models of visual attention: Implication for the design of head-up displays. *Journal of Experimental Psychology: Applied,* 1(3), 179-193.

Wiley, R.W. (1989). Visual acuity and stereopsis with night vision goggles. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 89-9.

Wiley, R.W., and Holly, F.F. (1976). Vision with the AN/PVS-5 night vision goggle. Neuilly-sur-Seine, France: *AGARD*, 191, C7.1-C.12.

Winterbottom, M.D., Patterson, R., Pierce, B.J., and Taylor, A. (2006). Helmet- and Head-Mounted Displays XI: Technologies and Applications. Visual suppression of monocular presented symbology against a fused background in a simulation and training environment. *Proceedings of SPIE*, 6224, 1-10.

Witt, H. (2007). Human-computer interfaces for wearable computers: A systematic approach to development and evaluation. Doctorial Dissertation, University of Bremen.

Wood, R.B., and Howells, P.J. (2007). Head-up displays. In: *Avionics*, Spitzer, C.R., (Ed.). New York: CRC Press.

Woodworth, R.S., and Schologberg, H. (1954). Visual depth perception. In: Woodworth, R.S., and Schlosberg, H. (Eds.), *Experimental Psychology: Revised Edition*. New York: Holt, Pinehart and Winston.

Yamane, N., Miyata, K., Samejima, T., Hiraoka, T., Kiuchi, T., Okamoto, F., Hirohara, Y., Mihashi, T., and Oshika, T. (2004). Ocular higher-order aberrations and contrast sensitivity after conventional laser in situ keratomileusis. *Investigative Ophthalmology and Visual Science*, 45, 3986-90.

Yeh, M., Merlo, J.M., Wickens, C.D., and Brandenburg, D.L. (2003). Head up versus head down: The costs of imprecision, unreliability, and visual clutter on cue effectiveness for display signaling. *Human Factors,* 45(3), 390-407.

Yeh, M., Wickens, C.D., and Seagull, F.J. (1998). Effects of frame of references and viewing condition on attentional issues with helmet-mounted displays (University of Illinois Institute of Aviation Tech Report No. ARL-98-1/ARMY-FED-LAB-98-1) Savoy, Ill: Aviation Research Laboratory.

Zalevski, A., Meehan, J.W., and Hughes, P.K. (2001). *Size Estimation with Night Vision Goggles.* Fishermans Bend, Victoria, Australia: Department of Defence, Defence Science and Technology Organisation (DSTO), Aeronautical and Maritime Research Laboratory. AR-011-775

Zhang, X., and Farrell, J.E. (2003). Sequential color breakup measured with induced saccades. *Proceedings of SPIE, Human Vision and Electronic Imaging VIII*, 5007, 210-217.

11 AUDITORY PERCEPTION AND COGNITIVE PERFORMANCE

Angélique A. Scharine
Kara D. Cave
Tomasz R. Letowski

Audition

Audition is the act of hearing a sound in response to acoustic waves or mechanical vibrations acting on a body. Sound also may result from direct electrical stimulation of the nervous system. The physical stimuli that are, or may become, the sources of sound are called *auditory stimuli*.

The human response to the presence of auditory stimulus and its basic physical characteristics of sound intensity, frequency, and duration is called *auditory sensation*. The three basic auditory sensations are loudness, pitch, and perceived duration, but there are many others. Auditory sensation forms the basis for discrimination between two or more sounds and may lead to some forms of sound classification (e.g., labeling sounds as pleasant or unpleasant). However, auditory sensation does not involve sound recognition, which requires a higher level of cognitive processing of the auditory stimuli. This higher level processing forms a conceptual interpretation of the auditory stimulus and is referred to as *auditory perception*. Auditory perception involves association with previous experience and depends on the adaptation to the environment and expected utility of the observation. Depending on the level of cognitive processing, auditory perception may involve processes of sound classification, e.g., on speech and non-speech sounds, sound recognition, or sound identification. More complex cognitive processing also may include acts of reasoning, selection, mental synthesis, and concept building involving auditory stimuli but extends beyond the realm of audition.

The study of audition is called *psychoacoustics* (psychological acoustics). Psychoacoustics falls within the domain of cognitive psychophysics, which is the study of the relationship between the physical world and its mental interpretation. Cognitive psychophysics is an interdisciplinary field that integrates classical psychophysics (Fechner, 1860), which deals with the relationships between physical stimuli and sensory response (sensation), and with elements of cognitive psychology, which involve interpretation of acting stimuli (perception). In general, cognitive psychophysics is concerned with how living organisms respond to the surrounding environment (Stevens, 1972b). For the above reasons, Neuhoff (2004) refers to modern psychoacoustics as ecological psychoacoustics.

In general, all content of our experience can be ordered by quantity, quality, relation, and modality (Kant, 1781). These experiences are reflected in perceptual thresholds, various forms of comparative judgments, magnitude estimation, emotional judgments, and scaling. These characteristics define the realm of psychoacoustics and, more generally, psychophysics. Various types of cognitive measurements and methodological issues addressing psychophysical relationships are described in Chapter 15, *Cognitive Factors in Helmet-Mounted Displays*, and are not repeated here. The current chapter presents psychoacoustic relationships and builds upon the information on the anatomy and physiology of the auditory system presented in Chapter 8, *Basic Anatomy of the Hearing System*, and Chapter 9, *Auditory Function*. It describes a variety of auditory cues and metrics that are used to derive an understanding of the surrounding acoustic space and the sound sources operating within its limits. Understanding how a particular sound is likely to be perceived in a particular environment is necessary for the design of effective auditory signals and to minimize the effects of environmental noise and distracters on performance of audio helmet-mounted displays (HMDs). Psychoacoustics provides the basic conceptual framework and measurement tools (thresholds and scales) for the discussion and understanding of these effects.

Sound Pressure and Sound Pressure Level

The main physical quantity that elicits auditory response is time-varying sound pressure. The other quantities are time-varying force (bone conduction hearing) and time-varying (alternating current [AC]) voltage (electric hearing). The unit of sound pressure is the Pascal (Pa), which is equal to a Newton/meter^2 (N/m^2), and the range of sound pressures that can be heard by humans extends from about 10^{-5} Pa to 10^2 Pa. The large range of values needed to describe the full range of audible sound pressure makes the use of Pascals, or other similar linear units, very cumbersome. In addition, human auditory perception is far from linear. Human perception is relative by nature and logarithmic in general, i.e., linear changes in the amount of stimulation cause logarithmic changes in human perception (Emanuel, Letowski and Letowski, 2009). Therefore, sound pressure frequently is expressed in psychoacoustics on a logarithmic scale known as the *decibel scale* from the name of its unit, the *decibel* (dB). The decibel scale has a much smaller range than the sound pressure scale and more accurately represents human reaction to sound. Sound pressure expressed in decibels is called sound pressure level. Sound pressure level (SPL) and sound pressure (p) are related by the equation:

$$SPL\,(\text{dB}) = 20\log\left(\frac{p}{p_o}\right) \quad [\text{dB SPL}] \qquad \text{Equation 11-1}$$

where p_o is the reference sound pressure value and equals 20×10^{-6} Pa. For example, a sound pressure (p) of 1 Pa corresponds to 94 dB SPL, and the whole range of audible sound pressures extends across about 140 dB SPL. An SPL of 1 dB corresponds to a sound pressure increase of about 1.122 times (~12%).

When dealing with complex continuous sounds, it is frequently more convenient to use energy-related magnitudes such as sound intensity (I) or sound intensity level (SIL) rather than sound pressure and sound pressure level. Such an approach allows one to refrain from the concept of phase that complicates physical analysis and has limited usefulness for many aspects of auditory perception. Sound intensity is the density of sound energy over an area, is expressed in units of Watts/meter^2 (W/m^2), and for a plane traveling sound wave, $I \sim p^2$. Therefore, the relation between sound pressure level (dB SPL) and sound intensity level (dB SIL) can be written as:

$$SPL\,(\text{dB}) = 20\log\left(\frac{p}{p_o}\right) = 10\log\left(\frac{I}{I_o}\right) = SIL\,(dB) \qquad \text{Equation 11-2}$$

where I_o is the reference sound intensity value of 10^{-12} W/m^2. I_o is the sound intensity produced by the sound pressure equal to p_o. This means that both values refer to the same sound, and the scale of sound pressure level in dB SPL is identical to the scale of sound intensity level in dB SIL. For that reason, the names *sound pressure level* and *sound intensity* are used interchangeably in this chapter and all the graphs labeled sound pressure level (dB SPL) would be identical if the label was replaced by sound intensity level (dB SIL).

Threshold of Hearing

Sensitive hearing is an important listening ability needed for human communication, safety, and sound assessment. An auditory stimulus arriving at the auditory system needs to exceed a certain level of stimulation to cause the temporary changes in the state of the system that result in the sensation and perception of sound. The minimum level of stimulation required to evoke a physiologic response from the auditory system is called the *threshold of hearing* or *detection threshold* and depends on the frequency, duration, and spectral complexity of the stimulus. Thus, the threshold value is the lowest intensity level (e.g., sound pressure level, bone conduction force level, electric voltage level) for which a particular auditory stimulus can be detected. When the sound is heard in quiet, the detection threshold is called the *absolute detection threshold* (absolute threshold of hearing), and when is presented together with other sounds, the detection threshold is referred to as the *masked detection threshold* (masked threshold of hearing).

The term threshold of hearing implies the existence of a discrete point along the intensity continuum of a particular stimulus above which a person is able to detect the presence of the stimulus. However, an organism's sensitivity to sensory stimuli tends to fluctuate, and several measures of the threshold value must be averaged in order to arrive at an accurate estimation of the threshold. Therefore, the threshold of hearing is usually defined as the sound intensity level at which a listener is capable of detecting the presence of the stimulus in a certain percentage, usually 50%, of cases.

Normal daily variability of the threshold of hearing can be assumed to be 6 dB or less. For example, Delany (1970) and Robinson (1986) used a Bekesy tracing procedure (Brunt, 1985) and supra-aural earphones to assess within-subject variability of the hearing threshold in otologically normal listeners and both reported an average standard deviation in the order of 5 dB at 4000 Hertz (Hz) for repeated measures of the threshold of hearing on the same person. Henry et al. (2001) used insert earphones ER-4B (see Chapter 5, *Audio Head Mounted Displays*) and the ascending method of limits with 1 dB steps and reported average within subject standard deviations between 1.9 dB and 5.3 dB depending on the stimulus frequency.

In this context, it is useful to differentiate between the physiological threshold defined by the inherent physiological abilities of the listener and the cognitive threshold limited by the listener's familiarity with the stimuli, interest in and attention to the task, experience with existing set of circumstances, and the numerous procedural effects in eliciting the listener's response (Seashore, 1899; Letowski, 1985). The cognitive thresholds can be made close to the physiological thresholds by an appropriate amount and type of training (Letowski, 1985; Letowski and Amrein, 2005); but the difference between potential physiological threshold and observed cognitive threshold always has to be taken into account when discussing any specific threshold data.

The need for a statistical approach to the threshold of hearing also exists for normative thresholds for specific populations because people differ in both their overall sensitivity to sound and the shape of the threshold of hearing as a function of frequency. For example, inter-subject (between-subject) standard deviations ranging from 3 dB to 6 dB were reported for low and middle frequencies in a number of studies (Moller and Pedersen, 2004) and the inter-subject data variability tends to increase with stimulus frequency. Thus, due to both individual (intra-subject) and group (inter-subject) variability of the threshold of hearing, human sensitivity to auditory stimulation needs to be defined in terms of a statistical distribution with certain measures of central tendency and dispersion (variability).

Air conduction threshold

The range of frequencies heard by humans through air conduction extends from about 20 Hz to 20 kHz and may possibly include even lower frequencies if stimulus intensities are sufficiently high (Møller and Pedersen, 2004). The average human threshold of hearing, standardized by the International Organization for Standardization (ISO, 2005) for people age 18 to 25 years with normal hearing, varies by as much as 90 dB as a function of the stimulus frequency and is shown in Figure 11-1. The specific threshold values are listed in Table 11-1.

The threshold curves in Figure 11-1 and numbers in Table 11-1 represent average binaural thresholds of hearing in a free sound field and a diffuse sound field. A free sound field is an acoustic field, free of reflections and scattering in which sound waves arrive at the listener's ears from only one direction identified by the position of the sound source relative to the main axis of the listener. A diffuse sound field is a sound field in which the same sound wave arrives at the listener more or less simultaneously from all directions with equal probability and level. The free-field thresholds were measured for pure tones with the subject directly facing the source of sound (frontal incidence). The diffuse-field thresholds were measured for one-third-octave bands of noise. In both cases, the threshold sound pressure level was measured at the position corresponding to the center of the listener's head with the listener absent.

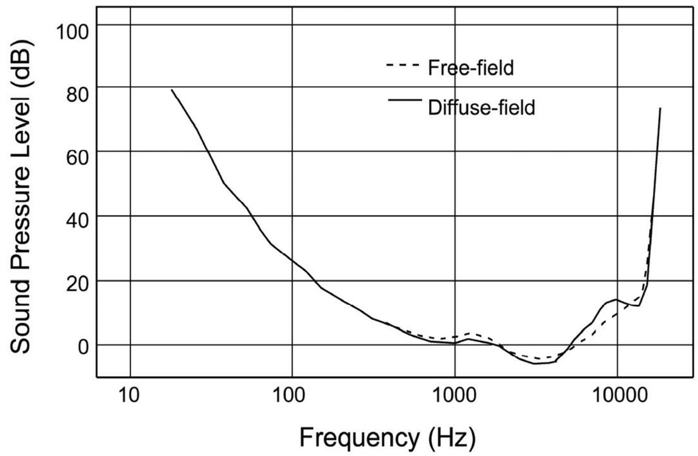


Figure 11-1. Binaural hearing threshold of hearing in a free field (frontal incidence) and in a diffuse field as a function of frequency (adapted from ISO, 2005).

Table 11-1.
Reference thresholds of hearing for free-field listening (frontal incidence) and diffuse-field listening in dB SPL (re: 20 μPa) (ISO, 2005).

Frequency (Hz)	Free-field listening (frontal incidence) (dB SPL)	Diffuse-field listening (dB SPL)
20	78.5	78.5
25	68.7	68.7
31.5	59.5	59.5
40	51.1	51.1
50	44.0	44.0
63	37.5	37.5
80	31.5	31.5
100	26.5	26.5
125	22.1	22.1
160	17.9	17.9
200	14.4	14.4
250	11.4	11.4
315	8.6	8.4
400	6.2	5.8
500	4.4	3.8
630	3.0	2.1
750	2.4	1.2
800	2.2	1.0
1000	2.4	0.8
1250	3.5	1.9
1500	2.4	1.0

Table 11-1. (Cont.)
Reference thresholds of hearing for free-field listening (frontal incidence) and diffuse-field listening in dB SPL (re: 20 µPa) (ISO, 2005).

Frequency (Hz)	Free-field listening (frontal incidence) (dB SPL)	Diffuse-field listening (dB SPL)
1600	1.7	0.5
2000	-1.3	-1.5
2500	-4.2	-3.1
3000	-5.8	-4.0
3150	-6.0	-4.0
4000	-5.4	-3.8
5000	-1.5	-1.8
6000	4.3	1.4
6300	6.0	2.5
8000	12.6	6.8
9000	13.9	8.4
10000	13.9	9.8
11200	13.0	11.5
12500	12.3	14.4
14000	18.4	23.2
16000	40.2	43.7
18000	73.2	--

The binaural thresholds of hearing shown in Figure 11-1 and Table 11-1 are approximately 2 dB lower than the corresponding monaural thresholds if both ears have similar sensitivity (Anderson and Whittle, 1971; Killion, 1978; Moore, 1997). This difference applies to pure tones of various frequencies as well as to speech, music, and other complex stimuli presented under the same monaural and binaural conditions.

Low frequency stimuli are felt as a rumble (Sekuler and Blake, 1994) and require a relatively high sound level to be detected. Their audibility does not vary much among individuals and depends primarily on the mechanical properties of the ear. The audibility of low frequency stimuli improves with increasing frequency at an average rate of about 12 dB/octave and typically decreases with age, from 20 to 70 years of age, by 10 dB or less for frequencies below 500 Hz (ISO, 2000). The audibility of stimuli with frequencies in the upper end of the frequency range varies quite a bit with individuals and decreases substantially with age (Stelmachowicz et al., 1989). The typical changes in the threshold of hearing with age, from 20 to 70 years of age, at frequencies above 8,000 Hz, exceed 60 dB in normally hearing population.

As demonstrated in Figure 11-1, the auditory system is especially sensitive to sound energy in the 1.5 to 6 kHz range with the most sensitive region being in the 3.0 to 4.0 kHz range (Moore, 1997). The high sensitivity of the auditory system in this frequency range results from acoustic resonances of the ear canal and concha described in Chapter 9, *Auditory Function*. The normal hearing threshold level in its most sensitive range is about -10 dB re 2 x 10^{-5} Pa (2 x 10^{-4} µbar)[1] and the amplitude of the tympanic membrane displacement is about 10^{-9} centimeter (cm), i.e., not much larger than the amplitude of the random motion of molecules in solution (Licklider, 1951).

[1] When reporting the relative intensity of a sound, it is important to not only say “dB” but to also add the reference level. This is often written as “dB re 20 µPa” for sounds in air that are measured relative (re) to 20 µPa.

Low hearing sensitivity at low frequencies is primarily due to poor transmission of low frequency energy by the mechanical systems of the outer and middle ears and limited mobility of the outer hair cells in the low frequency range (Moore, Glasberg and Bear, 1997). The presence of the second mechanism is probably due to the high level of low frequency internal body noise, such as that caused by the blood flow, which normally should not be heard.

When describing the threshold of hearing, it is important to consider not only whether the threshold is monaural or binaural but also how the sound is presented and where the level of arriving stimulus is measured. Two specific types of hearing threshold, the *minimum audible field* threshold and *minimum audible pressure* threshold are of special importance. The minimum audible field (MAF) threshold refers to the threshold of hearing for acoustic waves arriving at the ear in a free-field environment from a distal sound source, e.g., a loudspeaker. The minimum audible pressure (MAP) threshold refers to the threshold of hearing from a stimulus arriving from an earphone occluding the ear canal.

The difference between the MAF and MAP thresholds of hearing is illustrated in Figure 11-2. The average difference between both thresholds is in the order of 6 dB to 10 dB and has been sometimes referred to in the literature as the "missing 6 dB." It should be noted that especially large differences between the MAF and MAP thresholds in the 1.5 to 4 kHz frequency region. This difference is the effect of resonance properties of the ear canal and concha.

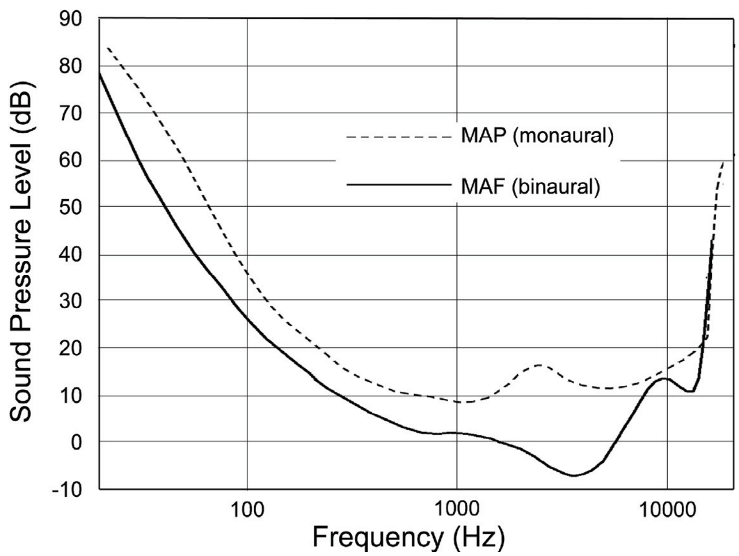


Figure 11-2. Comparison of the MAF (ISO, 2003) and MAP (Killion, 1978) threshold of hearings. The reference points for both measurements are the center of the listener's head with the listener absent (MAF) and a point close to the listener's tympanic membrane (MAP).

The main reason that the MAF and MAP thresholds of hearing, such as shown in Figure 11-2, differ in their values is the differences in the actual point of reference where the sound pressure is measured. In practice, the reference points for the MAF and MAP thresholds do not only differ from each other but they also differ within each threshold category. The typical reference points for MAF threshold are the point at the entrance to the ear canal at the tragus and the point representing the position of the center of the person's head when the person is absent. The MAP measurements are frequently made with a small probe tube microphone where the tip of the probe is located either close to the tympanic membrane, at various points along the ear canal, or in front of the earphone grill. The differences between thresholds obtained for sounds presented in the open field versus those presented through earphones are large due to the reflective properties of the human head and torso and different sound amplification by resonance properties of the ear canal and concha (Chapter 9, *Auditory Function*). In addition, the MAP values for the threshold of hearing are frequently estimated in various acoustic couplers and ear simulators in which the pressure often bears little resemblance to the actual pressure at the listener's tympanic membrane (Killion, 1978). In such cases, specific "reference equivalent threshold sound pressure levels" (RETSPLs) are established for various earphone-coupler combinations to reference threshold of hearing pressure to voltage value applied to the earphone. The RETSPL values are internationally standardized for several reference earphones including supra-aural earphones (e.g., Telephonics TDH-39 and Beyer DT-48), circumaural earphones (e.g., Sennheiser HD-200), and insert earphones (e.g., Etymotic Research ER-3). Each of the set of RETSPLs is referenced to a specific standardized acoustic coupler and is only valid when the appropriate coupler and appropriate calibration method are used.

Another factor contributing to the difference between MAF and MAP thresholds of hearing is that MAF thresholds are usually determined for binaural listening, while MAP thresholds are usually reported for the monaural condition. In addition, occluding the ear canal by an earphone was reported to cause an amplification of low frequency physiologic noise (e.g., blood flow noise) by the closed cavity of the ear canal and elevation of MAP threshold at low frequencies (Anderson and Whittle, 1971; Block, Killion and Tillman, 2004; Brodgen and Miller, 1947; Killion, 1978). The occluded ear canal also has different resonance modes than the open canal. Similarly, measurements of the MAF threshold in less than ideal (anechoic) conditions may cause room reflections to affect the real threshold values.

The threshold values discussed above have been determined for pure tones and are primarily used for clinical applications. However, for many practical field applications, room or transducer (e.g., audio HMD) frequency response considerations, and special population testing it is important to determine the threshold of hearing for speech signals and other complex acoustic stimuli such as narrow bands of noise and filtered sound effects. One class of such signals is 2% to 5% frequency modulated (FM) tones, called warble tones, which are used commonly in sound field audiometry. They result in the thresholds of hearing similar to the pure-tone thresholds but are less dependent on the acoustic conditions of a room. They are also the signal of choice for high frequency audiometry (Tang and Ltowski, 2007).

In the case of speech signals, there are two thresholds of hearing for speech that are of interest for practical applications: the threshold of speech intelligibility (speech recognition threshold, speech reception threshold) and the threshold of speech audibility (speech awareness threshold, speech detection threshold). The normative speech recognition threshold for English spondee (two-syllable) words is 14.5 dB SPL for binaural listening in a free sound field with the sound presented at 0° incidence (American National Standards Institute [ANSI], 2004). The speech awareness threshold (SAT) is approximately 8 to 9 dB lower (Dobie and van Hemel, 2004; Sammeth et al., 1989).

Auditory thresholds for narrow-band noises have been reported by Garstecki and Bode (1976), Mitrinowicz and Letowski (1966), Sanders and Joey (1970), and others. In all of these studies, the reported thresholds correlated very well with pure tone thresholds and were usually within ±3 dB of each other. Mitrinowicz and Letowski (1966) observed that the relation between the narrow-band noise thresholds and the pure tone thresholds was mitigated by the relation between the width of the noise band and the width of the corresponding critical band (to be discussed further in this chapter). Zarcoff (1958) also reported a good correlation between narrow-band noise thresholds at mid-frequencies and the speech recognition thresholds.

Environmental sound research is still in its beginning stages (Gygi and Shafiro, 2007), and there are few studies reporting thresholds of hearing for various environmental and man-made sounds. Many early reports dealing with environmental sounds are qualitative rather than quantitative in nature and the listening conditions used in these studies are not well described. Yet, they provide much information on human ability to differentiate various sounds, informational and emotional meaning of the sounds, and provide the case for further, more detailed studies. The few quantitative studies report threshold values that vary across more than 30 dB depending on the sound, listener, listening environment, and listening condition. Some of the more important early and quantitative reports related to detection and recognition of environmental sounds include: Abouchacra, Letowski and Gothie (2006); Ballas (1993); Ballas, Dick and Groshek (1987); Ballas and Howard (1987); Ballas and Barnes (1988); Fidell and Bishop (1974); Gygi (2001); Gygi, Kidd and Watson, (2004); Gygi and Shafiro (2007); Price and

Hodge (1976); Price, Kalb and Garinther (1989); and Shafiro and Gygi (2004; 2007). There are also reports on detection and recognition of octave-band filtered environmental sounds for warning signal and auditory icon applications (Myers et al., 1996; Abouchacra and Letowski, 1999).

Bone conduction threshold

Figures 11-1 and 11-2 refer to the air-conducted threshold of hearing. However, sound can also be heard through bone conduction transmission either directly, through contact with a vibrating object or indirectly, by picking up vibrations in the environment (Henry and Letowski, 2007). In real operational environments, bone conducted sound transmission is likely to occur through the use of bone conduction communication devices (direct stimulation) or from very loud sounds in the environment (indirect stimulation). In the former case, the effectiveness of bone conduction depends on the location of the vibrator on the head and the quality of the contact between the vibrator and the skin of the head (McBride, Letowski and Tran, 2005; 2008). In the latter case, bone conducted sounds may be masked by stronger air conducted sounds except for the cases when high-attenuation hearing protection is used (see Chapter 9, *Auditory Function*).

The threshold of hearing for bone conduction is defined as the smallest value of mechanical force (force threshold) or acceleration (acceleration threshold) applied to the skull resulting in an auditory sensation. Table 11-2 lists the force threshold values for bone conduction threshold as given in ANSI) standard S3.6 (ANSI, 2004) for a vibrator placed on the mastoid and on the forehead. Threshold values vary with frequency and are lowest in the 1 to 4 kHz region, similarly as for the air-conduction threshold.

Table 11-2.
Normal monaural force hearing thresholds for bone-conducted sounds at different frequencies for a B-71 vibrator placed on the mastoid and at the forehead (ANSI, 2004).

Frequency (Hz)	Mastoid Location (dB re 1μN)	Forehead Location (dB re 1μN)
250	67.0	79.0
315	64.0	76.5
400	61.0	74.5
500	58.0	72.0
630	52.5	66.0
750	48.5	61.5
800	47.0	59.0
1000	42.5	51.0
1250	39.0	49.0
1500	36.5	47.5
1600	35.5	46.5
2000	31.0	42.5
2500	29.5	41.5
3000	30.0	42.0
3150	31.0	42.5
4000	35.5	43.5
5000	40.0	51.0
6000	40.0	51.0
6300	40.0	50.0
8000	40.0	50.0

Similarly to air conduction thresholds, bone conduction thresholds may be measured using an artificial load, in this case a mechanical load such as an artificial mastoid, in lieu of the human head. In such cases, the bone conduction threshold is referenced by “reference equivalent threshold force levels” (RETFLs), which is the acoustic force needed for threshold sensation when applied to the artificial load.

At present the only well established direct-stimulation bone conduction thresholds are for pure tones. The only other bone conduction thresholds that were published are the thresholds for octave-band filtered sound effects that were published together with the corresponding air conduction thresholds by Abouchacra and Letowski (1999).

In the case of bone conduction stimulation by impinging sound waves, such sound waves need to be 40 to 50 dB more intense than those causing the same sensation though the air conduction pathways. More information on bone conduction mechanisms and the use of bone conduction hearing in audio HMDs is included in Chapter 9, *Auditory Function*, and Chapter 5, *Audio Helmet Mounted Displays*, respectively.

Threshold of Pain

The threshold of hearing is an example of the basic class of perceptual thresholds called the detection thresholds or absolute thresholds, which separate effective from ineffective stimuli. The other type of perceptual threshold is the terminal threshold. The terminal threshold defines the greatest amount of stimulation that can be experienced in a specific manner before it causes another form of reaction. Examples of auditory terminal thresholds are the threshold of discomfort, also referred to as the loudness discomfort level (LDL), and the threshold of pain.

The threshold of discomfort represents the highest sound intensity that is not uncomfortable or annoying to the listener during prolonged listening. According to Gardner (1964), the threshold of discomfort is almost independent of background noise level (in 30 to 70 dB SPL range) and exceeds 80 dB SPL. It depends on the listener, type of sound (tone, speech, noise band), and frequency content of the sound and varies in 80 to 100 dB SPL range for speech signals (Denenberg and Altshuler, 1976; Dirks and Kamm, 1976; Keith, 1977; Morgan et al., 1979). The typical difference between the most comfortable listening level and threshold of discomfort is about 15 dB for pure tones and speech (Dirks and Kamm, 1976) and about 25 dB for noises (Sammeth, Birman and Hecox, 1989).

The threshold of pain represents the highest level of sound that can be heard without producing a pain. The threshold of pain is practically independent of frequency and equals about 130 to 140 dB SPL. The nature of pain, however, varies with frequency. At low frequencies, people experience dull pain and some amount of dizziness, which suggests the excitation of the semicircular canals. At high frequencies, the sensation resembles the stinging of a needle.

There are reports indicating that tones of very low frequencies below 20 Hz, called *infrasound*, can be heard by some people at very high intensity levels exceeding 115 dB (Møller and Pedersen, 2004). In general, however, such tones having sufficiently high sound intensity levels are not heard but immediately felt causing disorientation, pain, and feeling of pressure on the chest (Gavreau, 1966, 1968). A similar situation may exist for high frequency tones exceeding 20 kHz, called *ultrasound*, although there are numerous reports indicating that people can hear ultrasound stimuli if they are applied through bone conduction (Lenhardt et al., 1991).

Area of Hearing

If the energy of an auditory stimulus falls within the sensory limits of the auditory system we can hear a sound, i.e., receive an auditory impression caused by the energy of the signal. The range of audible sound between the threshold of hearing and the threshold of pain displayed in frequency (abscissa) and sound pressure level (ordinate) coordinates is called the area of hearing. The area of hearing, together with smaller areas of music and speech sounds, is shown graphically in Figure 11-3.

The data presented in Table 11-1 and Figures 11-1, 11-2, and 11-3, show that the threshold of hearing changes considerably across the whole range of audible frequencies. Therefore, in some cases it is convenient to describe actual stimuli in terms of the level above the threshold of hearing rather than sound pressure level. This level is called the hearing level (HL), when referred to the average threshold of hearing for a population, or the sensation level (SL), when referred to the hearing threshold of a specific person. The average hearing threshold level (0 dB HL) for a given population is called in the literature the *reference hearing threshold level* or the *audiometric zero level*. For example, the level of 60 dB SPL in Figure 11-3 would correspond to 25 dB HL and 60 dB HL for a 100 Hz and a 1000 Hz tone, respectively. Keep in mind that these are the approximate values due to the conceptual form of Figure 11-3. The more exact numbers can be found from Figure 11-6 and the relation between dB SPL values and 0 dB HL values may be found is Table 11-1 (air conduction threshold) and Table 11-2 (bone conduction threshold).

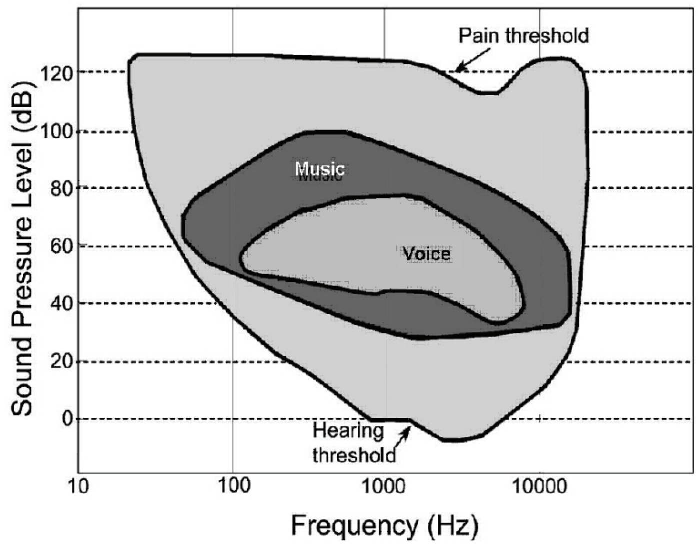


Figure 11-3. Area of hearing (light gray) together with areas of music (black), and speech (dark gray).

As shown in Figure 11-3, the dynamic range of human hearing extends from approximately -10 dB SPL to 130 dB SPL. To make the above numbers more practical, the range of sound intensity levels of the natural and man-made sounds that are generated in the environment is shown in Table 11-3.

The frequency range of hearing shown in Figure 11-3 extends from about 20 Hz (or even less) to 20 kHz. This range is slightly larger than the range of sounds of music but much larger than the range of speech sounds that extends from about 200 Hz to approximately 8 kHz. However, it is important to stress that hearing sensitivity, especially in the high frequency region, declines with age (Galton, 1883; Robinson and Dadson, 1957), exposure to noise, and use of ototoxic drugs, and the standardized threshold levels published in literature refer typically to the average threshold of hearing in young people, i.e., age 18 to 25 years.

The threshold of hearing is not the same for all populations and depends on both gender and ethnicity. Murphy, Themann and Stephenson (2006) evaluated hearing threshold in more than 5000 U.S. adults, age 20 to 69 years, and found: women have on average better hearing than men; non-Hispanic blacks have the best hearing threshold; and non-Hispanic whites have the worst among all ethnic groups evaluated in the study. In addition, Cassidy and Ditty (2001), Corso (1957; 1963) and Murphy and Gates (1999) reported that women of all ages have better hearing than men at frequencies above 2000 Hz, but with aging women have poorer capacity to hear lower frequencies than do men. This means that the pattern of hearing loss with aging for women and men is not the same. The data regarding changes in the threshold of hearing with age can be found in ISO standard 7029 (ISO, 2000).

From an operational point of view it is important to compare the frequency range of human hearing with the frequency ranges of other species. At its high frequency end human hearing extends above 10 kHz, as does the hearing of all other mammals with a few exceptions (e.g., subterranean mammals such as blind mole rat) (Heffner and Heffner, 1993). Birds do not hear sounds higher than 10 kHz and amphibians, fish, and reptiles do not generally hear sounds higher than 5 kHz (Heffner and Heffner, 1998). Dogs hear frequencies up to about 45 kHz and some bats and porpoises can hear sounds beyond 100 kHz. In mammals, smaller head size generally is correlated with better high frequency hearing of the mammal (Masterton, Heffner and Ravizza, 1967). This relationship is important for species survival since small head size produces a smaller acoustic shadow and good high frequency hearing is needed for effective sound localization and effective hunting (Heffner and Heffner, 2003). The importance of high frequency hearing for sound localization has been addressed previously in Chapter 9, *Auditory Function*, and will be further discussed in following sections in this chapter.

Table 11-3.
Sound intensity levels of some environmental and man-made sounds.
Adapted from Emanuel and Letowski (2009).

Sound Level (dB SPL)	Examples of Sounds
0	Quietest 1kHz tone heard by young humans with good hearing. Mosquito at 3 meters (9.8 feet).
10	Human breathing at 3 meters (9.8 feet). Wristwatch ticking at 1 meter (3.28 feet).
20	Rustling of leaves at 3 meters (9.8 feet). Whisper at 2 meters (6.6 feet). Recording studio noise level.
30	Nighttime in a desert. Quiet public library. Grand Canyon at night. Whisper at the ear.
40	Quiet office. Loud whisper. Suburban street (no traffic). Wind in trees.
50	Average office. Classroom. External air conditioning unit at 30 meters (98 feet).
60	Loud office. Conversational speech at 1 meter (3.28 feet). Bird call at 3 meters (9.8 feet).
70	Inside passenger car (65 mph). Garbage disposal at 1 meter (3.28 feet). [1 μbar = 74 dB SPL]
80	Vacuum cleaner at 1 meter (3.28 feet). Noisy urban street. Power lawn mower at 3 meters (9.8 feet).
90	Heavy track (55 mph) at 1 meter. Inside HMMWV (50 mph). [1 pascal = 94 dB SPL]
100	Pneumatic drill at 2 meters (6.6 feet). Chain saw at 1 meter (3.28 feet). Disco music (dancing floor).
110	Symphonic orchestra (tutti; forte fortissimo) at 5 meters (16.4 feet). Inside M1 tank (20 mph).
120	Jet airplane taking off at 100 meters (328 feet). Threshold of pain. [1W/m^2 = 120 dB SPL]
130	Rock band at 1 meter. Civil defense siren at 30 meters (98 feet). [1 mbar = 134 dB SPL]
140	Aircraft carrier flight deck Human eyes begin to vibrate making vision blurry.
150	Jet engine at 30 meters. Formula I race car at 10 meters (32.8 feet).
160	M-16 gunshot at shooter's ear (157 dB SPL). Windows break at about 160 dB SPL
170	Explosion (1 tone TNT) at 100 meters (328 feet). Direct thunder. [1 psi = 170.75 dB SPL]
180	Explosion (0.5 kg TNT) at 5 meters (16.4 feet). Hiroshima atomic bomb explosion at 1.5 kilometers (0.93 miles).
190	Ear drums rupture at about 190 dB SPL. [1 bar = 194 dB SPL]
200	Bomb explosion (25 kg TNT) at 3 meters (9.8 feet). Humans die from sound at 200 dB SPL
210	Space shuttle taking off at 20 meters (65.6 feet). Sonic boom at 100 meters (328 feet).
220	Saturn 5 rocket take off at 10 meters (32.8 feet).

At the low frequency end, human hearing is relatively extensive and very few species (e.g., elephants and cattle) hear lower frequencies than humans. It is noteworthy that the low frequency limit of the mammal hearing is either lower than 125 Hz or higher than 500 Hz. Only very few species that have been reported to have a low-frequency limit of hearing in 125 to 500 Hz range do not fit this dichotomy (Heffner and Heffner, 2003). This is an important finding because its supports the existence of dual mechanisms of pitch (frequency) perception in mammals (including humans), i.e., place coding and temporal coding (see Chapter 9, *Auditory Function*). It has been argued that temporal coding operates up to less than 300 Hz (Flanagan and Guttman, 1960; Shannon, 1983) while place coding fails to account for good frequency resolution at low frequencies. Thus it seems possible that the mammals that do not hear below 500 Hz may use only place mechanism for pitch coding while the mammals that hear below 125 Hz may be using both place and temporal coding for pitch perception. The list of frequency ranges of selected mammals is given in Table 11-4.

Table 11-4.
Approximate hearing ranges of various species (Fay, 1988; Warfield, 1973).

Species	Low Frequency Limit (Hz)	High Frequency Limit (Hz)
Beluga whale	1,000	120,000
Bat	2,000	110,000
Bullfrog	100	3,000
Cat	45	64,000
Catfish	50	4,000
Chicken	125	2000
Cow	25	35,000
Dog	65	45,000
Elephant	16	12,000
Horse	55	33,000
Owl	125	12,000

Auditory Discrimination

The third, and last, class of perceptual thresholds is the differential thresholds. The differential threshold is the smallest change in the physical stimulus that can be detected by a sensory organ. Such threshold is frequently referred to as a *just noticeable difference* (jnd) or *difference limen* (DL).[2] The size of the differential threshold increases with the size of the stimulus and this relationship is known as Weber's Law, named after Erich Maria Weber who formulated it in 1834 (Weber, 1834). Weber's Law states that the smallest noticeable change in stimulus magnitude (ΔI) is always a constant fraction of the stimulus magnitude (I):

$$\frac{\Delta I}{I} = c = const. \qquad \text{Equation 11-3}$$

where *c* is a constant called the Weber fraction (Weber, 1834). This expression leads to a logarithmic function describing the dependence of noticeable change in stimulus magnitude on stimulus magnitude.

When Weber's Law is applied to the stimulus intensity, it holds across a large range of intensities except for those intensities close to the threshold of detection. At the low levels of stimulation, the actual differential thresholds are larger than predicted by Weber's Law due to the presence of internal noise. This effect has been

[2] *Difference limen* (as the *jnd*) is the smallest change in stimulation that an observer can detect.

termed the "near miss to Weber's law" (McGill and Goldberg, 1968). Differential thresholds for stimulus characteristics other than stimulus intensity do not demonstrate any notable departure from Weber's Law.

Differential thresholds can be further classified as single-event (step) thresholds and modulation thresholds depending on the nature of the change. Single-event thresholds are the thresholds of detection of a single change in signal property. Modulation thresholds are the thresholds of detection of signal modulation, that is, periodic changes in signal property. If the single-event stimulus change has the form of a step without any interstimulus pause, the differential threshold is generally smaller than the modulation detection thresholds (Letowski, 1982). The tendency of the auditory system to smooth out (ignore) small frequently repeated changes in an auditory stimulus can be, among others, observed in the decrease of audibility of modulation with increasing modulation rate. Since single-event and modulation thresholds correspond to different natural phenomena and may need to be differentiated for some practical applications, it is important to know the specific methodology of the data collection that lead to specific published values of DLs.

After the auditory stimulus is detected, it can be discriminated from other stimuli on the basis of a number of auditory sensations that can be treated as the attributes of an internal image of the stimulus. The three basic auditory sensations are loudness, pitch, and perceived duration. These sensations are highly correlated with the physical properties of sound intensity, sound frequency, and sound duration, respectively, but they are affected by the other two physical properties of sound as well, e.g., loudness does not only depend on sound intensity but also on sound frequency and sound duration. However, when all physical properties of sound except for the property of interest are held constant, the smallest changes in sound intensity, frequency, or duration can be detected using the sensations of loudness, pitch, and perceived duration. These DLs, when obtained, can be used to measure the acuity of the hearing system with respect to a specific physical variable, e.g., intensity resolution, frequency (spectral) resolution, and temporal resolution, or to determine the smallest change in the stimulus that has practical value for signal and system developers.

Intensity discrimination

The two most frequently discussed differential thresholds in psychoacoustics are the differential threshold for sound intensity (intensity DL) and the differential threshold for sound frequency (frequency DL). The DL for sound intensity is the smallest change in sound intensity level that is required to notice a change in sound loudness.

The DL for sound intensity is typically about 0.5 to 1.0 dB within a wide range of intensities (greater than 20 dB above the threshold) and across many types of stimuli (Chochole and Krutel, 1968; Letowski and Rakowski, 1971; Riesz, 1928). This means that Weber's law holds for both simple and complex sounds and applies to both quiet and natural environments. The intensity DL can be as small as 0.2 dB for pure tones in quiet and sound levels exceeding 50 dB SPL (Pollack, 1954) and reaches up to about 3 dB for natural sounds listened to in natural environment. An example of the relation between the DL for sound intensity and the intensity of the pure tone stimulus is shown in Figure 11-4.

The exponential character of the intensity discrimination function can be approximated for pure tones by:

$$\frac{\Delta p}{p} = \frac{1}{4}\sqrt[6]{\frac{p_o}{p}}\,, \qquad \text{Equation 11-4}$$

where p is sound pressure, Δp is just noticeable increase in sound pressure and p_o is sound pressure at the threshold (Green, 1988).

The intensity DL for pure tones exceeding 50 dB SPL is fairly independent of frequency but increases for low and high frequencies for sound levels lower than 50 dB SPL, especially lower than 20 dB SPL. When the tone is presented in the background of wideband noise, the intensity DL depends on the signal-to-noise ratio (SNR) for low SNRs but is independent of SNR for SNRs exceeding 20 dB. For SNRs close to 0 dB, the intensity DL is equal about to 6 to 8 dB (Henning and Bleiwas, 1967). Similar values for intensity DL are reported for the threshold of hearing in quiet.

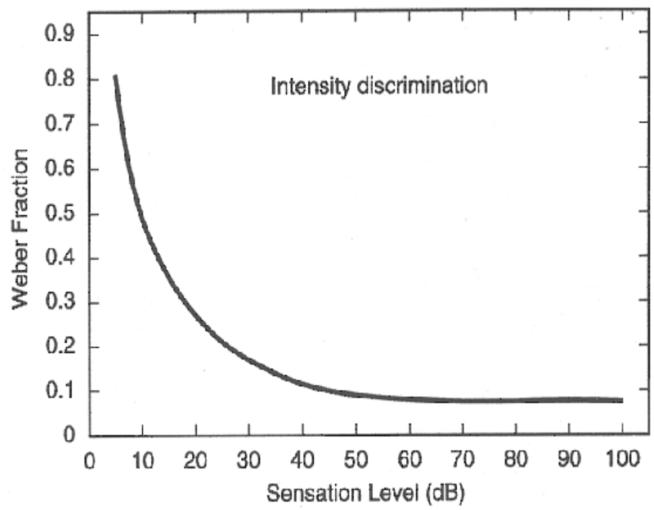


Figure 11-4. Weber fraction for intensity DL as a function of sensation level (i.e., number of decibels above threshold) for a 4 kHz tone. Data from Riesz (1928).

The intensity DL for wideband noises varies from about 0.4 to 0.8 dB depending on the type of noise (Miller, 1947; Pollack, 1951) and rises up to 1 to 3 dB for octave band noises depending on the center frequency of the noise (Small, Bacon and Fozard, 1959).

The Intensity DL depends also on the signal duration. This relationship is exponential and analogous to that of the dependence of stimulus loudness on signal duration (Garner and Miller, 1947b). The intensity DL (in dB) is independent of stimulus duration for durations exceeding 200 millisecond (ms) and increases at a rate of about 3 dB per halving the duration for durations shorter than 200 ms.

Frequency discrimination

The DL for frequency is defined as the minimum detectable change in frequency required detecting a change in pitch. Figure 11-5 presents frequency DL data reported by Wier, Jesteadt and Green. (1977). As can be seen in Figure 11-5 at low frequencies (below 500 Hz) the DL for frequency (in Hz) is relatively independent of frequency and increases logarithmically with frequency at mid and high frequencies. This shape is consistent with Weber’s law, i.e., the smallest noticeable change in frequency is a logarithmic function of frequency. For example, the smallest detectable change in frequency is about 1 Hz at 1000 Hz and about 10 Hz at 4000 Hz. In relative terms, the difference threshold at 1000 Hz corresponds to a change of about 0.1% in frequency. However, if expressed in logarithmic units, e.g., *cents* [3] (see the Pitch section later in this chapter), this difference is about 5 cents and remains constant across frequencies.

As shown in Figure 11-5, the frequency DL is dependent on the frequency and intensity of the stimuli being compared. It also depends on the duration and complexity of the stimuli. For tonal stimuli with intensity exceeding 30 dB SPL, average frequency DLs are about 1 to 2 Hz for frequencies below 500 Hz and 0.1 to 0.4%

---

[3] The *cent* is a logarithmic unit of measure used for musical intervals, often implemented in electronic tuners. Cents are used to measure extremely small intervals or to compare the sizes of comparable intervals in different tuning systems.

for frequencies above 1000 Hz (Koestner and Schenfeld, 1946; König, 1957; Letowski, 1982; Shower and Biddulph, 1931).[4] All these values are typical for average sound intensity levels, and they are the same or slightly smaller for increases in sound intensity up to about 80 dB SL (Wier, Jesteadt and Green, 1977). Similarly, the frequency DL decreases with increasing duration of short auditory stimuli and becomes independent of duration for stimuli exceeding 100 to 200 ms (Grobben 1971; Moore, 1973; Walliser, 1968; 1969). In addition, low frequency sounds need longer duration to be discriminated than high frequency sounds (Liang and Christovich, 1961; Sekey, 1963). Note also a very profound effect of training on frequency discrimination and recognition (Letowski, 1982; Moore, 1973; Smith, 1914).

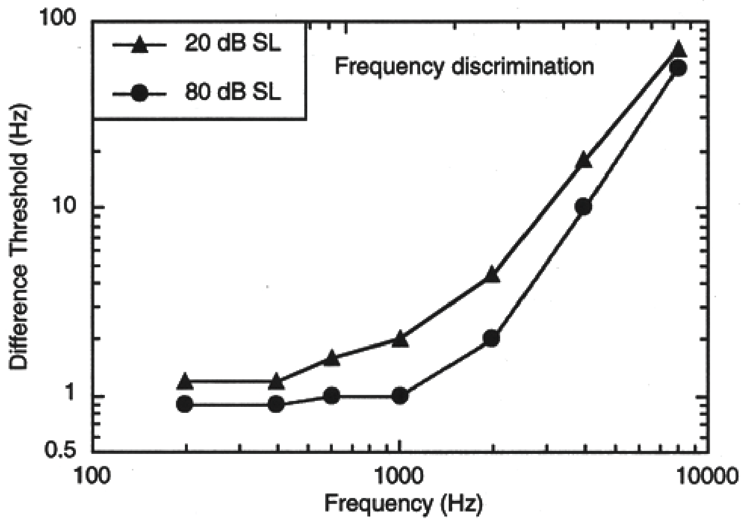


Figure 11-5. Frequency DL as a function of frequency. Data for pure tones presented at 20 and 80 dB SL (i.e. decibels above the threshold of hearing) (adapted from Wier, Jesteadt and Green, 1977).

The frequency DLs for narrow bands of noise are higher that corresponding DLs for pure tones and depend on the bandwidth of noise. According to Michaels (1957), frequency DLs for narrow band noises centered at 800 Hz vary from approximately 3 to 4 Hz for very narrow noises ($\Delta f < 12$ Hz) to more than 6 Hz for a noise band that is 64 Hz wide.

Frequency discrimination for complex tones (fundamental frequency with harmonics) is the same or better than for pure tones (Goldstein, 1973; Henning and Grosberg, 1968). Gockel et al. (2007) reported the frequency DLs as 0.1% and 0.2% for a complex tone and single harmonic, respectively. These values are representative of the frequency DLs found for music notes and vowel sounds, but the actual thresholds vary quite a bit depending on the fundamental frequency of the note and the type of music instrument producing it or the type of voice production, i.e., spectral and temporal envelopes of the sound (Kaernbach and Bering, 2001). However, for practical applications, it can be assumed that frequency DLs for complex tones are approximately constant in the 100 to 5000 Hz range (Wier, Jesteadt and Green, 1977).

Temporal discrimination

Auditory temporal discrimination has various forms and various discrimination thresholds. It refers to the human ability to distinguish between acoustic stimuli or silent intervals of different length, to detect a silent gap in an otherwise continuous stimulus, to resolve between one or two clicks presented in a succession, and to identify

[4] At low frequencies, the DL is constant in Hz; at mid and high frequencies, it is constant in percent (%).

temporal difference and order in the onsets of two overlapping stimuli. The corresponding temporal discrimination measures are called *sound duration DL*, *gap detection threshold*, *temporal resolution*, and *temporal order discrimination*, respectively.

The sound duration DL is the most commonly measured temporal discrimination capability. It depends on sonic content, temporal envelope of sound, and whether it applies to the sound itself or to the pause (gap) between two sounds. Abel (1972) reported duration DLs ranging from approximately 0.4 to 80 ms for stimulus durations of 0.2 and 1000 ms, respectively. In general, the duration DL of uniform (steady-state) sounds follows Weber's Law with a Weber fraction of around 0.1 to 0.2 for time durations greater than about 20 ms (Woodrow, 1951). Sounds with ramped down temporal envelopes are perceived as shorter, and sounds with ramped up temporal envelopes are perceived as longer than those with a uniform envelope (Schlauch, Ries and DiGiovanni, 2001).

A different type of auditory temporal resolution can be assessed by measuring the minimum detectable duration of a gap in a continuous sound. The gap detection is in the order of 2 to 3 ms for tones at moderate and high sound pressure levels (Exner, 1875; Ostroff et al., 2003; Plomp, 1964). Zwicker and Feldtkeller (1967) reported values of 1.5 ms and 5.0 ms for gap detection in tonal signals and noise, respectively. A gap detection threshold exceeding 15 ms is considered abnormal (Keith, Young and McCroskey, 1999). Experiments on gap detection in octave bands of noise have shown that temporal resolution is better at high frequencies than at low frequencies (Shailer and Moore, 1983). At low sound levels, minimum detectable gap duration increases considerably (Florentine and Buus, 1984). If the gap is presented periodically with a frequency $f_{int} \leq 25$ - 40 Hz in a continuous noise, then the noise is heard as a series of separate impulses. However if the frequency f_{int} increases above 25 to 40 Hz, the noise is heard as a continuous noise with a defined pitch corresponding to the frequency of interruptions. The sense of pitch decreases gradually for $f_{int} \geq 250$ Hz and disappears completely for f_{int} above 1000 Hz (Miller and Taylor, 1948).

The minimum detectable gap duration in continuous signals is very similar to the gap duration required to hear two similar clicks as separate events. Such temporal resolution is about 1.5 to 3 ms (Hirsch, 1959; Wallach, Newman and Rosenzweig, 1949), but it may increase to 10 ms for clicks that are greatly dissimilar (Leshowitz, 1971).

Temporal order discrimination requires substantially longer time intervals than temporal resolution or gap detection. The time difference required to determine the order of two sound onsets is to the order of approximately 30 to 60 ms. The actual time depends on gender (shorter for male listeners), age (shorter for young listeners), sound duration and temporal envelope, and whether both stimuli are presented to the same ear or to the opposite ears (dichotic task easier than monotic task) and varies between 20 and 60 ms (Rammsayer and Lustnauer, 1980; Szymaszek, Szelag and Sliwowska, 20006). However, temporal resolution does not seem to be much affected by a hearing loss (Fitzgibbons and Gordon-Salant, 1998). If the sounds overlap in time but have different onset times, they are heard as starting at the different points in time if their onset times differ by more than about 20 ms (Hirsh, 1959; Hirsch and Sherrick, 1961).

It is important to note that perception of the duration of a single acoustic event is affected by the temporal durations of the preceding events as well as by the rhythmic pattern of the events (Fraisse, 1982; Gabrielsson, 1974). For example, the duration of a short pause (gap) between two stimuli (e.g., 250 ms) is underestimated by 25% or more if it is preceded by another shorter pause (Suetomi and Nakajama, 1998). This effect is known as "time shrinking" and is an important element of music perception. It also has been reported that presentation of sound-distracter affects visual time-order perception (Dufour, 1999; McDonald et al., 2005).

Some information about temporal resolution of the auditory system also can be gleaned from data on the auditory detection of amplitude modulation (AM). Viemeister (1979) reported that detection of sinusoidal AM in noise is fairly independent of the modulation rate up to about 50 Hz and gradually decreases beyond this frequency, indicating that fluctuations of noise at higher frequencies are more difficult to detect, i.e., as the modulation rate increases and the time between amplitude peaks of noise becomes shorter, the depth of the modulation must be increased in order for the listener to detect the presence of modulation.

A good example of the practical limits of the auditory system in continuous processing of temporal information is auditory perception of Morse code. Morse code requires discrimination between long (dash) and short (dot) tone pulses separated by short (between symbols) and long (between blocks of symbols) pauses. The specific durations are relative and depend on the individual person, but they are usually in 1:3:1:3 relationships, respectively. Mauk and Buonomano (2004) reported that experts can understand Morse codes at rates of 40 to 80 words per minute (wpm), which for 40 wpm, results in timed events of 30, 90, 30, and 90 ms, respectively.

Cognitive discrimination

The differential thresholds discussed above apply to the smallest change in a single physical dimension that can be measured and assessed using one of the auditory sensations. These thresholds are important for signal and equipment designers and are used to optimize the usability of the products. They are also highly dependent on the overall cognitive capabilities of individual listeners and their familiarity with the situation to be assessed (Deary, Head and Egan, 1989; Helmbold, Troche and Rammasayer, 2005; Smith, 1914; Watson, 1991). However, in many cases the auditory stimuli to be compared differ in more than one physical characteristic, and the differences in their perception cannot be described by loudness, pitch and perceived duration alone. The perceived sounds also may be changing in time as their sources move across space. In such cases, other sensations, such as roughness, sharpness, or spaciousness can be used to differentiate and describe the stimuli of interest. These qualities are part of the domains of timbre and spatial character of sound and will be discussed in later sections of this chapter.

The above approach to evaluating sound events based on the perception of one or more physical dimensions may be quite effective in some situations but will not be sufficient for others. In the latter case, the stimuli can be differentiated at the sensory level using same-different criterion or the differentiation may require higher level processing and cognitive discrimination. For example, an HMD system designer may want to determine if the users will be able to differentiate between the old and new bandwidth of an audio HMD system using a pulsation threshold technique described in Chapter 13, *Auditory Conflicts and Illusions* (Letowski and Smurzynski, 1980). In another study, Nishiguchi et al. (2009) used a paired comparison technique to demonstrate that some listeners are able to discriminate between sounds with and without very high frequency components (f >20 kHz), while the majority of the listeners cannot. Both these studies are examples of the auditory discrimination task. A similar task is involved in differentiating between the sounds of two weapons or two vehicles. However a higher mental processing is required to recognize or identify the specific weapons or vehicles on the basis of their sounds. In an even more complex task performed daily people need to identify large number of speech phonemes in order to communicate by speech. In all these tasks, the listener is required to assign a specific sound to one or more nominal classes of sounds based on the listener's knowledge of the class characteristics. The cognitive processes involved in such decision making are usually described as *classification*, *recognition*, or *identification*.

Loudness

Loudness is an auditory sensation in terms of which sounds may be ordered on a scale extending from soft to loud (ISO, 2006). Loudness depends primarily upon the sound pressure of the stimulus but also depends upon its frequency, temporal envelope, spectral characteristics, and duration (International Electrotechnical Commission [IEC], 1995). Therefore, two sounds that have the same physical intensity (sound pressure, force of vibration) but differ in other physical characteristics may result in different sensations of loudness.

In its common, everyday usage, loudness is a categorical sensation that can be expressed in a number of terms such as *very loud*, *loud*, *soft* and *very soft*. If such a categorical (e.g., Likert) rating scale is used for scientific purposes, it is recommended that the scale has seven steps labeled (ISO, 2006):

Extremely Loud (100) - *Very Loud* (90) - *Loud* (70) -*Medium* (50) -*Soft* (30) - *Very Soft* (10) - *Not Heard* (0)

The numbers in parentheses are numeric values recommended for converting the loudness rating scale into numeric values suitable for averaging several ratings of a single person or a group of judges. In such cases, the minimum number of ratings being averaged should be 20 or higher (ISO, 2006) in order to approximate a Gaussian (normal) distribution in the data set.

Loudness level

Loudness level is a psychoacoustic metric that was developed to determine if sounds that differ in sound pressure (sound intensity) as well as other physical characteristics are equally loud without making a direct comparison for every combination of two of them. The unit of loudness level has been named the *phon*. A sound is said to have a loudness level of N phons if it is equal in loudness to a 1000 Hz tone having a sound pressure (intensity) level of N dB SPL (ANSI, 1994). Thus, a 1-kHz pure tone having a sound pressure level of 60 dB SPL and all other sounds that are equally loud have a loudness level of 60 phons.

The concept of loudness level was introduced primarily to compare the loudness of pure tones of different frequencies. Listeners were given a reference tone of 1 kHz and a test tone of a different frequency and asked to adjust the intensity level test tone until it matched the loudness of the reference tone. Such comparisons lead to the development of equal-loudness contours (iso-loudness curves) (Figure 11-6). The original iso-loudness curves were published by Fletcher and Munson (1933) and became the basis for the current standardized curves approved by the ISO (ISO, 2003). Each curve in Figure 11-6 connects the intensity levels of tones of different frequencies that are equally loud, i.e., the same loudness level in phons. Note that equal-loudness curves flatten gradually with the increase of the sound pressure level. This means that at high intensity levels the ear is less sensitive to fluctuations in intensity as a function of frequency than at low intensity levels.

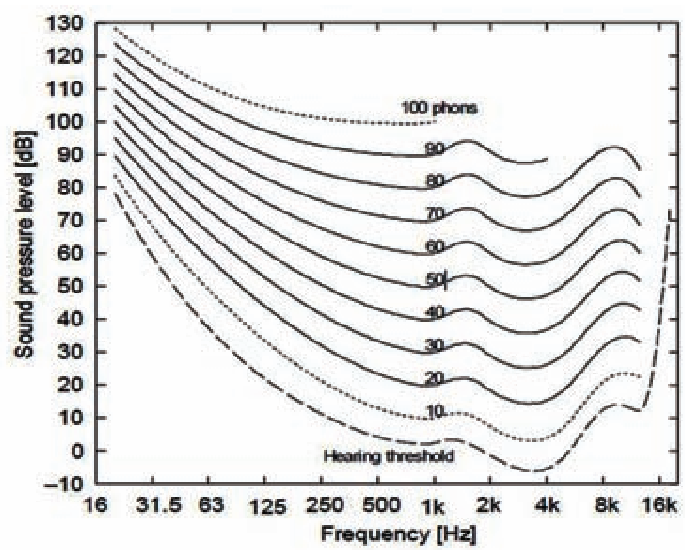


Figure 11-6. Equal-loudness contours for pure tones (adapted from ISO 226, 2003).

The equal-loudness contours for pure tones are not the only equal-loudness contours that have been developed. Similar equal-loudness contours for narrow band noises have been published by Pollack (1952). However, the equal-loudness contours for noises have never gained much popularity and are not widely used.

There are also approximate relationships between the eight formal music dynamic levels and loudness levels for various types of music. An example of such a relationship for symphonic music, based on observations made by Leopold Stokowski in the 1930s, is shown in Table 11-5. The weakness of this relationship is that the music dynamic levels are relative steps that can be different for each music piece and each music performance, and the relationship shown in Table 11-5 is only an approximation established for an average concert hall performance of symphonic music. The levels for chamber music will be much lower and the levels of rock music much higher (e.g., 140 phons at about 1 meter [3.28 feet] from a loudspeaker).

Table 11-5.
General relationship between music dynamics steps and the loudness levels for a typical concert hall performance of symphonic music (adapted from Slot, 1954).

Dynamic Level	Abbreviation	Loudness Level (phons)
forte fortissimo	Fff	90-100
fortissimo	Ff	80-90
forte	F	70-80
mezzoforte	Mf	60-70
mezzopiano	Pf	50-60
piano	P	40-50
pianissimo	Pp	30-40
piano pianissimo	Ppp	20-30

There also have been some attempts to apply the concept of equal-loudness contours to other perceptual attributes of sound. Fletcher (1934) introduced the concept of pitch level and equal-pitch contours to capture the effect of sound intensity on pitch of sound. Such contours were discussed later by Ward (1954) and Rakowski (1978; 1993). In addition, Thomas (1949) and Guirao and Stevens (1964) attempted to established iso-contours for auditory sensations of volume (Stevens, 1934a) and density (Stevens, 1934b), respectively. All of these attempts were short-lived, and neither triggered any wider interest in scientific community nor found practical applications.

Most comfortable loudness level

Most comfortable loudness (MCL) level has been defined as the listening level selected by the listener to optimize listening pleasure or communication effectiveness. It refers primarily to listening to natural sounds such as music, environmental sounds, and speech. MCL is important for audio HMD design because of the dependence of many perceptual responses on the level (loudness) of incoming stimuli. In almost all practical situations, listening to sound at the MCL results in the best and most consistent human performance. Listening at levels other than the MCL also demands increased attention resources and causes the listener to become fatigued more rapidly. Too high listening levels also may lead to temporary or even permanent hearing loss.

For most listeners the MCL for listening to speech in quiet or low levels of background noise is approximately 60 to 65 dB SPL, which corresponds to the level of normal conversational speech heard at a 1-meter (3.28-foot) distance (Denenberg and Altshuler, 1976; Gardner, 1964; Hochberg, 1975; Kopra and Blosser, 1968; Richards, 1975; Sammeth et al, 1989). This level corresponds roughly to 50 dB HL, which is used in most of the clinical evaluations of speech communication ability. Thus, the MCL of the listener should be the preferred level for speech stimuli delivered through audio HMDs in quiet environments. Speech stimuli also can be presented at both lower and higher levels if they were naturally produced at these levels, and the transmission is intended to truly reproduce the behavior of the talker. For example, natural levels for raised voice (raised speech level), loud speech, and shouting are about 65 to 75 dB SPL, 75 to 85 dB SPL and 85 to 95 dB SPL, respectively (Pearson, Bennett and Fidell, 1977).

One of the most important factors affecting the MCL of a listener for a given listening situation is the level of background noise. Kobayashi et al. (2007) reported that that noise levels up to 40 dB SPL have a negligible effect on the MCL for speech. Above this noise level, the MCL for speech appears to be the level that results in a SNR of approximately 15 dB. However, the fact that conversational speech is at 60 to 65 dB SPL combined with the 15 dB SNR requirement brings the noise levels that are negligible for speech communication to about 50 dB SPL. In addition, at high noise levels, the 15 dB SNR rule cannot be met. Richards (1975) and Beattie and Culibrk (1980) studied MCL levels for speech in noise and concluded that the MCL increases about 7 dB per 10 dB of increase in noise levels, up to about 100 dB SPL.

The MCLs for listening to music are substantially higher than those for speech and depend on the type of music, surrounding acoustics, and type of music instrument. Individual differences in MCLs for music are larger than those for speech and can vary from about 70 to 95 dB SPL.

MCLs are usually expressed as the sound intensity (pressure) level selected by the listener. However, they also can be expressed in phons. When expressed in phons, they become less dependent on the specific sound and are easily transferable to other listening situations. The typical MCL (in phons) for various types of music as calculated by the authors on the basis of several MCL studies (Gabrielsson and Sjögren, 1976; Martin and Grover, 1976; McDermott, 1969; Sone et al., 1994; Suzuki, Sone and Kanasashi, 1982; Staffeldt, 1974; Steinke, 1958) are:

- Symphonic and big-band music: 85 phons
- Solo and chamber music: 75 phons
- Artistic speech and solo singing: 65 phons

The selection of a very high listening level (above 85 phons) frequently makes the listening experience more exciting as opposed to remote (Freyer and Lee, 1980). However, it also makes the perceived sound image less clear due to nonlinear distortions generated in the middle ear and is more tiring (Kameoka and Kuriyagawa, 1966).

One area that requires special attention in respect to MCL is the perceptual assessment of sound, i.e., perceived sound quality (PSQ). Illényi and Korpassy (1981) conducted a number of listening tests of loudspeakers and demonstrated that louder sounds lead to higher ratings of the sound quality of the loudspeaker. This requires very careful loudness balance in PSQ assessment of sounds produced by different sound sources. It is also important for proper PSQ judgments that the sounds need to be reproduced at their natural levels (Gabrielsson and Sjögren, 1976; 1979) or at the ultimate listening levels, if such levels are known (Toole, 1982).

Loudness scale

Sensation of loudness is a perceptual representation of the amount of stimulation and depends primarily on sound intensity. In order to determine the effect of sound intensity on loudness, some type of psychophysical relationship between these two variables needs to be determined. One type of such a relationship is provided by the loudness level that allows comparing loudness of two or more sounds by comparing them to the equivalent loudness of a 1 kHz tone.

However, it does not allow one to determine how much louder one sound is with respect to another one. For example, the fact that one sound has a loudness level of 75 phons and another sound has a loudness level of 83 phones does not make it possible, by itself, to determine how much louder the second sound is. Such a comparison requires the direct representation of both sounds on a quantitative psychophysical loudness scale.

The first attempt to create a quantitative loudness scale was by Fechner (1860), who extended Weber's Law and assumed that the sensation of loudness increases by a constant amount each time the stimulus is increased by one DL. This dependence results in a logarithmic relationship between loudness (L) and sound intensity (I) having interval scale properties and is referred to as Fechner's Law or Weber-Fechner's Law:

$$L = a \times \log(I) + b, \qquad \text{Equation 11-5}$$

where *a* and *b* are constants dependent on the type of sound and a particular listener. The unit of loudness on Fechner's loudness scale is 1 DL, and the change of the stimulus intensity by 3 dB results in doubling of loudness. This relationship has been experimentally confirmed at low intensity levels at and slightly above the threshold of hearing where doubling of loudness requires a 2 to 4 dB increase in sound intensity. However, it overestimates the growth of loudness at higher levels. Research by Newman, Volkmann and Stevens (1937), Stevens (1955) and others led to the observation that for moderate and high intensity levels the loudness of a 1-kHz tone doubles when its sound pressure level increases by about 10 dB (Stevens, 1955). Thus, the shape of the loudness scale for the 1 kHz tone has been determined by Stevens to be a power function of the tone sound pressure level described as:

$$L = kI^{0.3} = kp^{0.6}, \qquad \text{Equation 11-6}$$

where *L* is loudness of sound, I is sound intensity, *p* is sound pressure, and *k* is the coefficient of proportionality that accounts for individual differences (Stevens, 1972a). This functional relationship sometimes is referred to in the literature as the Power Law of Loudness.

Since the 1-kHz tone serves as a reference sound for the loudness level, this means that loudness doubles when the loudness level increases by 10 phons. Therefore, in order to determine how much louder one sound is than another, one needs to determine the loudness levels of both sounds and compare them on the loudness scale for the 1-kHz tone.

The unit of loudness expressed by Equation 11-6 is a *sone*, defined as the loudness of a 1-kHz tone having a sound level of 40 dB SPL. Thus, the loudness of 1 sone corresponds to the loudness level of 40 phons. A sound that is N times louder has a loudness of N; and a sound that is N times softer has a loudness of 1/N sones. The relationship between loudness (L) and loudness level (LL) is such that a doubling of L in sones occurs for each increase in LL of 10 phons and can be written as:

$$L = 2^{\frac{LL-40}{10}}. \qquad \text{Equation 11-7}$$

The actual functional relationship between L and LL based on the data collected by Hellman and Zwislocki (1961) and other researchers is shown in Figure 11-7.

The function described by Equations 11-6 and 11-7 is shown in Figure 11-7 as a straight line and it matches experimental data very well for loudness levels of 30 phones or more. The curved portion of the loudness function indicates that at the threshold of hearing loudness grows more rapidly than at higher levels. This growth can be approximated by the modified loudness function equation:

$$L = k(p - p_0)^{0.6}, \qquad \text{Equation 11-8}$$

where p_o is the sound pressure at the threshold (Scharf, 1978). Although physiologic mechanisms behind the growth of the loudness function are not entirely clear, they have been related to both the overall number and timing of neural discharges (Carney, 1994; Fletcher, 1940; Relkin and Doucet, 1997) and the nonlinearities of both cochlear and central parts of the auditory system (Schlauch, DiGiovanni and Ries, 1998; Zeng and Shannon, 1994).

It has to be added that individuals with neurophysiologic hearing loss have elevated thresholds of hearing but still have about the same or even lower threshold of pain. These shifts in thresholds result in a narrower dynamic range of hearing and in a loudness function that has to have a steeper slope than that of normally hearing individuals. This rapid increase in loudness function associated with neurophysiologic hearing loss is called *recruitment*.

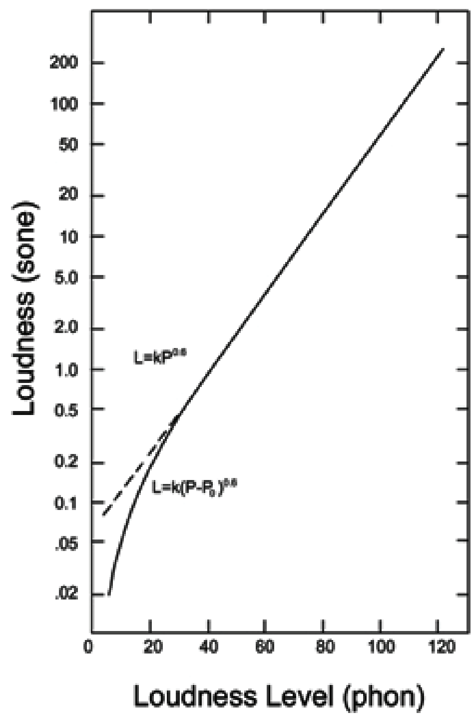


Figure 11-7. Binaural loudness of a 1 kHz tone as a function of loudness level (adapted from Scharf, 1978).

The discussion above assumes one-ear listening and the concept of monaural loudness. There is still a debate in the literature regarding the difference between *monaural* and *binaural* loudness. Marozeau et al. (2006) demonstrated that the difference between monaural and binaural loudness is practically independent of the sound pressure level. However, some researchers (e.g., Fletcher and Munson, 1933; Hellman, 1991; Marks, 1978; Pollack, 1948) reported that an increase in sound loudness due to binaural listening is equivalent to a 3 dB change in sound intensity received monaurally (doubling of sound intensity) while some others concluded that this change is more likely to be in 1.3 to 1.7-dB range (Scharf and Fishken, 1970; Wilby, Florentine, Wagner and Marozeau, 2006; Zwicker and Zwicker, 1991). This summation process seems to parallel an approximate 1.1 times (0.4 dB) binocular visual acuity and a 1.4 times (1.5 dB) contrast sensitivity advantage phenomenon in binocular vision (Rabin, 1995).

Temporal integration

The thresholds of hearing presented in Figures 11-1 and 11-2 were determined using continuous (long) pure tone stimuli; therefore, they are independent of sound duration. The same is true for the loudness functions expressed by Equations 11-6 and 11-8. However, for short sounds, both the threshold of hearing and sound loudness are affected by sound duration. The relationship between stimulus duration and the perceptual effects of the stimulus is referred to in the literature as *temporal integration* or *temporal summation*, and the changes in perceptual effects with stimulus duration have been attributed to temporal summation of excitations in the auditory system (Zwislocki, 1960).

The maximum duration of the stimulus through which the temporal summation effect operates is called *critical duration.* According to many studies, the critical duration for pure tone signals is approximately 200 to 300 ms, although this value depends somewhat on sound frequency (Miskolczy-Foder, 1959; Sanders and Honig, 1967, Zwislocki, 1960). The threshold of hearing is higher for durations shorter than the critical duration and decreases at a rate of about 3 dB per doubling of duration (Zwislocki, 1960). For example, for 100-μsec square-wave clicks presented at the rate of 10 Hz, the threshold of hearing is in the order of 35 dB SPL (Stapells, Picton and Smith, 1982), while the hearing threshold for continuous white noise is near 0 dB SPL. The functional relationship between the threshold of hearing and the stimulus duration for a 1 kHz tone is shown in Figure 11-8.

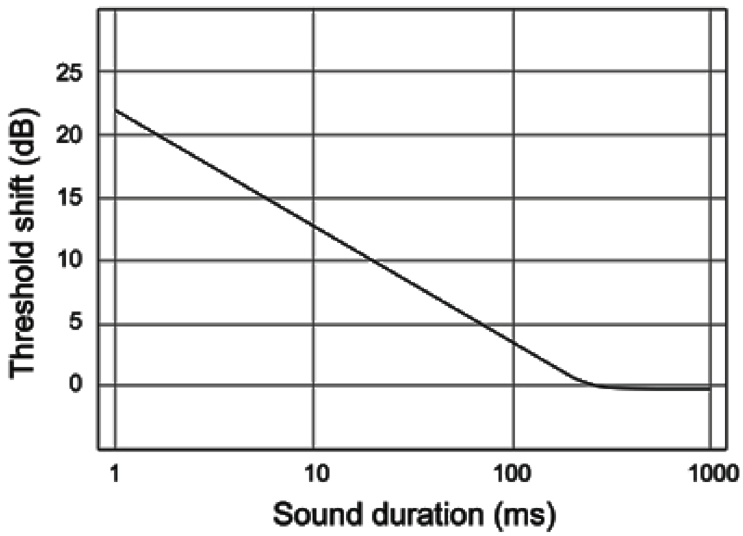


Figure 11-8. The effect of stimulus duration on the threshold of hearing for a 1-kHz tone (adapted from Zwislocki, 1960).

The temporal integration of energy in the auditory system also operates at the above-threshold (suprathreshold) levels, affecting sound loudness. For sounds shorter than critical duration, the loudness of sound increases with sound duration and this relationship can be described as:

$$SIL \times T = \text{constant loudness} \quad \text{Equation 11-9}$$

where SIL is sound intensity level in dB, and T is stimulus duration (in seconds) (Garner and Miller, 1947). Plomp and Bouman (1959) concluded that loudness is an exponential function of the duration of the sound and depends on the relationship between the stimulus duration and the time constant of the ear (determined to be 50 ms). According to these researchers, tonal stimuli that last for durations of 50 ms and 200 ms produce sensations of loudness that are equal to 62.70% and 99.98% of the loudness produced by a continuous sound, respectively.

The signal does not need to be a single short sound impulse to be affected by the mechanism of temporal integration. Series of clicks or short bursts of noise also are affected by the mechanism of temporal summation. However, bursts of higher repetition rate and shorter duration have been reported to sound louder than the same bursts of longer duration and slower repetition rate (Garner, 1948; Pollack, 1958). This effect may be attributed to an increasing neural firing rate by a group of neurons with increasing number of sound onsets. This increase in the firing rate seems to more than offset the effect of time latency (rest period) in a single neuron firing rate and should result in a decrease in sound loudness (Zwislocki, 1969).

Loudness summation

One important factor affecting loudness of a stimulus is the distribution of sound energy across the auditory frequency range. The loudness of a sound depends on where along the frequency scale the sound energy is located and how concentrated or dispersed is its allocation. Sound energy located in the area of greatest ear sensitivity (the lowest region for a given equal-loudness contour) contributes the most to sound loudness. The distribution of sound energy along the frequency scale affects the manner in which the auditory system integrates spectral components of the stimulus. This process is called *loudness summation* or, more accurately, *spectral integration* of sound energy by the auditory system.

Several algorithms have been proposed to model spectral integration of sound energy process in the development of the sensation of loudness. Some of the algorithms have been proposed by Fletcher and Munson (1933), Beranek et al. (1951), Howes (1971) and Stevens (1956). Further research led to observations that the process of spectral integration of sound is closely associated with the concept of critical bands (discussed later in this chapter). Briefly, if the sound components are located within a narrow frequency band smaller than a single critical band, the total loudness of sound is proportional to the total sound energy contained within the band. If the sound components are separated further apart than a critical band, the sound loudness is the sum of the loudnesses of the individual components. The two modern algorithms of loudness summation based on the general concept of critical band have been developed by Zwicker (Zwicker and Feldtkeller, 1955; Zwicker, 1960; Zwicker and Scharf, 1965) and Moore and Glasberg (Moore and Glasberg, 1996; Moore, Glasberg, and Baer, 1997) (see sections on Critical Bands and Loudness Scale for additional discussion on spectral summation and binaural summation, respectively.)

Auditory adaptation and fatigue

Auditory adaptation, or loudness adaptation, is a gradual decrease in hearing sensitivity during sustained, fixed-level, auditory stimulation. As shown in Figure 11-8, due to the effect of temporal integration, the sensation of loudness increases gradually with sound duration and reaches its terminal value for sounds longer than 200 to 300 ms. However, if the auditory stimulus acts for a prolonged period of time, the sensation of loudness slightly decreases. The decrease in sound loudness is accompanied by some decrease in hearing sensitivity for frequencies outside the frequency range of stimulation (Thwing, 1955).

The amount of adaptation is dependent on the frequency, level and duration of the auditory stimulus and increases with decreasing level of the stimulus and increasing frequency (Scharf, 1983; Tang, Liu and Zeng, 2006). Several early studies indicated strong auditory adaptation at all signal levels (e. g., Hood, 1950), but more recent studies demonstrated that under most listening conditions the auditory adaptation at high intensity levels is relatively minimal (Canévet et al., 1981). For example, Hellman, Miśkiewicz and Scharf (1997) reported that over the period of several minutes, the loudness of a continuous fixed level pure tone can decrease by 70% to 100% at 5 dB SL, 20% at 40 dB SL and stays practically constant at higher SLs. The exceptions are frequencies above 10 kHz, where the auditory adaptation effect is quite strong at both low and high stimulation levels (Miśkiewicz et al., 1992).

The physiologic mechanism responsible for auditory adaptation is still not clear. One possibility is a “restricted excitation pattern” mechanism proposed by Scharf (Miśkiewicz et al., 1992; Scharf, 1983). According to this concept, all low-level stimuli regardless of their frequency and all high-frequency stimuli regardless of their level produce a more restricted excitation pattern along the basilar membrane that is subjected to more adaptation than respective high-level and low-frequency stimuli.

Auditory adaptation needs to be differentiated from auditory fatigue. Fatigue is a loss of sensitivity as a result of auditory stimulation, manifesting itself as a temporary shift in the auditory threshold after termination of the stimulus. It is often referred to as a *temporary threshold shift* (TTS) and appears gradually for sounds exceeding 70 dB SPL. It differs from adaptation in two important ways. First, it is measured after the auditory stimulus has ended (poststimulatory fatigue); whereas auditory adaptation is measured while the adapting stimulus is still present (peristimulatory adaptation). Second, as a loss of sensitivity (rather than a shift in perception), it is a traumatic response to excessive stimulation by intense auditory stimuli (continuous noise above 85 dB or impulse noise above 140 dB). Exposure to recurring or extreme acoustic trauma can result in permanent hearing loss.

Masking

Sounds very rarely occur in isolation. They are usually heard as signals in the background of other sounds or are themselves a part of the background. The concurrent, or in close succession, presence of two or more sounds causes the audibility of the individual sounds to be adversely affected by the presence of other sounds. This adverse effect is called *masking*. Masking is defined as: (a) a process by which the threshold of hearing for one sound is raised by the presence of another sound and (b) the amount by which the threshold of hearing for one sound is elevated by the presence of another sound (ANSI, 1994).

A *masker* is a sound that affects audibility of another sound, the *target sound* (or *maskee*). More intense sounds mask less intense sounds. Masking effect of a target sound by a masker may be total, making the target sound inaudible, or partial, making it less loud. It should be noted that the masking phenomenon affects not only other sounds but also all individual components of a single complex sound. If the target sound is not completely masked by a given masker, the additional amplification of the masker needed to completely mask the target sound is called the *masking margin* (MM). The concept of the MM applies, among others, to the design of sound masking systems intended to provide privacy and security of acoustic information without creating excessively high noise levels.

The maskers can be of two types: *energetic* maskers, which physically affect the audibility of the target sound, and *informational* maskers, which have masking capabilities due to their similarity to the target sound. In general, both of these masking phenomena may exist together and may be caused by the same stimulus, but they are frequently considered separately due to the difference in the way they affect the audibility and identity of the target sound. Energetic masking is peripheral masking caused by the overlap of the excitation patterns created by the target sound and the masker along the basilar membrane and is considered to be a peripheral type of masking. Informational masking is related to non-energetic characteristics of the masker and may take place even if there is no overlap in the excitation patterns caused by the target stimulus and the masker along the basilar membrane. This type of masking is considered to originate in the central auditory nervous system.

A phenomenon very similar to masking and difficult to differentiate from masking is *perceptual fusion*. The concept of fusion applies mostly to complex sounds that have several qualities that need to be attended separately. In fusion and in masking, the distinct qualities of a target sound, or its partial loudness, are lost, and the physiological mechanisms underlying both phenomena are the same. Thus, both phenomena are most likely two different views of the same physiological process. In the masking approach, the focus of the observation is on the audibility of a single target sound; while in the fusion approach, the focus is on both the masker and the target sound, i.e., whether the masker and the target (masked) sound the same as the masker alone, or not. (Bregman, 1990; Schubert, 1978)

As with most of the auditory phenomena, masking can be monaural or binaural. However, if both the masker and the target sound are delivered to both ears, the target sound audibility is very much the same as in the case of monaural listening assuming that both ears are fairly identical. A common situation is that the masker affects both ears, and the target sound is only available at one of the ears. The reverse situation is also possible and, in such case, may affect localization of the sound source producing the target sound by masking the sound in one of the ears.

In addition, masking can be ipsilateral (masker and target (masked) sound in the same ear) or contralateral (masker and target sound in the opposite ears) (also known as peripheral and central masking, respectively). Ipsilateral masking is much stronger than contralateral masking, but the latter is frequently used to prevent sound leakage to the opposite ear (e.g., in bone conduction hearing tests). The difference in the effectiveness of both masking modes is in the order of 50 dB.

Energetic masking

The basic form of masking is related to sound energy and its distribution in frequency and time domains. This form of masking is called *energetic masking* (EM). There are two basic forms of energetic masking: *simultaneous* masking and *temporal* masking. Temporal masking is further divided into forward and backward masking. The other types of energetic masking discussed in the psychoacoustic literature, such as an overshoot masking, are just combinations of the two basic forms of energetic masking.

Simultaneous masking

Simultaneous masking is masking caused by a masker that is present throughout and possibly beyond the duration of the target sound. It is the most effective form of energetic masking. The amount of masking is dependent on the sound intensity of the masker and its spectral proximity to the target sound. Therefore, this form of masking is sometimes also referred to as spectral masking.

When the masker contains sufficient energy in the frequency region of the target sound, the masked threshold increases about 10 dB for every 10 dB increase in masker. Such relation between masker and masked threshold of the target sound can be observed when a pure tone is masked by wideband noise (Hawkins and Stevens, 1950; Zwicker and Fastl, 1999). This situation is shown in Figure 11-9.

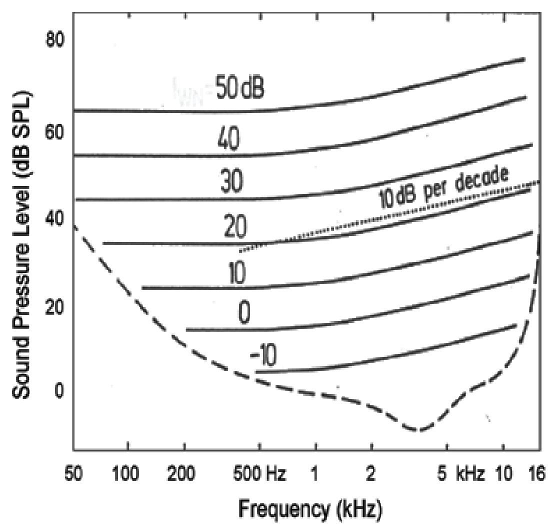


Figure 11-9. Detection thresholds of pure tones masked by white noise as a function of the frequency. Horizontal lines show masked thresholds for noise density levels from -10 to 50 dB and their relation to the threshold of hearing in quiet (dashed line) (adapted from Zwicker and Fastl, 1999).

The noise spectrum (spectral density) levels listed in Figure 11-9 indicate the density per Hz of white noise stimulus used as the masker. The masked threshold curves produced by white noise are fairly independent of frequency up to about 500 Hz and then increase with frequency at a rate of approximately 3 dB/octave (10 dB/decade). The equally-masking noise, which has constant density per Hz up to 500 Hz and then constant density per octave, i.e., density per Hz decreasing at a rate 3 dB/octave, would result at higher frequencies in practically frequency-independent masked threshold curves being parallel to the frequency axis. Other noises will result in quite different masking contours.

Masking produced by a continuous stationary noise is the simplest and most common form of energetic masking. The most common broadband noises that can be used as maskers in audio HMD testing (depending on the field application) are listed in Table 11-6. White noise and pink noise – noise that has the same power per relative (Δf/f) bandwidth – together with the equally-masking noise are frequently used as maskers in laboratory studies because they are well defined mathematically, and their effects on the audibility of the individual frequency components in the target sound are relatively easy to quantify. In addition, white noise and pink noise represent two important classes of real world maskers, e.g., thermal noise (e.g., heat noise, power generator noise, fan noise) and environmental noise (1/f noise).

Table 11-6.
Common wideband noises used for research purposes (adapted from Internet webpage the *Colors of Noise* [http://en.Wikipedia.org/wiki/Colors_of_noise.html]).

Noise Name	Description	Comments
Black Noise	No noise	Silence
Blue Noise	Noise that has a frequency spectrum envelope that changes proportionally to frequency. Blue noise has a spectral power density that increases by 3 dB per octave.	
Brown Noise	Noise with frequency spectrum envelope that changes proportionally $1/f^2$. Brown noise has a spectral power density that decreases by 6 dB per octave.	This name refers to Brownian motion that has these specific properties; also called Red Noise
Equally Masking Noise	Noise that equally masks tones of all frequencies	Also called Gray Noise
Pink Noise	Noise with frequency spectrum envelope that changes proportionally $1/f^2$. Pink noise has a spectral power density that decreases by 6 dB per octave.	
Purple Noise	Noise that has a frequency spectrum envelope that changes proportionally to f^2. Purple noise has a spectral power density that increases by 6 dB per octave.	Also called Violet Noise
White Noise	Noise that has flat frequency spectrum envelope. White noise has a constant spectral power density per Hz.	Acoustic analog of white light

Masking situations where a pure tone is masked by a narrow band of noise or another pure tone are shown in Figure 11-10. When the masking stimulus is a narrow band of noise the elevation of the threshold of hearing for pure tone target sounds is the greatest about the centre frequency of the noise. The masked threshold gradually and smoothly decreases for both low and high frequency target tones.

When both masker and the target sound are pure tones and have similar frequencies, they create beats (Egan and Hake, 1950; Wegel and Lane, 1924). Beats are periodic changes in sound intensity resembling amplitude modulation. When beats appear, they are heard as a tone with basic frequency f_o , which is the mean frequency of the two beating frequencies of the masker f_2 and target f_1:

$$f_o = (f_1 + f_2)/2 \qquad \text{Equation 11-10}$$

and the frequency of (f_{beats}) is equal to the difference between the frequencies of the masker and target:

$$f_{beats} = f_2 - f_1 \ . \qquad \text{Equation 11-11}$$

The presence of beats makes it easier for the listener to detect the target tone, even when the masker level is relatively high. These situations are shown in Figure 11-10 as dips in the masking curve around the frequency of the masker (400 Hz) and its harmonics (800 and 12000 Hz). The presence of beatings at the harmonic frequencies of the masker reveals the existence of nonlinear processes in the ear and the presence of the aural harmonics in the processed sound.

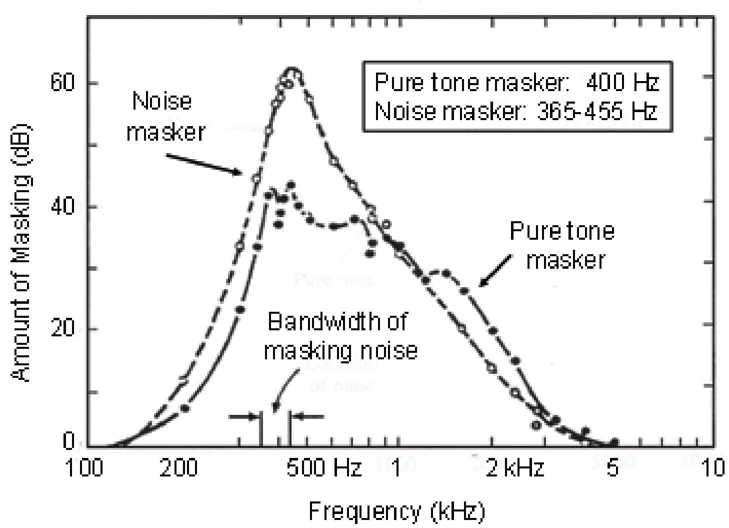


Figure 11-10. Masking effects of a 400 Hz pure tone and a narrow band of noise centered at 400 Hz (adapted from Egan and Hake, 1950).

The shape of the masked thresholds in Figure 11-10 shows that masking effect extends further in the high frequency region than in the low frequency region. In other words the upward spread of masking is much greater than the downward spread of masking, and this disproportional growth increases with the increase in the intensity of the masker. This situation is shown in Figure 11-11. The presence of the upper spread of masking also means that low frequency stimuli mask better high frequency stimuli better than the reverse.

In general, masking varies as a function of the frequency content of the masker. The closer the masker and target sound are on the frequency scale, the greater the masking. Thus, a narrowband noise centered on the frequency of a pure tone will have the greatest masking effect on that pure tone. As the bandwidth of the narrowband masker increases, its masking effectiveness increases until its bandwidth exceeds the limits of the critical band (see the later section on Critical Bands). However, further increase of the bandwidth of noise beyond the width of the critical band does not increase the masking power of the noise (Fletcher, 1940; Hamilton, 1957; Greenwood, 1961a,b). This can be explained by the fact that noise energy within the critical band prevents detection of the target sound because both the target sound and the masker are being passed to the same auditory system filter (auditory channel). However, noise energy outside of the critical band has no effect on detection of target sound because they pass through different filters. In addition, Buus et al. (1986) reported a 6 to 7 dB difference between the thresholds of detection for a pure tone (220, 1110, or 3850 Hz) and for an 18-tone complex tone[5] of uniform intensity when both were being masked by the same 64 dB SPL equally masking noise. The complex tone was detected easier. This finding indicates that simultaneous presence of signal energy in several critical bands aids signal detection.

---

[5] A *complex tone* is a sound consisting of several, usually harmonically related, pure tones.

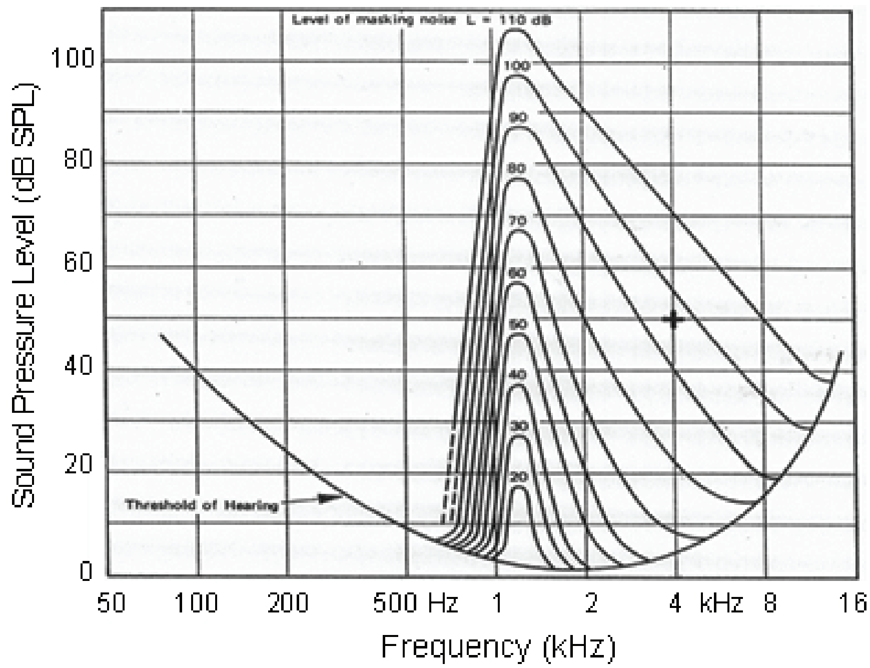


Figure 11-11. Masking effect of a narrow band of noise centered at 1200 Hz. The level of masking noise is shown next to each masked threshold (adapted from Zwicker and Feldtkeller, 1967).

Temporal masking

Masking caused by sounds that are not simultaneous with the target sound is called *temporal masking*. When two sounds arrive at the listener in short succession, the listener may hear only one sound event due to limited temporal resolution of the hearing system. However, if one of the two sounds has much higher sound intensity than the other, the listener still may hear only the more intense sound, even if the time difference between the sounds is above the temporal resolution limit of the hearing system.

There are two forms of temporal masking: forward (post-stimulatory) masking and backward (pre-stimulatory) masking. Forward masking appears when a short target sound is played after the end of the masker sound. If the time difference between the offsets of masker and target sound is very short, the sensory trace left by the masker decreases hearing sensitivity to the target stimulus resulting in its masking. The level of forward masking is dependent on the intensity of the masker, spectral similarity between the masker and target sound, and the time difference between the offsets of both sounds. Masking decreases as the intensity of the masker decreases, the separation between the sounds increases, and the time difference between the two offsets increases. In general, the increase of the wide band noise masker by 10 dB causes the increase of the detection threshold for immediate following tone by about 3 dB. Little masking occurs for times longer than 200 ms (Fastl, 1976; Jesteadt, Bacon and Lehman, 1982). The time difference between the offset of a masker and the onset of the target sound is inappropriate as a variable describing forward masking because the listener may still detect the target sound by hearing its end.

It has been demonstrated that the level of forward masking exerted by one tone on the subsequent tone can be decreased if additional tone is added to the masker in the region outside of the critical band of the target tone (Houtgast, 1974; Shannon, 1976). A similar but smaller effect can be observed in simultaneous masking (Fastl and Bechly, 1983). This phenomenon has been labeled *spectral unmasking* (Shannon, 1976) and is probably a result of physiological suppression of the excitatory response to the first tone by the addition of another tone (Sachs and Kiang, 1968).

Backward masking appears when the target sound is presented just before the masker. As with forward masking, the amount of backward masking is dependent on the intensity of the masker, spectral similarity between the masker and target sound, and the time difference between the offsets of both sounds. However, the time interval between the onsets of both sounds during which backward masking takes place rarely exceeds 25 ms. Although a large number of studies have been published on backward masking, the physiologic basis of this phenomenon is still largely unknown. Moore (1997) observed that, contrary to forward masking, the amount of backward masking decreases substantially with listener's experience and argued that backward masking may result from some kind of "confusion" between the target sound and the masker. The observed effects of forward and backward masking are illustrated in Figure 11-12.

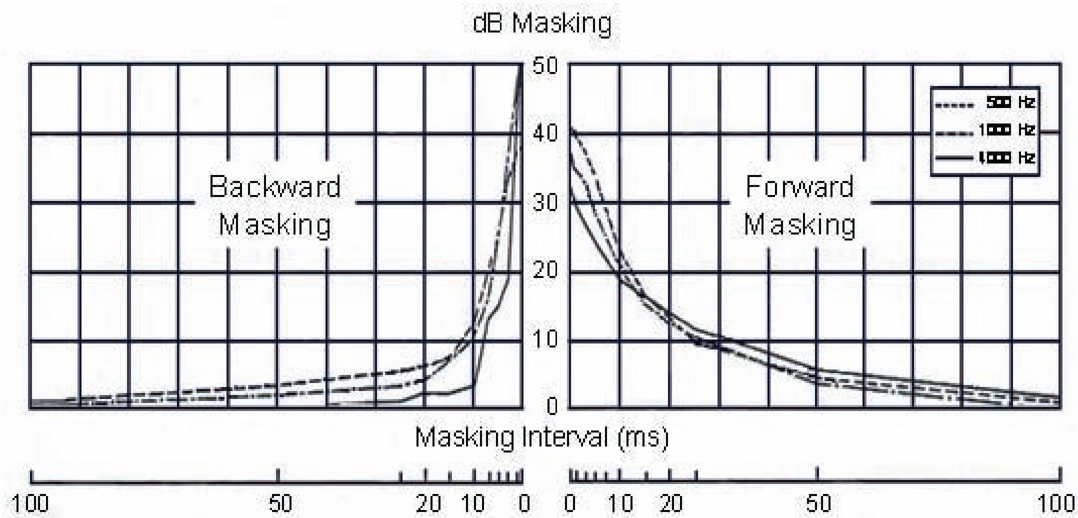


Figure 11-12. The relationship between the amount of backward (left panel) and forward (right panel) masking and time interval between the masker and target sound (adapted from Elliott, 1962).

Temporal masking, and especially forward masking, plays an important role in auditory perception because it degrades temporal cues in perceived stimuli. For example, in speech perception, a strong vowel may mask a weak consonant following or preceding the vowel. In addition, if speech communication takes place in a sound field, a strong reflection from a nearby wall may mask subsequent weak sound arriving along the direct pathway.

Informational masking

Informational masking (IM) is the amount of masking of one stimulus by another that cannot be explained by the presence of energetic masking. In other words, informational masking is the masking caused by the characteristics of the masker other than its energy. The amount of informational masking can be determined as a difference between the overall masking level and the masking level due to the energetic masking only. For example, *multitalker noise* (MTN), also known as speech babble, can serve as both energetic and informational maskers of a speech target, whereas a random (or frozen) white noise with speech spectrum envelope of the actual MTN can serve as an approximation of pure energetic masker.

Two main causes of informational masking are similarity between the masker and the target sound and the variability (uncertainty) of the masker (Durlach et al., 2003). The concept of informational masking originated in 1970s and was initially associated with the effect of spectro-temporal variability (uncertainty) of the masker or target on detection of the target stimulus (Dirks and Bower, 1968; Pollack, 1995; Watson, Kelly and Wroton, 1976). This concept later was expanded to include similarity between the masker and the target sound and spatial uncertainty regarding the location of the masker (Durlach et al., 2003). It has been demonstrated that the decrease in the degree of similarity between the target sound and the masker reduces substantially the amount of informational masking affecting the target sound (Kidd et al., 1994; Micheyl et al., 2000).

The reason that MTN is such an effective informational masker of speech is its overall similarity to the target speech. However, its actual effectiveness depends on the number of voices constituting the MTN, gender of the talkers, synchrony and rate of speech of the MTN voices, and the overall similarity of speech patterns of the MTN and the target speech. For example, masking effectiveness of an MTN increases with the number of voices, reaches its plateau for about 10 voices and then declines. Conversely, the content of the spoken messages, being a positive, neutral, or negative content, does not seem to have bearing on masking effectiveness of a MTN (Letowski et al., 1993; Letowski et al., 2001).

Informational masking due to masker uncertainty may be a result of either spectro-temporal uncertainty, spatial uncertainty, or both. Random variations in a masker spectrum outside of the protected zone located in close vicinity of the target stimulus have been reported to cause as much as 20 to 40 dB of additional masking. Numerous studies demonstrating the presence of additional masking caused by spectro-temporal uncertainty have been cited by Durlach et al. (2005). However, this increase reflects the joint effect of masker-target similarity and masker uncertainty. It can be argued that masker-target similarity is still the main cause of the masking increase shown in the reported studies. For example, Durlach et al. (2003) observed that masker-target similarity seems to greatly increase the effect of masker uncertainty on its masking effectiveness. Lufti (1990) analyzed a number of masking studies with naturally varying masking noise in each masking trial and concluded that the amount of informational masking in these studies was about 22% of the overall masking. The effect of spectro-temporal variability of the masker on the overall amount of masking also has been shown by Pfafflin and Matthews (1966), Pfafflin (1968), Lufti (1986) and others who compared effectiveness of natural random noise with that of the fixed (frozen) noise played in each masking trial.

Similarly, it has been shown that uncertainty of the spatial position of the masker can reduce speech intelligibility of the speech target (Kidd et al., 2007) or detection of the nonspeech target (Fan, Streeter and Durlach, 2008). Evidence of informational spatial masking in speech-on-speech masking situations can be found in frequent errors in substituting target words with words contained in the masking message. However, as compared to the effects of masker-target similarity or even spectral uncertainty, the effect of spatial uncertainty is very small. It can be argued that spectro-temporal and spatial uncertainties of the masker cause uncertainty about the target sound template or distract the attention of the listener, drawing it away from the target sound (Best et al., 2005; Conway, Cowan and Bunting, 2001).

It is also important to note that the amount of informational masking caused by a specific masker-target relationship is highly dependent on the listener, and that the inter-subject differences in respect to informational masking are very large (Durlach et al., 2003). Equally important is that the effectiveness of informational masking increases with age even in people with otologically normal hearing. Rajan and Cainer (2008) reported that with aging, independent of any hearing loss, older individuals (age 60 or greater) performed as well as younger individuals in speech recognition in an energetic masking background but performed much poorer in the presence of informational maskers. The authors attributed this difference to an age-related increase in competing-signal interference in the processing of auditory and phonetic cues.

Critical Bands

The concept of *critical band* is central to understanding the mechanisms of sound processing in the auditory system. This concept was introduced by Fletcher (Fletcher and Munson, 1933, 1937; Fletcher, 1948, 1940) to account for filtering actions of the human auditory system. Fletcher et al. studied loudness summation and masking of tones by various wideband noises and observed that only noise energy contained in a certain frequency band centered on the frequency of the pure tone contributes to the tone masking. They also noticed that the loudness of the tones separated by the width of this band is additive, while the loudness of the tones within this bandwidth is not. They called this bandwidth the critical band.

Fletcher (1940) originally assumed that to mask a tone, the total power of the masking noise has to be equal to the power of the tone and defined the critical band as a bandwidth of noise having power equal to the power of the tone:

$$P = N \times CB\,, \qquad \text{Equation 11-12}$$

where P is the power of a tone, N is the noise spectrum (noise spectrum density) level, and CB is the bandwidth of the noise that contributes to the masking effect, i.e., the critical band width. This concept is shown graphically in Figure 11-13.

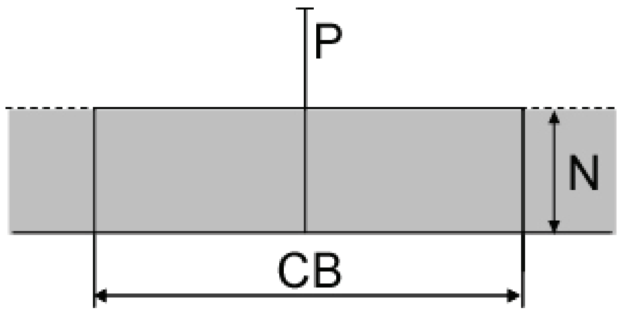


Figure 11-13. Fletcher's concept of the critical band. P – power of the tone (dB), N – noise spectrum level (dB), CB – critical band (Hz).

Equation 11-12 also can be written as:

$$CB = \frac{P}{N} \qquad \text{Equation 11-13}$$

And, after taking the logarithm of both sides and multiplying it by 10:

$$10\log CB = 10\log\frac{P}{N} = CB(dB) = CR \qquad \text{Equation 11-14}$$

where CB(dB) is a critical band expressed in dB, which is currently called the *critical ratio* (CR).

Critical ratio specifies the number of dB by which the power of the tone needs to exceed the noise spectrum level in order for the tone to be detected. For example, according to Fletcher's concept of critical bands, for a tone with a frequency of 1000 Hz, CB equals 65 Hz , and CR equals 18.1 dB.

Fletcher's concept of the critical bands was revised in 1950s by Zwicker when it was determined that in order to make a tone inaudible, the power of the masking noise needs to be about 2.5 times (4 dB) greater than the power of the masked tone (Zwicker, 1952;1961; Zwicker, Flottorp and Stevens, 1957). This finding extended the width of the critical bands by a factor of approximately 2.5. The new width of the critical bands also was confirmed in experiments on the threshold of hearing (Gässler, 1954; Hamilton, 1957; Zwicker and Feldtkeller, 1955) and loudness (Gässler, 1954; Zwicker, 1952) of complex sounds. For example, the relationship between the threshold of hearing at 1100 Hz and the bandwidth of the auditory stimulus reported by Gässler (1954) is shown in Figure 11-14. Gässler measured the threshold of hearing for a multi-tone complex composed of from 1 to 40 equal-amplitude pure tones evenly spaced 10 or 20 Hz apart. As the tones were added sequentially to the complex, the overall sound pressure at the threshold of hearing remained constant up to some defined width of the bandwidth. When the tones were added beyond this width, the overall sound pressure needed to elicit a threshold sensation increased with a slope of 3 dB per doubling of the signal bandwidth outside of the critical band. When additional components were added symmetrically on both sides of the critical band, the threshold increased at a rate of 1.5 dB per doubling of the signal bandwidth outside of the critical band.(Spiegel, 1979). These findings are consistent with the predictions of an energy-detector model of the auditory system.

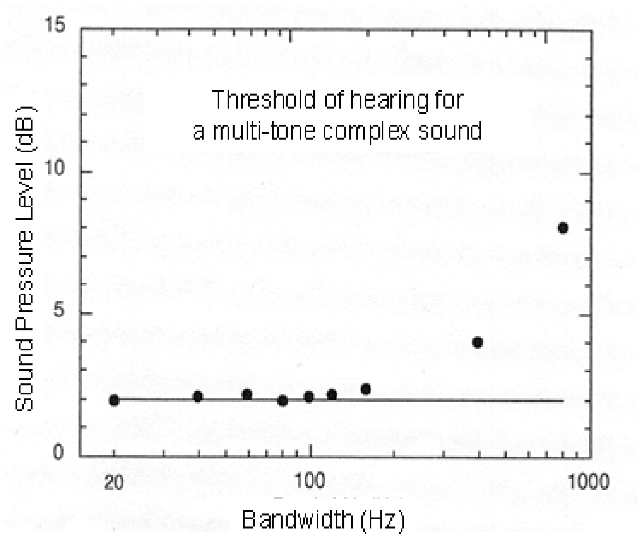


Figure 11-14. Threshold of hearing for a multi-tone complex as a function of bandwidth. Data are shown for tones added every 20 Hz below 1100 Hz. Continuous line shows the threshold of hearing for a single 1100 Hz tone (adapted from Gässler, 1954).

Similarly, the loudness of the complex auditory stimulus, with sound energy contained within a single critical band, is independent of the distribution of sound energy within the band. The effects of critical band on the loudness of a narrowband noise with a bandwidth changing from very narrow one to one that is wider than a critical band is shown in Figure 11-15.

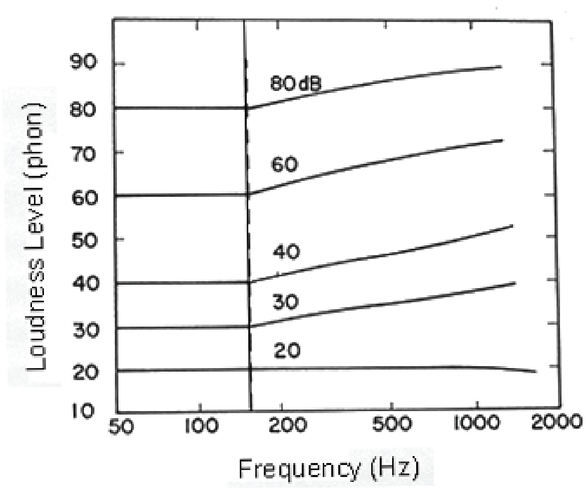


Figure 11-15. Loudness level of a narrow band of noise as a function of noise bandwidth. Numbers on the curves indicate the overall sound intensity level of the band (adapted from Scharf, 1978).

There is also very little effect on loudness due to the number of spectral components contained within a single critical band as long as the total energy of the complex remains unchanged. For example, several researchers have reported no difference in the loudness of two-tone complexes, four-tone complexes, and a broadband stimulus for stimuli contained within the same critical band (Zwicker and Feldtkeller, 1955, Feldtkeller and Zwicker, 1956; Zwicker et al., 1957, Scharf, 1959). Others have found a slightly higher loudness of a broadband noise in comparison to the loudness of the tonal stimuli, particularly at loudness levels near 65 phons (Florentine, Buus and Bonding, 1978). The overall loudness of two tones that are separated by less than 20 Hz is affected by the audible changes in sound intensity caused by beats and is dependent on the phase relationship between the tones (Zwicker, Flottorp and Stevens, 1957). More information about auditory system sensitivity to phase is included in the later section Phase and Polarity.

The size of Zwicker's critical bands (*Frequenzgruppen*) is about 100 Hz for frequencies below 500 Hz and increases with frequency *f* at about the 0.2*f* rate; this relationship is shown in Figure 11-16. Thus, the bandwidth of the critical band above 500 Hz can be roughly approximated by the bandwidth of 1/4 octave filters ($\Delta f = 0.18f$) with the same center frequency.

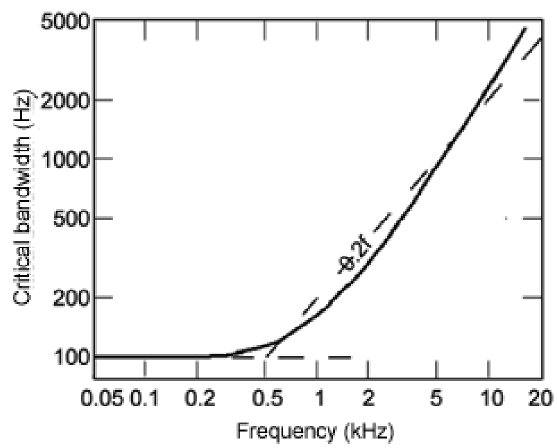


Figure 11-16. Critical bandwidth as a function of frequency (adapted from Zwicker and Fastl, 1999).

Since von Békésy's studies of basilar membrane and its tonotopic organization in 1920s and 1930s (Békésy, 1960), the term critical band is used also to denote regions of the basilar membrane that respond to stimulation by a sine wave input. Zwicker (1952) observed that when subsequent 24 CBs are placed back-to-back they cover almost the whole rage of hearing (0 to 15,500 Hz) and can be represented conveniently along the basilar membrane. He also demonstrated that a CB takes a relatively constant length of 1.3 mm along the basilar membrane and can be used as a unit of frequency along the basilar membrane. It also corresponds to about 1300 neurons in the cochlea (Zwislocki, 1965). This unit has been named the *bark* in honor of German physicist Heinrich Barkhausen who initiated perceptual measurements of loudness (Zwicker, 1961). The bark scale extends from 1 bark to 24 barks, and its functional relationship with frequency is shown in Figure 11-17.

The relationship between CB (in Hz) and the specific frequency *f* of the tonal stimulus, i.e., the center of the CB, as well as the distance *x* (in mm) of the point of maximal excitation on the basilar membrane from the oval window, can be calculated using a formula proposed by Greenwood (1961b):

$$CB = 22.9\,(0.006046\,f + 1) = 22.9 \times 10^{0.06x} \qquad \text{Equation 11-15}$$

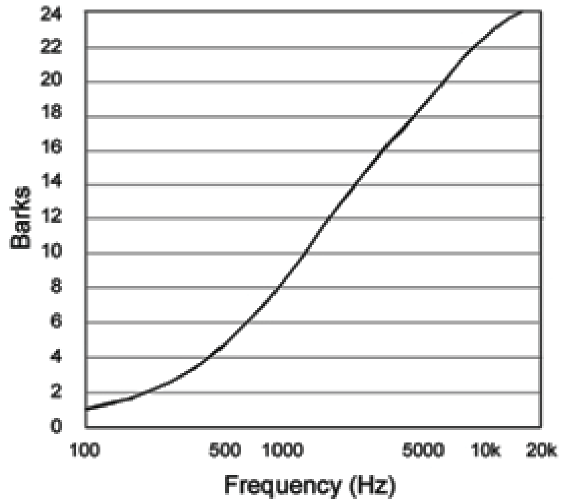


Figure 11-17. Critical band rate or barks as a function of frequency (adapted from Zwicker and Terhardt, 1980).

This relationship between the bark scale and the frequency scale shown in Figure 11-17 can be expressed mathematically as (Zwicker and Terhardt, 1980):

$$z = [13\arctan(0.76f) + 3.5\arctan\left(\frac{f}{56.25}\right)^2], \qquad \text{Equation 11-16}$$

where z is the distance along the basilar membrane in barks, and f is frequency of the stimulus in kHz. Other formulae to calculate the width of CB in Hz and in barks for specific frequencies have been published by Zwicker and Terhardt (1980) and Traunmüller (1990). This equation can be reformulated to calculate stimulus frequency for a known location on the bark scale and expressed as (Lubman, 1992):

$$f = \{[(\frac{e^{0.219z}}{352} + 0.1)z] - [0.032e^{[-0.15(z-5)^2]}]\} \qquad \text{Equation 11-17}$$

Barks are used frequently in modeling and simulations as an input to models of pitch perception, masking, and loudness summation, and noise hazard. The widths and lower and upper limits of critical bands for the 24 steps of the bark scale are listed in Table 11-7.

It is still unclear what the shape of the critical band filters is and whether it depends on sound intensity (e.g., Fletcher and Munson, 1937 (Figure 17); French and Steinberg, 1947 (Figure 8); Glasberg and Moore, 1990; Greenwood, 1961a,b). As with each mechanistic entity, such a filter has to have skirts with finite slopes. However, for many practical applications, it is convenient to assume that critical bands are brick-wall filters[6] with rectangular shapes. In their revision of Zwicker's loudness model, Moore and Glasberg (Glasberg and Moore, 1990; Moore and Glasberg, 1983; 1996; Moore, Glasberg and Baer, 1997) derived such a filter shape for critical bands in order to better account for the shape of the equal-loudness contours in low frequency range and the loudness of partially masked sounds. In their model of loudness summation Moore and Glasberg introduced the concept of the *equivalent rectangular bandwidth* (ERB) as a replacement for the critical band (bark) scale. The ERB is the bandwidth of a rectangular filter that has the same peak transmission as the auditory filter for that

[6] *Brick-wall filter* is an informal term for an idealized electronic filter, having full transmission in the pass band, complete attenuation in the stop band, and an abrupt transition(s) between the two bands.

frequency and passes the same total power for a white noise input (Moore, 1997). Its bandwidth varies as a function of frequency as:

$$ERB = 24.7(4.37f + 1) \qquad \text{Equation 11-18}$$

where f is the center frequency of the ERB filter. The function in Equation 11-18 has the same shape as the function (Equation 11-17) proposed by Greenwood for CBs and differs only in respect to constant values. A comparison of the critical bandwidths and the ERBs is shown in Figure 11-18.

Table 11-7.
Critical bands corresponding to 24 steps of the bark scale (adapted from Zwicker and Feldtkeller, 1967).

Bark band	Lower limit Frequency (Hz)	Center Frequency (Hz)	Bandwidth Δf (Hz)	Upper limit Frequency (Hz)
1	20	50	80	100
2	100	150	100	200
3	200	250	100	300
4	300	350	100	400
5	400	450	110	510
6	510	570	120	630
7	630	700	140	770
8	770	840	150	920
9	920	1000	160	1080
10	1080	1170	190	1270
11	1270	1370	210	1480
12	1480	1600	240	1720
13	1720	1850	280	2000
14	2000	2150	320	2320
15	2320	2500	380	2700
16	2700	2900	450	3150
17	3150	3400	50	4700
18	4700	4000	700	4400
19	4400	4800	900	5300
20	5300	5800	1100	6400
21	6400	7000	1300	7700
22	7700	8500	1800	9500
23	9500	10500	2500	12000
24	12000	13500	3500	15500

The shape of the critical bands (CBs) function in Figure 11-18 ERB is the same as in Figure 11-16 and the shape of ERB function is described by Equation 11-18. In some cases it is also convenient to think about ERBs as units of the frequency scale analogous to barks. The functional relationship between the number of ERBs, and the specific frequency is given by:

$$E = 21.4 \times \log(4.37f + 1), \qquad \text{Equation 11-19}$$

where E is the number of ERBs, and f is frequency in kHz. The constants of integration have been chosen to make E=0 where f=0 (Glasberg and Moore, 1990).

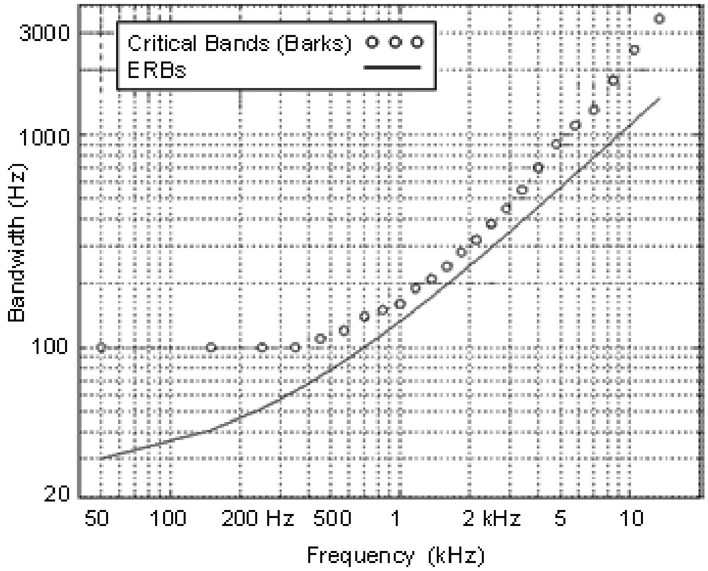


Figure 11-18. Critical bandwidth (Bark scale) and equivalent rectangular bandwidth (ERB) as a function of frequency (adapted from Smith and Abel, 1999).

Pitch

Pitch is the perceptual correlate of frequency. It is a sensation that the sound has a specific physical frequency. According to the formal ANSI standard definition of pitch, it is an auditory sensation that can be ordered on a scale extending from low to high (ANSI, 1994). Thus, low frequency pure tones are heard as being low in pitch, and high frequency pure tones are heard as being high in pitch. However, most sounds that occur are not pure tones, and yet many of them, but not all, have an identifiable pitch. Thus, pitch and frequency are not related in a simple one-to-one manner but depend on other physical properties of the stimulus, such as spectral complexity and intensity. Pitch is also a much more complex sensation than the sensations of loudness or perceived duration and actually has a multidimensional character

The sensations of pitch and rhythm are the foundation of music, which is a deliberate rhythmic sequence of sounds that differ in their pitch and/or timbre. The concept of pitch in music is closely related to the concepts of the musical scale and music intervals. The musical scale is a succession of selected notes (frequencies) arranged in ascending or descending order. The scale represents a specific music system and includes all the notes that are allowed to be used in this system, e.g., pentatonic system (5 notes in an octave), diatonic system (7 notes), or chromatic system (12 notes). The *key* or *tonic* of the scale is the first tone in the scale, and all subsequent tones are defined by simple ratio multiples of the tonic, e.g., 2:1 (octave), 3:2 (Major 5th), 3:4 (Major 4th), 5:4 (Major 3rd), or 6:5 (minor 3rd). The frequency ratios (pitch differences) within a given scale are referred to as intervals, and in many traditional music systems, they are not the exact multiples of each other. For example, in the diatonic scale, there are two unequal whole tone intervals 9:8 (major whole tone) and 10:9 (minor whole tone). Thus, because of these strict ratio relationship requirements, a musical instrument tuned to a particular key (tonic) would require retuning if one changed the key up or down a step. For instruments like the piano, or its earlier cousins, the clavier or the harpsichord, this was an onerous task. In the early 18th century, it became common for Western music to be written using an equally-tempered scale in which each octave is divided into 12 steps (semitones) (Helmholtz, 1863). These semitones are further divided into *cents*. Each semitone is 100 cents, and an octave is 1200 cents. The advantage of the equally-tempered scale is that any key can be played without changing the tuning of the instrument, and any song using the Western musical system can be written out using this notation.

The lower the frequency ratio, i.e., the smaller the numbers describing this ratio, the more similar in the character are the two notes separated by the interval (Galilei, 1638). The smallest possible frequency ratio is the ratio 2:1=2, which is called an octave. The octave has a special meaning in music because all sounds that are separated by one or more octaves fuse together very well and are sometimes very hard to differentiate from one another (Shepard, 1964). All other music intervals, such as semitone, tone, major third, or perfect fifth, are well defined within an octave and have the same sonic quality, called tone chroma, when repeated in other octaves. This octave equivalence led to the naming convention used in Western music, such that the notes (frequencies) that are an octave apart are named with the same letter (e.g., C, D, E) or syllable (e.g., do, re, mi) (Justus and Bharucha, 2002).

The concepts of music scale and octave similarity (tone chroma) led to the recognition of the two-dimensional character of pitch: pitch height and pitch class. Pitch height is the pitch defined in the ANSI standard (used at the beginning of this section). It is a continuous dimension logarithmically related to stimulus frequency. Therefore, a sound at 440 Hz (A4) is perceived as being equidistant from both 220 Hz (A3) and 880 Hz (A5). As all other auditory sensations, pitch height depends also to some degree on other basic physical parameters of sound, e.g., sound intensity and duration.

Pitch class, or tone chroma, is a dimension arranging music intervals from the smallest to the largest within a single octave. So, a middle C in the Western music system is lower in pitch height than the C one octave above it, but they occupy the same position on the pitch class scale. This terminology captures the circular nature of pitch that is a foundation of most of Western music. Both pitch dimensions, pitch height and pitch class, can be combined together in one helical representation of pitch.

One of the most remarkable properties of the human auditory system is its ability to extract pitch from complex tones. If a group of pure tones, equally spaced in frequency are presented together, a pitch corresponding to the common frequency distance between the individual components will be heard. For example, if the pure tones with frequencies of 700, 800, and 900 Hz are presented together, the result is a complex sound with an underlying pitch corresponding to that of a 100 Hz tone. Since there is no physical energy at the frequency of 100 Hz in the complex, such a pitch sensation is called *residual pitch* or *virtual pitch* (Schouten 1940; Schouten, Ritsma and Cardozo, 1961). Licklider (1954) demonstrated that both the *plac*e (spectral) pitch and the *residual* (virtual) pitch have the same properties and cannot be auditorally differentiated. In a harmonic tone, such as descried above, the residual pitch is often described as pitch corresponding to a *missing fundamental* frequency. The sensation of residual pitch is the main evidence that the auditory system must be able to code frequency based on its periodicity (see Chapter 9, *Auditory Function*). It also invalidates the so-called Ohm's Acoustic Law, which states that "Each tone of different pitch in a complex sound originates from the objective existence of that frequency in the Fourier analysis of the acoustic wave pattern."

Note, also, that a listener may listen to a complex sound in two different ways: analytically and synthetically. When listening analytically, the listener is focused on individual components of the sound and may hear their individual pitches. When listening synthetically, or holistically, the listener perceives the sound as a whole and pays attention only to the fundamental (or residual) pitch (Smoorenburg, 1970). So, the listener who listens analytically to a complex tone with a missing fundamental may not immediately recognize its residual pitch.

In reality, the dimensions of pitch height and pitch class (tone chroma) are not the only two dimensions of pitch. Another dimension of pitch is the pitch strength. Pitch height and pitch class are sufficient to describe the relationship between pure tones but not between complex natural, synthetic, and speech sounds. Sounds are usually composed of a number of frequency components that may be in harmonic or inharmonic relationships. It is the relationship of these components that determines whether a sound is tonal – i.e., carries the pitch of its fundamental frequency, or atonal. Most, although not all, of the music sounds are presumed to be tonal, however, outside of the realm of music many sounds contain frequencies that are not in harmonic relations. There are also musical instruments that produce sounds with inharmonic overtones (partials). The degree to which a specific sound has an identifiable pitch is called its *pitch strength* (Rakowski, 1977). Fastl and Stoll (1979) asked listeners to complete a magnitude estimation task for a number of test sounds, including pure tones, low-pass complex tones, complex tones, narrow-band noise and various other kinds of modulated or filtered noises. The general findings were that sounds with an orderly pattern of harmonics had the strongest pitch, as well as those containing a narrow band of frequencies. The pitch strength ranking of some test sounds investigated by Fastl and Stoll (1979) is shown in Table 11-8. The sounds with a more random and/or broadband spectral content have only a faint or no pitch strength. Shofner and Selas (2002) summarized their findings by stating that pitch strength depends primarily on the fine structure of the waveform and secondarily on the stimulus envelope. The relative perceptual salience of pitch in tonal complexes can be also estimated using an algorithm developed by Terhardt et al. (1982). In the case of residual pitch, the pitch strength decreases with an increase of the average frequency of the tonal complex and is the strongest for harmonics in the region of the third, fourth, and fifth harmonic (Boer, de, 1956; Ritsma, 1967; Ritsma and Bilsen, 1970).

Table 11-8.
Pitch strength rankings for 11 test sounds as obtained by Fastl and Stoll (1979).

Pitch Strength Ranking	Test Sound
1	Pure tone
2	Complex tone: -3 dB/octave low pass
3	Complex tone: -3 dB/octave
4	Narrow-band noise: Δf = 10Hz
5	AM tone: m=1
6	Complex tone: Band Pass
7	Band-pass noise: 96 dB/octave
8	Low-pass noise: 192 dB/octave
9	Comb-filtered noise: d=40 dB
10	AM noise: m=1
11	High-pass noise: 192 dB/octave

The Western music tonal system has influenced heavily the human concept of the pitch height scale, which is based on the logarithmic scaling of frequency perceived as pitch. Octave intervals are said to have the same pitch class (tonal chroma) and serve as equal steps of the music scale. However, it does not mean that they are perceptually equal although they are frequently treated that way. In order to answer this question Stevens, Volkmann and Newman (1937) constructed perceptual scale of pitch asking listeners to adjust a pure tone stimulus until it sounded half as high as a comparison stimulus (ratio scaling). They also proposed the *mel* (from the word “melody”) as a unit of the pitch scale. The mel has been defined as 1/1000th of the pitch of a 1000 Hz tone presented at 40 dB HL. Thus, the pitch of a 1000 Hz tone at 40 dB HL is equal to 1000 mels and equal numeric distances on the pitch scale were defined as equal perceptual distances although the developed scale should be treated more like a ratio scale. Later, Stevens and Volkmann (1960) conducted a similar study asking the listeners to divide frequency range from 200 to 6500 Hz into four equal intervals (interval scale). This new pitch scale used the same reference point of 1000 mels at 1000 Hz as the previous scale but was truly an interval scale. Due to the difference in testing methodology, the scales ware not identical, and the new scale was compressed heavily above 1000 Hz in comparison to the old scale. The relationship between pitch and frequency arrived at by Stevens and Volkmann (1940) is:

$$m = 1127 \times \ln\left(1 + \frac{f}{700}\right), \qquad \text{Equation 11-20}$$

where m is pitch in mels, and f is the frequency in Hz.

The pitch scale based on mels differs from the music scale and has been criticized for this reason. The musical objections to the pitch scale are that it is counterintuitive and counterproductive for different octaves to have different perceptual size (Hartmann, 1997). An explanation of confusion between pitch doubling and octave similarity might be found in the complex tones that are generated by musical instruments, which commonly consist of frequency components having a harmonic relationship to each other. These same harmonic relationships are the basis of the scales upon which musical structure is formed. Further, music pieces consisting of multiple voices depend on these same harmonic relationships to create consonance and dissonance. Thus, the entire structure of Western music depends on the mathematical relationships of frequency components of the complex tones within it. It does not depend on whether or not those pitches are perceived as being equidistant. In reality, an octave from 50 to 100 Hz sounds perceptually smaller than the octave from 2000 to 4000 Hz, and the frequency of 1300 Hz is reported by several investigators as having a half of the pitch of frequency of 8000 Hz (Zwicker and Fastl, 1999). These observations support the concept that pitch has actually two separate dimensions: pitch height (measured in mels) and pitch class (measured in music intervals).

In addition to two pitch scales developed by Stevens and his colleagues, Zwicker and Fastl (1999) constructed a third scale, also based on mels, using ratio scaling and a reference point of 125 mels set at 125 Hz. The scale extends form 0 to 2400 mels for the frequency region from 20 Hz to about 16 kHz and is shown in Figure 11-19.

Below about 500 Hz, the mel scale and the Hz scale are roughly equivalent, and the standard tuning frequency of 440 Hz has pitch of 440 mels. Above 500 Hz, larger and larger intervals are considered to be equivalent. As a result, four octaves on the frequency scale above 500 Hz correspond to about two octaves on the pitch mel scale. For example, the frequency of 8000 Hz has pitch equal to 2100 mels, while the frequency of 1300 Hz has pitch of 1050 mels. This relationship agrees well with the earlier experimental finding that a tone of 1300 Hz has half of the pitch of a 8000 Hz tone.

The pitch scale developed by Zwicker and his colleagues is highly correlated with the bark scale and with the distribution of excitation patterns along the basilar membrane. One can note the remarkable similarity between the bark scale (Figure 11-17) and the mel scale shown in Figure 11-19. This similarity indicates that the mel scale is practically parallel to the bark scale, and the unit of 1 bark corresponds to 100 mels. Both scales also are related to the ERB scale (Moore and Glasberg, 1983) and to the distance along the basilar membrane (Greenwood, 1961; 1990). It must be stressed, however, that all these scales have been developed for pure tones and do not directly apply to speech and music. Their major role is to help to understand frequency decoding and pitch encoding by the auditory system.

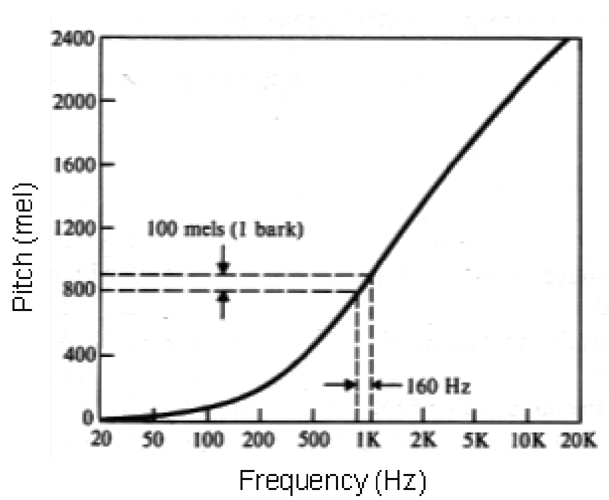


Figure 11-19. The relationship of the mel scale to frequency (adapted from Wightman and Green, 1974).

The relationships between frequency, pitch, critical bands, and the distance along the basilar membrane are shown together in Figure 11-20.

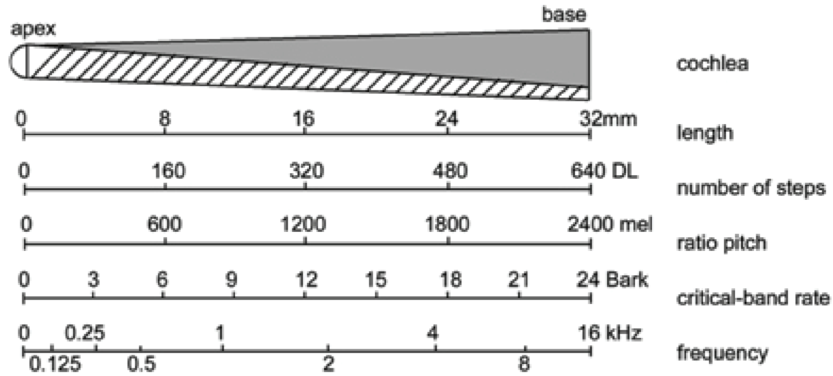


Figure 11-20. Similarity between various psychophysical scales and distribution of neurons along basilar membrane. Note that the scales of the length of the basilar membrane, numbers of DL steps, ratio pitch, and barks are linearly related while scale while frequency is not (adapted from Zwicker and Fastl, 1999).

One of the important concepts in music and everyday sound perception is the concept of consonance and dissonance. Music intervals, chords, or any combination of frequencies may be pleasant or unpleasant to the listener. The pleasant sounds are called *consonant sounds* and those that are unpleasant are called *dissonant sounds*. Dissonance occurs if the frequency separation between the individual tones of the sound is smaller than a critical band with its maximum for tones separation equal about ¼ of the critical band (Plomp and Levelt, 1965). This separation corresponds to about 20 Hz for lower and about 4% for higher sound frequencies. Helmholtz (1863) attributed the perception of dissonance to the sensation of beats or the roughness of sound, and Stumpf (1911) attributed it to perception of sound fusion, i.e., to ability of two sounds to assume a new identity, independent of their individual identities, when heard together (see section on Masking). Roughness and the dissonance, according to Helmholtz, are more likely represented in the auditory cortex by neural responses phase-locked to the amplitude-modulated temporal envelope of complex sound (Fishman et al., (2001).

In addition to the relationship to frequency discussed above, the pitch of a sound is dependent on its intensity and duration. Stevens (1935) and Gullick (1971) demonstrated that for middle frequencies (1 to 2 kHz in Stevens' and 2.5 kHz in Gullick's case), the pitch of a pure tone is independent of the stimulus intensity. However, for tones of higher frequencies, increased sound intensity produces an increase in pitch. Conversely, for tones of lower frequencies, increased sound intensity produces a decrease in pitch. Gullick (1971) reported that both shifts are similar for frequencies equidistant form the reference frequency of 2.5 kHz tone if expressed in terms of frequency DLs but not Hz. For example, a change in sound intensity of 40 dB resulted in similar but opposite shifts by 7 DLs for tones of 700 and 7000 Hz. For music, the effect of sound intensity on sound pitch is much smaller than for pure tones and is of the order of 17 cents for 30 dB change in sound intensity (Rossing, 1989). The direction of the change depends on the dominant components of the sound (Terhardt, 1979).

The effects of sound intensity on perceived pitch reported by Stevens (1936) and Gullick (1971) were measured by presenting two static tones of different frequencies and intensities and asking the listener to adjust the frequency or intensity of one of the tones until they seemed equal in pitch. However, sounds also can shift dynamically in both the frequency and intensity like in the case, for example, of the Doppler effect (Doppler shift). The Doppler effect is the change in the frequency of the arriving at the listener sound produced by a moving sound source. As the sound source approaches the listener, the compressions and rarefaction of the produced sound wave become compressed, making the frequency of the sound reaching the listener's ears higher than the frequency of the actually emitted sound. As the sound source passes the listener and moves away, the distances between the compressions and rarefactions of the sound wave became stretched, making the frequency of the sound reaching the listener's ears lower that the frequency of the sound produced by the departing sound source. So, if a sound source emitting a sound of frequency f_o is traveling at a constant velocity directly toward the listener, the sound that reaches the listener's ears has higher frequency than the frequency of the sound produced by the sound source but the difference between both frequencies is constant until the sound source reaches the listener's position. As the sound source passes the listener, the frequency of the propagating sound will drop suddenly. As a result, as the sound source moves away from the listener, the frequency of the sound that reaches the listener's ears is lower than that of the emitted sound. During the same time, the intensity of sound arriving at the listener's ear will gradually rise as the sound source is approaching the listener and gradually fall as it sound source moves away. A common example given of this effect is that of the sound of a passing vehicle using a classical siren.

Neuhoff and McBeath (1996) studied the effect of the Doppler shift on the pitch perceived by the listeners and found that the majority of the listeners reported that a Doppler shift consists of a rising (sound source is approaching the listener) and then falling (sound source is moving away from the listener) pitch shift. They then presented listeners with 27 Doppler shifted tones, created using three frequencies (220, 932 and 2093 Hz), three levels of spectral complexity (sinusoid, square wave, tone complex), and three velocities (10, 15 and 20 meters/second [m/s]). Listeners tracked the pitch using a pitch wheel. Listeners reported hearing a rise in pitch on 70% of all trials. The only conditions where a rising frequency was not reported were those of the lowest frequency pure tone. The probability of reporting a rise in pitch increased as a function of frequency and spectral complexity. The contour of the reported rise in the perceived pitch occurred synchronously with the rise in intensity suggesting that listeners perceived the rising sound intensity as an upward shift in pitch. This was true even at the two lowest frequencies, i.e., the frequencies where there should be no shift or downward shift in frequency according to Stevens (1955) and Gullick (1971). This situation is shown in Figure 11-21. Note also that if the sound source is traveling slightly at the angle to the listener's position, there is a slight relative decrease in the velocity of the sound source as it gets closer to the listener. However, this change will result in be a small decrease and not increase in sound frequency arriving at the listener's ears.

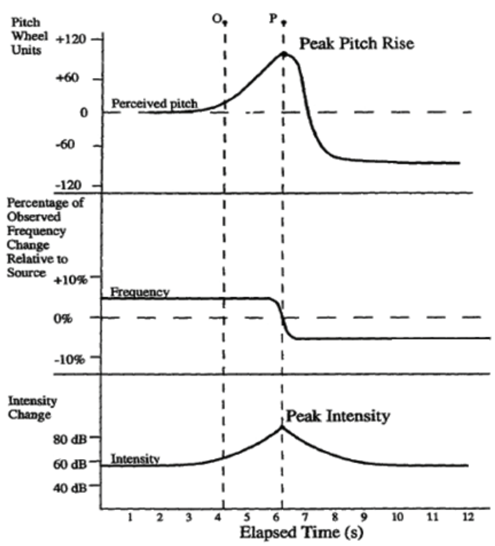


Figure 11-21. Schematic representation of the stimuli used in Neuhoff and McBeath's (1996).study. The bottom frame shows the intensity of the sound at the listener's ears. The middle frame shows the frequency at the listener's ears. The top frame shows the perceived pitch as listeners reported it using a pitch wheel (used with permission from Neuhoff and McBeath, 1996).

To test the hypothesis that the reported effect was due to dynamically changing sound intensity, Neuhoff and McBeath (1996) then asked the listeners to select the higher pitch tone of pairs of static tones consisting of a loud, lower frequency tone and a soft, higher frequency tone. For static tones, listeners accurately judged pitch, suggesting that the dynamic changes in both the intensity and frequency of the Doppler shifted tones are responsible for their perceptual interaction. Neuhoff's data suggest that pitch and loudness are perceived integrally (i.e., changes in one dimension can be perceived as changes in the other), a finding supported later by other research (Grau and Kemler-Nelson, 1988: Scharine, 2002). From a practical standpoint, the interrelationship of two perceptual dimensions underscores the complexity of pitch scaling and suggests that signal designers must exercise caution in using frequency as the basis for presenting dynamically changing information as its perception can be easily influenced by secondary factors. These situations are discussed further in Chapter 14, *Auditory Conflict and Illusions.*

The minimal duration of a pure tone needed to develop the full sensation of pitch depends primarily on the frequency of the stimulus and to a smaller degree on its intensity (Doughty and Garner, 1947). In general, for frequencies below 1000 Hz, a tone needs about 6 to 10 periods (cycles) to develop a sense of tonality, the so-called click-pitch sensation. For frequencies above 1000 Hz, the minimal duration needed to develop a click-pitch sensation is about 10 ms (Gullick, Gescheider and Frisona, 1989). The strength of pitch of short tonal and harmonic stimuli increases gradually up to about 100 to 250 ms (Bürck, Kotowski and Lichte, 1936; Moore, 1973; Turnbull, 1944). For unresolved complex tones, i.e., the tones consisting of only high order harmonics, pitch perception depends primarily on the repetition rate of the sound envelope and sound duration (White and Plack, 2003).

Phase and Polarity

The perception of phase and polarity has been a long-debated topic in audition. In general, numerous studies have shown that people are sensitive to neither absolute nor relative phase difference between various components of the periodic stimulus if the components are separated in their frequencies by more than one critical band. Hartman (1997) observed that phase difference between two pure tones separated by more than one critical band is irrelevant to audition because there is no single auditory neuron that responds to both tones. For example, changes in the phase relationship between the fundamental frequency and its lower resolved harmonics (separate by more than one critical band) have no audible effect, despite the fact that these changes greatly affect the temporal properties of the signal waveform. The fact that people are in general insensitive to the phase of the signal supports the general concept that the auditory system is a power (energy) detector rather than a pressure detector (Howes, 1971).

However, if two frequency components, e.g., harmonics, fall into the same critical band and their difference in frequency is rather small, the changes in their phase relationship are audible and affect both pitch value and pitch clarity (Lundeen and Small, 1984; Moore, 1977; Moore and Peters, 1992), There are also reports that for tone-on-tone modulation the AM is easier to detect than the frequency modulation (FM), if in both cases the carrier frequency and modulation frequency differ by less than a half of the critical band and the FM modulation index is less than 1 (Dau, 1997; Zwicker, 1952; Zwicker, Flottorp and Stevens, 1957; Schorer, 1986). In such cases, the spectra of modulated signals differ only by one sideband shifted in phase by 180°; this is the case of very low modulation rates. For higher modulation rates, where the frequency components are separated by more than one critical band, the detectability is the same. This view about the importance of the critical band for detecting phase differences is challenged by the results of some other studies, where the researchers demonstrated that the listeners were able to hear phase changes even if the frequency components were separated by more than one critical band (Lamore, 1975; Patterson, 1987).

Several authors report that humans cannot detect short term phase reversal of the stimulus (Warren and Wrightson, 1981; Sakaguchi, Arai and Murahara, 2000) or that their threshold of hearing is different for rarefaction or condensation clicks (Stapells, Picton and Smith, 1982). However, there are also reports that short clicks may be heard differently depending on their polarity. In addition, there are reports indicating differences in auditory brainstem responses (ABRs) to sequences of condensation and rarefaction clicks (Berlin et al., 1998).

Timbre

Auditory image and timbre

Physical sounds stimulating the auditory system generate *auditory images* in our perceptual space (Letowski and Makowski, 1977; Letowski, 1989). McAdams (1984) defined an auditory image as a "psychological representation of a sound exhibiting an internal coherence in its acoustic behavior." A single auditory image can be analyzed perceptually by a listener focusing attention on the individual sensations or details of the image.

Auditory images are commonly described in terms of loudness, pitch, perceived duration, spatial character (spaciousness), and timbre (Letowski, 1992). The first three dimensions are perceptual reflections of basic physical properties of simple sounds, i.e., sound intensity, frequency, and duration, and have been discussed above. Timbre and spaciousness are multidimensional characteristics carrying information about the sound source and its acoustic environment, respectively.

Timbre has been defined by the ANSI as that attribute of an auditory image "in terms of which a listener can judge that two sounds, similarly presented and having the same loudness and pitch are dissimilar" (ANSI, 1994; Moore, 1997). A footnote to the definition explains that the term 'similarly presented' refers foremost to sound duration and spatial presentation. In a similar definition listed by Plomp (1970) loudness and pitch are supplemented by perceived duration.

In other words, timbre is the characteristic other than loudness, pitch, and perceived duration that makes two sounds perceptually different. Unfortunately, such a definition of timbre is not very useful in practical applications since it tells what timbre is *not*, rather than what timbre *is*. It also makes it unclear whether or not loudness, pitch, and perceived durations are the dimensions of timbre (Letowski, 1989). Therefore, in addition to the standardized, theoretical definition of timbre, many authors introduce another working definition of timbre. According to this definition, timbre is the perceptual property of sound that reflects unique properties of the sound and its sound source. This definition of timbre focuses on the perceptual representation of a specific pattern of temporal and spectral characteristics of a sound resulting from the specific operational principles of the sound source and facilitates identification and storage of auditory images. For example, Roederer (1974) defined timbre as "the mechanism by means of which information is extracted from the auditory signal in such a way as to make it suitable for: (1) storage in the memory with an adequate label of identification and (2) comparison with previously stored and identified information." The basic sensations of loudness, pitch, and auditory duration usually do not convey information about sound source behavior, so the above definition does not seem to contradict the formal definition of timbre. In addition, timbre defined in this way is less restrictive and allows for the differences in loudness, pitch, and auditory duration to be taken into account by the listener in assessing timbre, if needed. In other words, the sounds may differ in any or in all three of those characteristics and still have distinct differences in timbre. It also clarifies the requirement of equal loudness, pitch, and perceived duration in the standardized definition of timbre as an attempt "to bring other dimensions into focus" (Gabrielsson and Sjögren, 1979).

Timbre and pattern recognition

Two physical factors are commonly mentioned as physical correlates of timbre: the spectral envelope and the temporal envelope of the sound. Two complex tones creating the same sensations of pitch and loudness can differ greatly in their spectral content and temporal envelope. An example of such a difference is shown in Figure 11-22, which compares spectral properties of the sounds of guitar, bassoon, and alto saxophone having the same loudness and pitch. The sound of the guitar has a very dense pattern of harmonics, even in the upper range of frequencies, while the sound of the alto saxophone has very little energy above about 5 kHz. It is this spectral pattern that helps one to hear and recognize differences among musical instruments and talkers.

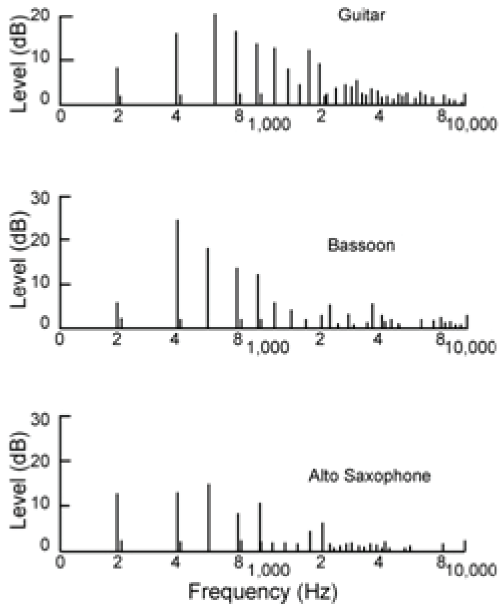


Figure 11-22. Line spectra of three instruments playing a tone with the same pitch and loudness (adapted from Olson, 1967).

It is perceptually easy to differentiate continuous sounds that differ in their spectral pattern. However, it is the temporal envelope of sound that is the main property of sound leading to sound source identification. For example, there are reports indicating that it takes at least 60 ms to recognize the timbre of continuous sound after its onset. Thus, although the differences between stationary sounds can be heard and may lead to general sound source classification (recognition), they are usually not sufficient for sound source identification. To account for this deficiency, stationary timbre is frequently referred to as sound color (noises) or tone color (periodic sounds).

Intensity changes occurring over time form the temporal envelope of a sound. In general, the temporal envelope of an isolated sound includes three distinct portions – the onset (rise), steady state (sustain), and offset (decay). In Figure 11-23, panels (a) and (b), two sound waveforms of a violin tone resulting from the vibration of a violin string actuated by (a) plucking and (b) bowing are shown. Note that the onset is quite abrupt for the plucked tone, and gradual for the bowed tone. Further, the offset begins quite early for the plucked tone; there is very little steady state content. Speech also can be described as a series of spectral changes that create the different phonemes. In Figure 11-23, panels (c) and (d), the waveforms of two different consonant-vowel syllables /*ba*/ and /*wa*/ are shown; they differ only in their initial consonant. The consonant /*b*/ is a voiced stop created by the closing of the vocal tract that produces an abrupt onset similar to that of the plucked violin. The consonant /*w*/ is an approximant, a sound produced by an incomplete constriction of the vocal tract. Although both syllables have nearly the same pitch due to a common fundamental frequency, the other peaks in the spectral content (formants) shift as the utterance shifts from the initial consonant to the vowel, and this causes a timbre difference between the two syllables.

The two examples of different spectral and temporal patterns that result in timbre differences are an indication that timbre is an important perceptual cue in pattern recognition and the dominant cue in differentiating between various music instruments. They also can be used to explain the importance of timbre judgments for sounds that differ in loudness or pitch. The sounds of the same music instrument played in different registers of the instrument may have different timbre, but since these differences in timbre are expected for sounds of different pitch played on this instrument, they are recognized as coming from the same sound source.

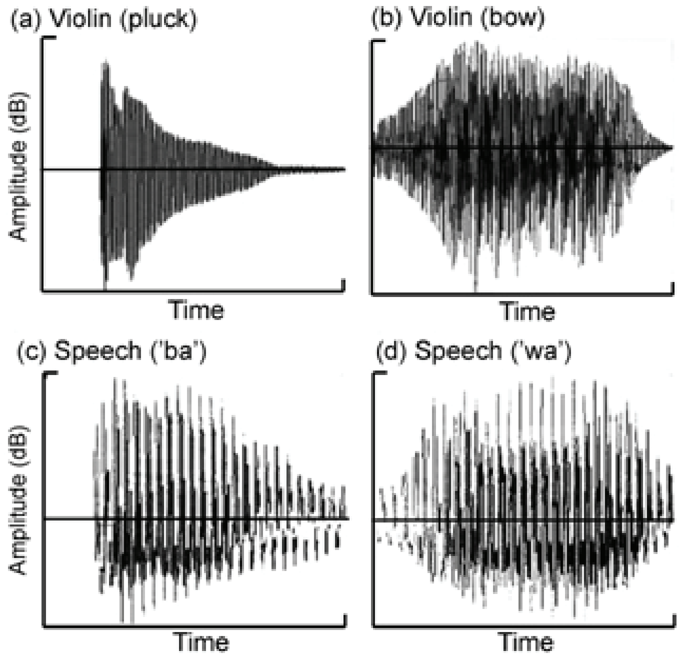


Figure 11-23. Two examples of a violin tone produced by plucking (a) or bowing (b) and of speech (c-d) (see text for details).

Timbre dimensions

Timbre is a sensation of the sound pattern (structure), which together with spaciousness reflecting the environment surroundings of the sound source and the listener, forms an auditory image of the acoustic reality outside of the listener. However, as a complex multidimensional sensation, it cannot be described well by a global assessment alone.

It is very important to realize that the multidimensional character of timbre is not a combination of a small number of well-defined subservient sensations but rather a rich language that consists of a myriad of terms with overlapping or even redundant meaning and that has a number of dialects used by engineers, musicians, psychologists, journalists, and other professional groups. Therefore, due to the richness of timbre terminology, it is necessary to identify and define some basic dimensions of timbre in order to establish universally accepted although limited timbre terminology needed for scientific and human communication purposes.

Many theoretical and experimental studies have been devoted to the identification of the dominant sensations that constitute timbre. The majority of studies used either factor analysis (FA) techniques applied to ratings made on the numerous semantic differential scales or multidimensional scaling (MDS) techniques applied to similarity judgments (Letowski, 1995). The common goal of these studies was to establish a set of meaningful descriptive adjective-based scales permitting quantitative description of timbre changes. For example, Stevens and Davis (1938) and Lichte (1941) investigated timbre dimensionality by using a semantic differential method and identified the following sensations as the main dimension of timbre: loudness, pitch (pitch height), volume, density, brightness (spectral balance), vocality (vowel-likeness), and tonality (strength of pitch). The spatial character (spaciousness) of sound usually was not addressed in the semantic differential and similar studies, with the exception of studies dealing with sound reproduction systems and stereophonic music recording (Eisler, 1966; Gabrielsson and Sjögren, 1979) or the sound character of concert halls (Hawkes and Douglas, 1971). Therefore, although the majority of the proposed systems are limited to the timbre dimensions, there are some systems that are applicable to the overall sound image. Another complicating factor is that in many of these systems sound character (timbre) and sound quality (pleasantness) criteria were mixed together resulting in poorly designed systems. Some examples of the semi-orthogonal linear systems of bi-polar timbre or auditory image dimensions proposed by various authors are listed in Tables 11-9 to 11-12 (Letowski, 1995). The tables list the proposed dimensions and the adjectives defining both ends of the bi-polar scales.

None of the systems listed in Tables 11-9 to 11-12 seem to fully capture the dominant aspects of either timbre or auditory image, but they are listed here as examples of systems available in the literature.

One attempt to identify timbre dimensions involved the division of the spectral range into 18 one-third octave bands, assessing loudness of each of these bands, and defining timbre as a perceptual spectrum of a sound. Another attempt involved creating several perceptual dimensions based on combinations of on-third octave bands and applying them to a specific class of sounds, e.g., vowel sounds (Plomp, Pols and van der Geer, 1967; Pols, van der Kamp and Plomp, 1969; Plomp, 1970, 1976). Such approaches led to several advances in signal processing techniques, but they did not enhance our knowledge of timbre dimensions.

Table 11-9.
The system of timbre dimensions proposed by Bismarck (1974a, 1974b) for the assessment of complex tones.

Dull	Sharpness	Sharp
Compact	Density	Scattered
Empty	Fullness	Full
Colorless	Coloration	Colorful

Table 11-10.
The system of timbre (sound quality) criteria proposed by Yamashita et al. (1990) for the assessment of automotive noises.

Pleasant	Annoyance	Annoying
Weak	Powerfulness	Powerful
Dull	Sharpness	Sharp

Table 11-11.
The system of timbre dimensions developed by Ptratt and Doak (1976).

Dull	Sharpness	Brilliant
Cold	Warmth	Warm
Pure	Richness	Rich

Table 11-12.
The system of auditory image dimensions proposed by Gabrielsson and Sjögren (1979) and Gabrielsson and Lindström (1985) for the assessment of audio systems.

Dull	Sharpness	Sharp
Unclear	Clarity	Clear
Distant	Nearness	Near
Closed	Spaciousness	Open
Dull	Brightness	Bright
Soft	Loudness	Loud
Thin	Fullness	Full
Absent	Disturbance	Present

There were also several attempts, e.g., Solomon (1958), to divide the entire frequency range into a number of bands and assign timbre dimensions to sounds characterized by the dominant energy in each individual band. An example of such a system based on octave bands proposed by Letowski and Miskiewicz (1995) is shown in Table 11-13.

In addition to one-level semi-orthogonal systems of timbre dimensions, there were some attempts to create hierarchical systems in which auditory image or timbre was gradually divided into more and more detailed descriptors forming separate layers of auditory image dimensions (Clark, 1987; Steinke, 1958; Szlifirski and Letowski, 1981). An example of this type of system for two-dimensional auditory images, called MURAL, proposed by Letowski (1989), is shown in Figure 11-24.

Table 11-13.
A system of timbre dimensions for description of stationary sounds (Letowski and Miskiewicz, 1995).

Center frequency of the octave band (Hz)	Timbre Dimension
63	Boom
125	Rumble
250	Powerfulness
500	Hollowness
1000	Nasality
2000	Presence
4000	Sharpness
8000	Brilliance
16000	Rustle

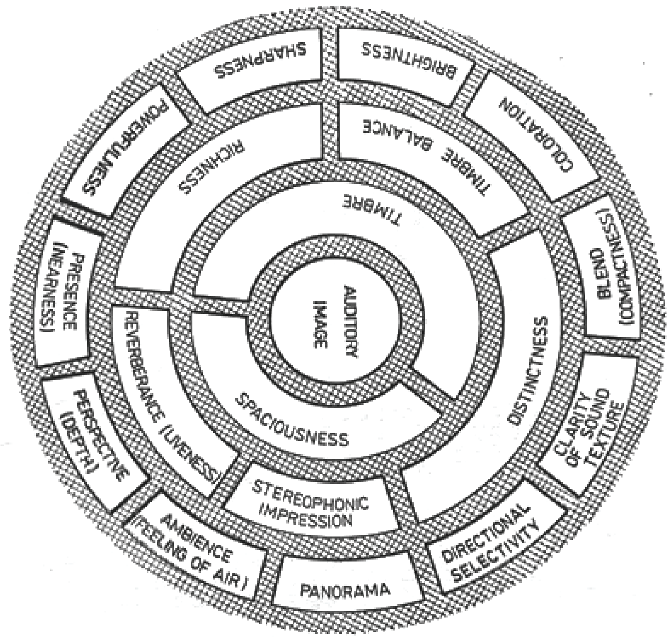


Figure 11-24. **MU**ltilevel audito**R**y **A**ssessment **L**anguage (MURAL) for timbre and sound quality assessment (Letowski, 1989).

Sound Quality

It should be recognized that the effects of auditory stimulation involve not only quantitative judgment of sensations and the subsequent perception of the acting stimulus, but also the emotional judgment of the stimulus' aesthetic value (beauty) and the assessment of the degree of the listener's satisfaction (utility). These types of judgments are together called the sound quality judgments.

Sound quality can be broadly defined as a set of properties of a given sound that determines the capability of the sound or its source to fulfill a particular function or need. As defined above, the sound quality may be either objective (technical) or subjective (perceptual). If the sound quality fulfills a perceptual need, it is sometimes called *perceived sound quality* (PSQ) to clearly identify its origin.

Letowski (1989) described PSQ as the emotional aspect of the overall auditory image. One auditory image is not better than another, they are just different. However, one auditory image may fit better a particular need than another or be closer to the desired standard than another. This underlines the basic difference between the *sound character*, expressed in terms of auditory image, timbre, spaciousness, and a multitude of other auditory sensations (e.g., roughness, breathiness, or ambience) and the *sound quality*. Sound character is expressed on scales from *more* to *less*, while sound quality is expressed on scales from *better* to *worse*.

There are two fundamental forms of sound quality assessment: global assessment and parametric assessment. Global assessment of sound quality can be made according to one of the three basic aspects of quality:

- *Fidelity* (accuracy), which reflects similarity of a given auditory image to a specific auditory standard or to another auditory image,
- *Naturalness*, which reflects an agreement of a given auditory image with general expectations of the listener, and
- *Pleasantness*, which reflects the degree of the listener's satisfaction with a given auditory image.

All these aspects of sound quality also may be expressed in terms of the opposite end of the respective perceptual scale, i.e., inaccuracy (instead of fidelity), awkwardness (instead of naturalness), and annoyance or unpleasantness (instead of pleasantness). Focusing on the positive or negative end of the scale allows the ability to differentiate better small differences between stimuli occupying a particular end of the scale, but it does not change the general order (ranking) of the stimuli quality (Letowski and Dreisbach, 1992).

Note that while an auditory image or timbre cannot be assessed globally on a more-less sound character scale, sound quality, as expressed above, can be assessed globally on a better-worse quality scale. Whether the assessment has a form of fidelity, naturalness, or pleasantness depends entirely on the application of such judgment. Audio HMD assessment may be performed with either of these criteria, however, in most practical cases, it will be done in the form of a fidelity assessment of transmitted speech (speech intelligibility), spatial auditory image (localization accuracy), or sound source identification (signature identification).

While the value of the global assessment of sound quality should not be underestimated, such assessment does not provide information about specific aspects of the auditory stimulus. Recall the multidimensional character of the auditory image and timbre discussed above. The multidimensional character of sound requires multidimensional (parametric) assessment of its sound quality in such processes as audio equipment design or selection. In order to conduct parametric assessment of sound quality, one of the timbre or auditory image dimensions subsystems discussed in the section on timbre dimensions, or any other arbitrary or experimental selection of auditory dimensions, can be used. The data collection process may have two forms: (1) classical assessment of timbre and spaciousness of a number of subservient more-less sound character scales or (2) on the same system of scales converted to better-worse sound quality scales. In the first case, the users (listeners) make classical psychophysical judgments, and the designers or researchers interpret the data as subject matter experts (SMEs) and make the decision whether more or less is good or bad. In the second case, the users themselves make these decisions.

An example of a dedicated system of objective criteria for parametric sound quality assessment is the system of metrics proposed by Zwicker and his coworkers. This system consists of one global assessment scale (sound pleasantness or annoyance) and five subordinate dimension scales: loudness, sharpness, fluctuation strength, roughness, and tonality (Aures, 1984, 1985; Bismarck, 1974b; Terhardt, Stoll and Seewann, 1982; Zwicker and Fastl, 1999). Although all the above criteria have the same names as the perceptual dimensions of the auditory image they are only certain approximations of the perceptual dimensions and are frequently referred to as, for example, calculated roughness rather than roughness to stress their objective character. Calculated loudness, sharpness, fluctuation strength, roughness, and tonality are briefly described in Table 11-14.

Various implementations of listed above calculated sound quality metrics are available in many major sound analysis software packages (e.g., PULSE by Bruel and Kjaer, dBFA32 by 01dB, Artemis and SQLab II by HEAD Acoustics, DATS by Prosig). Other objective metrics of sound quality proposed in literature include, among others, booming (sound level in 22.4 to 224 Hz band) and impulsiveness (crest factor). It may also be helpful to indicate that there are two basic methods to calculate the tonality (pitch strength) of sound used in sound analysis systems. The first method uses the concept of *tone-to-noise ratio*, defined as the ratio of the power of the tone of interest to the power of the critical band centered on that tone (excluding the tone power) (ANSI, 2005). Usually the tone is audible at tone-to-noise ratios above approximately -4 dB. Recall that noise within the critical band is masking the tone, so this is more a measure of the effective SNR than a measure of tonality. The second method uses the concept of *prominence ratio*, defined as the ratio of the power in the critical band centered on the tone of interest to the mean power of the two adjacent critical bands (ANSI, 2005). According to this metric, a tone is prominent if this ratio is above 7 dB. Neither of these metrics says much about whether the sound is perceived as a coherent tone, but rather whether a noisy

Table 11-14.
System of objective sound quality metrics developed by Zwicker and his coworkers (Zwicker and Fastl, 1999)

Dimension	Definition	Comments
Loudness	Perceptual impression of the intensity of sound.	See section on loudness. The unit of loudness is *sone*
Sharpness	Sensation caused by acoustic energy concentrated in a narrow band around relatively high center frequency of sound; perceptual metric related to the spectral center of gravity.	The unit of sharpness is *acum* (Latin for sharp). One acum is defined as the sharpness of a 1 kHz narrowband (one critical band wide) noise at 60 dB SPL.
Roughness	Perceptual impression created by amplitude and frequency modulations in sound at high modulation rates, above about 20 Hz. Roughness notably decreases for modulation frequencies higher than about 50 Hz (Terhardt, 1974).	The unit of roughness is *asper* (Latin for rough). One asper is defined as the roughness of 1 kHz tone at 60 dB SPL that is 100% modulated at 70 Hz. (Aures, 1985)
Fluctuation strength	Perceptual impression created by amplitude and frequency modulations in sound at low modulation rates, up to about 20 Hz. The greatest amount of fluctuation strength is perceived at modulation frequency of 4 Hz.	The unit of modulation strength is *vacil (*Latin for vacillate). One vacil is defined as the fluctuation strength of a 60 dB SPL, 1 kHz tone 100% amplitude modulated at 4 Hz.
Tonality	Degree to which a sound has a distinct pitch; strength of pitch.	See section on pitch. The unit of tonality is *tu* (tonality unit). One tu is defined as tonality of 1 kHz tone at 60 dB SPL
Annoyance	Combination of sharpness, fluctuation strength, roughness, and loudness.	Global assessment of sound quality The unit of calculated (unbiased) annoyance is *au*.
Pleasantness	Combination of roughness, sharpness, tonality and loudness.	Global assessment of sound quality.

Perceived Duration

Acoustic events may appear perceptually shorter or longer than the actual physical events. This phenomenon is generally described as *time distortion* or time warping, and the amount of time assigned by a person to a specific physical event is called *perceived duration*. In the case of long lasting events, exceeding several seconds, perceived duration is primarily dependent on emotional state, expectations, and activity of a person and is very difficult to generalize. The only rule that can be generally applied is that pleasant events appear to last shorter (time contraction) and unpleasant events longer (time dilation) than their actual physical durations.

In the case of very short acoustic events, humans have a general tendency to overestimate sound duration, and the amount of overestimation seems to be inversely proportional to the actual duration of the sound. When the sound duration exceeds about 200 to 300 ms and is less than several seconds, perceptual duration is very close to physical duration and generally assumed to be identical (Zwicker and Fastl, 1999).

One important condition where perceived duration differs greatly from the physical duration is perception of short silent intervals. Again, if the intervals are longer than several hundred milliseconds (500 to 1000 ms) but shorter than a few seconds the perceived duration and physical duration are about the same. Similarly, for shorter pauses, the pause duration is overestimated. However, short pauses seem to last as much as 2 to 4 times longer than short sound bursts of the same duration (Zwicker and Fastl, 1999). The higher the frequency of the sound, the greater this perceptual difference. This perceptual difference has direct impact on music perception (rhythm perception) as well as the design of periodic, fast-rate changing, signals for industrial and military applications.

Time Error

The duration of the gap between two stimuli is not only an object of detection itself, but it also moderates the effect that separated stimuli may have on each other. When the stimuli are separated by a very short period of time, they are subjected to the effects of temporal masking. If they are far apart, their comparison is affected by the decaying memory trace of the first stimulus. These phenomena are not unique to audition but occur throughout human perception.

The error in sensory judgment resulting from sequential presentation of stimuli is referred to as the *time error* (TE) or sequential error. The TE was originally observed and described by Fechner (1860) and has been studied extensively for more than a century. The type and size of TE depends on the duration of the gap between the stimuli, duration of stimuli, and the property being judged (Hellström, 1977). In the case of short time gaps when the TE is primarily a result of a forward masking, the TE is positive (+TE). In the case of long time gaps, when the time error is due to decaying memory trace of the first stimulus, the TE is negative (-TE). The duration of the time interval between the stimuli when +TE changes into –TE has been of great interest to psychologists because such stimulus separation seems to eliminate the need for consideration of TE in comparative studies.

Köhler (1923) investigated the effect of temporal gap on comparative judgment of loudness and observed -TE for gap of 1.5 seconds and +TE for gaps of 6 and 12 seconds. Based on these observations, he concluded that the optimum time gap for pair comparison of loudness should be about 3.0 seconds. The results of later studies by Needham (1935) and Pollack (1954) shortened this time to about 1.0 to 1.5 seconds.

According to Stevens (1956, 1957), the TE in pitch comparison should be very small or not present due to the relative (metathetic; associated with a quality) character of pitch as opposed to loudness that has an absolute (prothetic; associated with the quantity) character. Small TE values for pitch also were reported by Köenig (1957) who observed that the optimum gap duration for pitch comparisons were the same as for loudness comparison. Other studies concluded that as long as the temporal gap is within 0.3 and 6.0 seconds, the effect of TE on pitch perception seems negligible (Jaroszewski and Rakowski, 1976; Koester, 1945; Massaro, 1975; Postman, 1946; Truman and Wever, 1928). A similar conclusion was reached for the comparative judgment of auditory brightness, a timbre dimension very close in its character to pitch, by Letowski and Smurzyński (1980). Note that these gap durations are about the same as the silent intervals, which durations are perceived without substantial time distortions.

Unlike the rather wide range of time gaps that can be used for pitch and brightness comparisons, successive presentation of complex sounds for sound quality assessment may require gaps similar to those used for loudness comparisons (Letowski, 1974). Qualitative assessment of complex, usually time-varying sounds, seems to require shorter temporal gaps, as the listener tends be biased toward the “preferential” treatment toward the second stimulus (Choo, 1954; Brighouse and Koh, 1950; Koh, 1962, 1967; Saunders, 1962). In addition, regardless of whether the judgment is quantitative or qualitative, longer temporal gaps between signals lead to larger variability in listener judgments (Bindra, Williams and Wise, 1965; Shanefield, 1980).

Speech Perception

Speech is a system of sounds produced by the vocal system that allow human-to-human communication. Simple sounds, called phonemes, are combined together in more complex structures (strings) to convey thoughts, feelings, needs, and perceptions. Small structures are combined in larger and larger structures, called syllables, words, phrases, sentences, and stories, respectively, depending on the complexity of the intended message. Each spoken language has a certain limited number of phonemes that form the basis of speech communication and has a practically infinite number of higher order structures that can be constructed with these phonemes.

Liberman et al. (1967) stated that the perception of speech is different from the perception of all other complex sounds because it is mentally tied up to the process of speech production. However, if the speech sounds are unfamiliar to the listener (e.g., listening to an unknown foreign language), the speech looses its special character caused by coupling between speech perception and speech production of the listener; such sounds should be treated as non speech sounds.

Speech production

Speech sounds can be spoken or sung, especially the voiced phonemes. Depending on the range of frequencies that the singer can produce, singer voices are typically classified as soprano, mezzo-soprano, and alto (female voices) and tenor, baritone, and bass (male voices), starting from the highest through the lowest. In addition to spoken and sung speech sounds, human vocal production includes whistling, crying, murmuring, tongue clicking, grunting, purring, kissing sounds and laughing.

The process of speech production is called articulation and involves the lungs, larynx and vocal folds, and vocal tract. The vocal tract is the air tube that begins at the mouth's opening and ends at the vocal folds with branches off to the nasal cavity. In the process of speech production the stream of air controlled by the lungs and vocal folds is processed by the set of three articulators located in the mouth cavity – tongue, teeth, and lips – and becomes a string of speech sounds, i.e., phonemes. The process of combining phonemes into larger structures, i.e., the process of chaining the phonemes together into strings, is called *coarticulation*.

Two basic classes of phonemes are vowels and consonants, which can be divided further in many subclasses depending on the form and degree of activation of the vocal folds and mouth articulators. The vowels are usually classified based on the tongue position and lips openness. The consonants are classified on the basis of their voicing (voiced and unvoiced), place and manner of their production. All vowels and voiced consonants are the results – but not solely – of the acoustic filtering by the vocal tract of the saw tooth-like periodic waveform generated by vocal folds in a process of phonation. The momentary positions of speech articulators during the process of phonation divide the vocal tract into a series of resonance tubes and cavities that produce local concentrations of energy in the spectrum of output signal. These concentrations are called formants, and their relative positions on the frequency scale identify individual vowels. Vowels are very important to speech production, but it is the consonants, i.e., the very movement of articulators, which make the speech rich in meanings and contexts.

In addition to the factors discussed above, the emotional and ecological conditions during speech production lead to various forms and levels of speech: a soft whisper (30 to 40 dB SPL), a voiced whisper (40 to 55 dB SPL), conversational speech (55 to 65 dB SPL), raised speech (65 to 75 dB SPL), loud speech (77 to 85 dB SPL), and shouting (85 to 95 dB SPL). These values correspond to the sound pressure levels at about 1 meter (3.28 feet) from the talker's lips. Directly at the lips, these values are much higher. A list of selected basic factors affecting speech production is presented in Table 11-15.

The Lombard effect (Lombard, 1911) is a phenomenon in which a talker alters his or her voice in noisy environments. Generally, there is an increase in vowel duration and voice intensity (Summers et al, 1988; Junqua, 1996). In addition, Letowski, Frank and Caravella (1993) reported changes in the fundamental frequency of the voice (male voices) and spectral envelope of the long term spectrum (female voices). These changes to the speech produced in noise are most likely caused by the talker's attempt to improve audibility of the sidetone (i.e., audibility of the talker's own voice) and result in improved speech intelligibility (Lane and Tranel, 1971; Letowski, Frank and Caravella, 1993). There observations are in agreement with the reports that signal processing techniques that replicate the Lombard effect improve the intelligibility of speech in a noise environment (Chi and Oh, 1996). However, the human tendency to alter speech in this way is largely automatic (Pick et al., 1989), and individuals has no control over the Lombard effect. The existence of the Lombard effect also affects the accuracy of speech recognition software (Junqua, 1993). Because of this, the presence of the Lombard effect is worth considering when designing audio HMDs that will be used in conjunction with speech recognition software in noisy environments.

Table 11-15.
Basic factors affecting talker's speech production.

Factors Affecting Speech Production
Fundamental frequency of the voice
Language (primary vs. secondary)
Articulation and coarticulation
Breathing (emotions)
Vocal effort (whisper to shout)
Auditory feedback (sidetone)
Ambient noise (Lombard effect)
Hearing loss of the talker

Speech communication

Speech communication refers to the processes associated with the production and perception of sounds used in spoken language. Humans are able to understand speech produced by an infinite variety of voices in an infinite variety of combinations. However, individuals differ in their hearing ability and language proficiency, environments are noisy or unpredictable, and equipment supporting speech communication may be noisy or problematic.

The highest level of speech understanding is referred to as speech comprehension. Speech comprehension is a function of environmental conditions, the communication channel and its capacity, and peripheral hearing ability and higher order cognitive factors of the listener. Speech comprehension can only be approximated, and the process is time consuming and tedious. A few such tests have been developed, but are not commonly used.

Speech recognition (SR) is a lower level of speech understanding. SR is the human ability to understand speech. It is measured by the percent of correctly recognized speech items (phonemes, syllables, words, phrases, or sentences). The result can be expressed as percent correct responses for the whole speech test (speech recognition score), as a speech intensity level for which a person is able to recognize 50% of the speech items (speech recognition threshold), or as a speech level at which an individual is able to recognize 50% of the test items as speech (speech detection level).

The two lowest levels of speech understanding are speech discrimination and speech detection. Speech discrimination tests are used very rarely, and they are intended to measure the degree to which a person is able to hear the differences between speech items, even if they are meaningless. One practical application of speech discriminations tests is prediction of potential problems in acquiring a second language. Various pairs of phonemes or syllables in a new language are played to a person before the language training begins to determine which sounds would be the most difficult for this person to differentiate and, because of it, clearly produce.

The *speech detection threshold* (SDT), frequently referred to as the *speech awareness threshold* (SAT), has been introduced above in the section about the air conduction threshold. This metric is used mainly to determine minimum required levels of masking stimulus to mask speech, e.g., in open office situation. SDTs in quiet and in noise are used also for testing human ability to hear speech for those who does not speak a given language and in lieu of more time consuming tonal audiometric tests to roughly assess hearing threshold of a person in a given environment.

Various speech communication terms used in speech communication testing are shown in Figure 11-25. Speech recognition is a measure of a person's ability to hear speech. Speech articulation is a measure of the clarity of speech production. Speech transmission is the measure of the effect of a communication channel on the clarity of speech. These three basic elements of speech communication assessment may be combined in different configurations resulting in speech intelligibility encompassing speech articulation and transmission or speech audibility encompassing speech transmission and recognition.

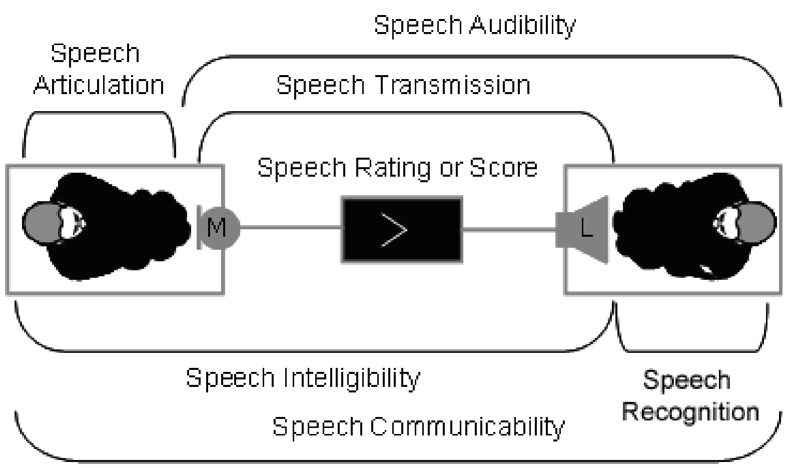


Figure 11-25. Speech communication terminology used in the assessment of the effects of various elements of the speech transmission chain on speech communication.

For example, *speech intelligibility* (SI) is the understanding of speech in a particular environment by expert listeners with normal hearing. SI testing is used to quantify the operational conditions of a speech communication environment in order to determine whether there are problems that threaten the transmission of spoken information. Speech intelligibility is affected by imperfect speech production and by properties of the communication channel between the talker and the listener, including environmental conditions surrounding both the talker and the listener. It varies as a function of SNR, reverberation time, rate of speech and other factors. It is measured usually as a word recognition score for a given transmission system or environment but it can be applied to sentences and connected speech as well.

In many cases, neither the talker's characteristics, environmental conditions, nor the listener's characteristics are ideal, and it is necessary to capture human ability to communicate under these conditions. Such speech tests are referred to in this chapter as speech communicability tests (Figure 11-25). Note, however, that regardless of the specific part of the communication chain being assessed, the same physical speech tests may be used for data collection. There is a very large selection of speech tests that differ in their redundancy, complexity, and vocabulary and result in fairly different test scores for the same auditory environment. Therefore, it is important that speech communication data are reported together with the name of the speech test used for the data collection and the name of speech communication measure to document what was actually measured and how. Unfortunately, there is a general lack of terminological discipline among the people developing speech tests and conducting speech assessments, and the described terms are frequently misused.

There are a number of perceptual tests of speech recognition, intelligibility, or communicability including perceptual tasks of recognition, discrimination, and detection of speech. In addition, speech intelligibility (clarity) can be also rated on the scale from 0% to 100%. This test procedure is called a *speech intelligibility rating* (SIR) test and may be applied not only to intelligibility testing but also to the other forms of speech assessment shown in Figure 11-25. It is a fast data collection procedure that provides data highly correlated with measures requiring much more effort and time consuming scoring procedures.

Speech perception and environment

The primary environmental effects on speech communication are those of noise and reverberation. Good understanding of speech requires high SNRs in the order of 15 to 30 dB. Smaller SNRs lead to reduced speech intelligibility scores. The exact SNR level required to achieve minimal required speech intelligibility depends on the speech material, type of noise, the acoustic environment, and the listeners themselves.

As long as the SNR is sufficiently high to allow enough of the speech signal to be heard in the noise, the absolute level of the noise has a minimal effect on speech understanding as long as the noise levels are below 85 dB SPL. Conversely, speech understanding depends to a large degree on the type of background noise. As discussed earlier in the section on masking, steady-state broadband noise causes primarily energetic type of masking. As long as a sufficient proportion of the speech energy is audible, speech is heard and understood. However, random, unpredictable noise, or noise where the temporal and spectral characteristics are similar to that of speech, can add informational masking to the energetic masking. Therefore, the most efficient masker of speech is other speech, such as a multitalker noise including a moderate number of voices. An example of functional relationship between speech recognition score and SNR is shown in Figure 11-26.

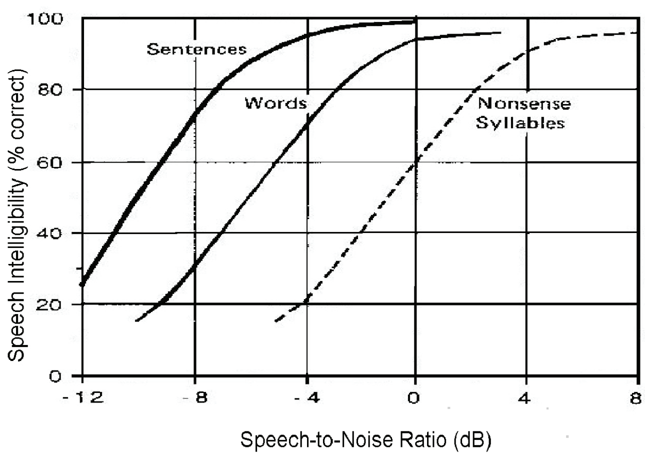


Figure 11-26. The effect of speech-to-noise ratio (SNR) on intelligibility of nonsense syllables, words, and sentences. Adapted from Levitt and Webster (1997, Figure 16.3).

An acoustic environment can also add reverberation to the speech signal. Reverberation consists of multiple reflections of the sound that mask the direct sound. In the case of speech, the masking is simultaneous and temporal. Thus, it adds noise to and alters the spectro-temporal envelope of the original speech signal. Some very large spaces have a sound decay time (i.e., reverberation times) to the order of 5 seconds and higher, especially at low frequencies, which can make normal speech communication in these spaces virtually impossible.

The spatial configuration of the speech source and the noises source(s) can also have an effect on intelligibility. If the talker and the noise are located in the same location, the listener can only use spectral and temporal characteristics of the two signals to parse the two signals. This is true also if the listener only has a monaural signal (e.g., phone and radio). If the two signals are separated in space, binaural cues can be used to separate the two locations, and the listener can selectively attend to the target speech signal.

All of these environmental characteristics affect speech intelligibility and can often be quantified and described to some degree. However, cognitive features of the speech message also affect comprehension to some degree as well. The most notable of these is known as the “cocktail party phenomenon” where the listener embedded in “party noise” can clearly hear his or her own name when spoken, even though other speech may not be audible (for a more thorough discussion of cocktail party effect, see Bronkhorst, 2000).

Speech recognition

The term *speech recognition* (SR) is often confusing because it has two related by separate meanings. In its narrow sense, it is a metric that provides information about individual’s ability to hear speech as shown in Figure 11-25. In its broadest meaning it is the score on any speech test regardless of the specific type of communication assessment. For example, one can use a speech recognition test to assess speech articulation, speech recognition, speech transmission, or speech communicatibility. It is this second, broader, meaning of the *speech recognition* term that we use throughout the rest of this chapter.

The *SR score* and *speech recognition threshold* (SRT) are two basic metrics of speech recognition. They are used to characterize SR ability of an individual listener under specific test conditions but in practice they are also dependent on speech material, the talker’s voice, and many procedural factors. However, as long as these test elements are kept constant, any speech tests can be used as a relative measure of human capabilities. It is the predictive value of the speech test for the specific operational conditions that makes various test more or less appropriate. It is important to recognize that all speech tests data are limited by the degree to which selected speech material is representative of the speech vocabulary and speech structures used in the operational environment for which performance is being predicted.

The SR score is simply the percentage of speech material understood by the listener. The ANSI standard S3.5 – 1997 (revised in 2007) (ANSI, 1997) gives speech recognition scores for a number of commonly accepted perceptual speech recognition measures and compares them to objective measures described later in this section

Basic test conditions and test material for SRT testing are addressed by ANSI standard S3.6, “Specification for Audiometers” (ANSI, 2004). As the speech test complexity gets lower and the background noise gets quieter, the SRT decreases. A number of other factors also affect SRT level. First, individuals differ in their hearing sensitivity. Hearing loss due to trauma and age typically occurs in the range of frequencies containing information about consonants. Second, there will be an effect on scores due to whether the speech material consists of syllables, words or sentences, as there is more disambiguating information for longer speech items. Third, the size of the speech vocabulary and the type of speech information used can affect both the SRT level and the SR score. If the operational vocabulary is relatively small and the items within the set phonemically distinct, scores will remain high and SRT levels low even under poor listening conditions. Further, if the items are presented as a closed set (e.g., multiple choice and limited options), performance will be higher than if the items are presented as an open set. Fourth, the quality of the speech presentation will have an effect on recognition performance. If speech material is presented over high-fidelity audio display equipment, scores will be higher than if it were distorted by poor quality radio-type transmissions with low bandwidth. Finally, the spatial arrangement of the speech source relative to that noise also will affect the degree of masking.

Speech perception tests

If speech transmission is to be characterized in terms of performance in a specific environment, either perceptual or objective, microphone-based predictive speech tests can be used. Perceptual measures entail the presentation of speech material at one or more fixed intensity levels to a group of listeners. Performance is given as the average percent correct recognition. Objective measures predict intelligibility by calculating an index based on a recorded sample of ambient environment. Perceptual measures are limited by the speech materials used, the talkers presenting the materials, and the listener sample involved in the study. Figure 11-27 graphs the relationship of scores obtained in a range of common perceptual tests to an objective measure of speech intelligibility called the *articulation index* (AI). The AI will be discussed below, but note that for a particular AI value, performances on the perceptual measures vary widely.

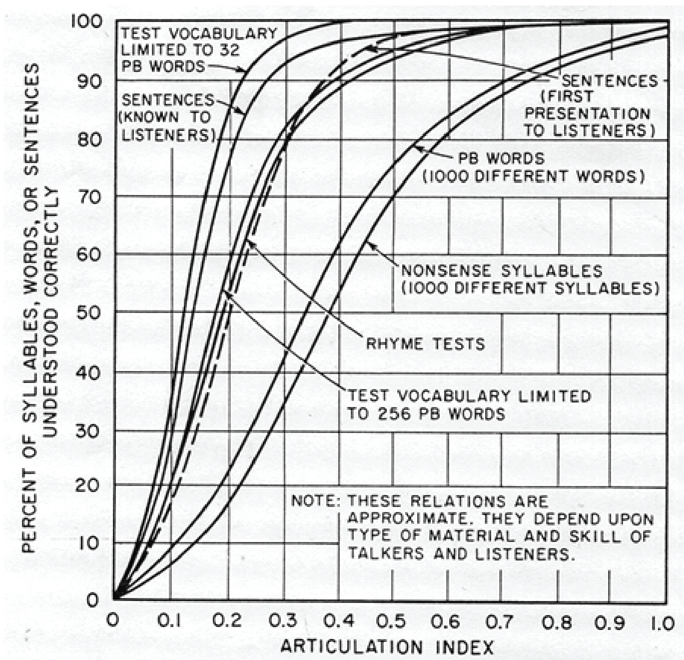


Figure 11-27. Relationship between various perceptual measures of speech intelligibility and articulation index (AI) (after ANSI S3.5 [1997]).

Speech intelligibility performance generally improves as the material becomes more complex and contains more contextual information and higher degree of redundancy (Egan, 1948; Hochhaus and Antes, 1973; Hudgins et al., 1947; Miller, Heise and Lichten, 1951; Rosenzweig and Postman, 1957). Representative tests used in architectural acoustics, communications, and audiology are included in Table 11-16. The tests are classified according to the speech material used in the test and whether the number of alternative answers to the test item was finite (closed set test) or infinite (open set text).

The tests listed in Table 11-16 differ not only by the speech units used for testing and by the open or closed set of possible answers but may also differ in the way they are administered. For example, they may differ by the presence or absence of carrier phrases (word tests), monotone or natural presentation (spondee, phrase, and sentence tests), recorded or live voice administration, and many other technical elements. Therefore, it is important that in a comparative evaluation not only that the same test is used for all audio HMD systems or listening conditions under comparison, but also that it is administered in the same manner.

Table 11-16.
A listing of speech tests using speech material of various degrees of speech complexity.

Speech Unit	Name of Test	Citation
digits and words (open set)	SDS Speech Discrimination Scale	Gaeth, 1970
Oo, ah, ee, sh, M, sss (open set)	Five Sound Test	Ling, 1978
phonemes (closed set)	Test of Phoneme Identities	Murray et al. 2000
phonemes (closed set)	Keating-Manis Phoneme Deletion Test	Keating andManis, 1988
nonsense syllables (closed set)	CUNY-NST Nonsense Syllable Test	Resnick et al., 1975
nonsense syllables (closed set)	DFD Distinctive Feature Difference Test	Feeney and Franks, 1982
words (open set)	PB-50 Phonetically Balance (PB) Word Lists	Egan, 1948
words (open set)	PBK-50 PB Kindergarten Lists	Haskins, 1949
words (open set)	W-22 CV, VC, and CVC Words Test	Hirsh, 1952
words (open set)	CNC Consonant-Nucleus -Consonant Test	Lehiste and Peterson, 1959
words (open set)	NU-6 Northwestern University Lists	Tillman et al., 1963
words (open set)	ABL Test	Boothroyd, 1967
words (open set)	HFCDT High Frequency Consonant Discrimination Test	Gardner, 1971
words (open set)	SPRINT Speech Recognition in Noise Test	Cord et al., 1992
words (closed set)	RT Rhyme Test	Fairbanks, 1958
words (closed set)	Fry Lists	Fry, 1961
words (closed set)	MRT Modified Rhyme Test	House et al., 1965
words (closed set)	DRT Diagnostic Rhyme Test	Voiers, 1975; 1983
words (closed set)	DIP Discrimination by Identification of Pictures Test	Siegenthaler and Haspiel, 1966
words (closed set)	RMC Rhyming Minimal Contrasts Test	Griffiths, 1967
words (closed set)	MCDT Multiple Choice Discrimination Test	Schultz and Schubert, 1969
words (closed set)	WIPI Word Intelligibility by Picture Identification Test	Ross and Lerman, 1970
words (closed set)	OUCRT Oklahoma University Closed Response Test	Pederson and Studebaker, 1972
words (closed set)	CCT California Consonant Test	Owens and Schubert, 1977
words (closed set)	Perception of Words and Word Patterns	Erber and Witt, 1977
words (closed set)	NU-Chips Children's Perception of Speech	Katz and Elliott, 1978
words (closed set)	DFDT Distinctive Features Discrimination Test	McPherson and Pang-Ching, 1979

Table 11-16. (Cont.)
A listing of speech tests using speech material of various degrees of speech complexity.

Speech Unit	Name of Test	Citation
words (closed set)	ANT Auditory Numbers Test	Erber, 1980
words (closed set)	Picture Identification Task	Wilson and Antablin, 1980
spondees	Psycho Acoustic Laboratory (PAL) List #9	Hudgins et al., 1947
spondees	CID W-1 and W-2 (Spondaic Word Lists)	Hirsh et al., 1952
sound effects (closed set)	SERT Sound Effects Recognition Test	Finitzo-Hieber et al., 1980
sentences (open set)	Sentence Tests for Deaf People	Fry and Kerridge, 1939
sentences (open set)	CID Central Institute for the Deaf Test	Silverman and Hirsh, 1955
sentences (open set)	Fry Revised Sentence Test	Fry, 1961
sentences (open set)	SPIN Speech Perception In Noise	Kalikow et al., 1977
sentences (open set)	R-SPIN Revised Speech Perception In Noise	Bilger, 1984
sentences (open set)	BKB (Bamford-Kowal-Bench)	Bench et al., 1979
sentences (open set)	HINT Hearing in Noise Test	Nilsson et al., 1994
sentences (closed set)	KSUDT Kent State University Discrimination Test	Berger, 1969
sentences (closed set)	SSI Synthetic Sentence Identification Test	Speaks and Jerger, 1965
sentences (closed set)	PSI Pediatric Speech Intelligibility Test	Jerger et al., 1980
sentences (closed set)	CRM Coordinate Response Measure (AFRL)	(Brungart, 2001)
sentences (closed set)	NSMRL Tri-Test of Intelligibility	Sergeant et al., 1981
sentences (closed set)	DSI Dichotic Sentence Identification Test	Fifer et al., 1983
phrases (close set) 2 word-3 syllable phrases	CAT Callsign Acquisition Test	Rao and Letowski, 2004; Blue, Ntuen and Letowski, 2004
passages of speech (9-10 sentences)	Connected Speech Test (CST)	Cox et al., 1987

Tests that differ in units of speech and sets of available responses differ also in the test difficulty. Usually, open set tests are more difficult then closed set tests, and tests using meaningless syllables or sentences are more difficult than tests using meaningful items. For example, percent correct scores on nonsense syllables may not exceed 70% correct even at very high SNRs (Miller, Heise and Lichten, 1951). In contrast, scores on words in meaningful sentences may reach 100%, even at moderate SNR, and scores on digits may reach this limit at SNRs as low as 6 dB. The context existing in a meaningful sentence provides information about what kinds of words would be probable at a given place in the sentence, effectively limiting the listener's choices. Thus, even if a particular word is partially masked, enough information is available for the listener to fill in the missing information. In the case of the digits, the listener is limited to 10 or even less available numerical digits and has a high probability of guessing correctly even if a part of the digit is masked or distorted.

As discussed previously in this section, speech understanding depends, among other things on the talker. A trained talker such as radio announcer who speaks clearly will be more intelligible than a talker using normal conversational speech. Clear speech has been found to be slower, both the phonetic components and the pauses between words are drawn out more (Picheny, Durlach and Braida, 1985, 1986, 1989; Uchanski et al., 1996). Usually, clear speech is used for test materials; however, most speech in operational environments is conversational and will not be as intelligible. If testing is done using recordings made of a trained speaker using clear speech, it will probably overestimate performance in most settings. Further, there is a large difference between the intelligibility of different talkers (Black, 1957; Hood and Poole, 1980; Bond and Moore, 1994). For example, a female voice is generally more intelligible than a male voice (Bradlow, Torretta and Pisoni, 1996). Therefore, it is important to use several talkers in validating communication effectiveness of audio HMDs. The current ANSI S3.2-1989 standard for speech intelligibility testing requires that the number of talkers is at least equal the number of listeners.

It is important to recognize that the training and hearing sensitivity of the listener also affect speech intelligibility performance. Trained listeners who are familiar with the test and the speech material to be tested will perform best and have the most reliable scores (Hood and Poole, 1980). Listeners who have impaired hearing will perform differently than normal hearing counterparts, even if the average intensity levels are above threshold (Ching, Dillon and Byrne, 1998). It needs to be stressed that hearing loss is common in those working in high noise environments. Therefore, a measure of the speech intelligibility of a particular environment obtained using normal hearing listeners and professional talkers may overestimate performance by operators in that environment.

Speech intelligibility criteria used by the U.S. Army are listed in he MIL-STD-1472F (Department of Defense, 1999) and presented as Table 11-17. The criteria list the Phonetically Balanced (PB) Word List[7] and Modified Rhyme Test (MRT)[8] perceptual test scores and calculated value of the AI. The criteria are intended for voice communication over a transmission system such as radio and intercom and apply to audio HMDs. The criteria listed in the second row of Table 11-17 are acceptable for normal operational environments. AI should only be used to evaluation of natural speech but not synthetic speech, because some key acoustic features of speech are not present in synthetic speech.

All perceptual speech intelligibility measures discussed above and the speech tests listed in Table 11-16 are influenced by a number of factors that limit applicability of their scores to the operational environments in which they were obtained. Their results may be generalized to other similar environments but not to environments that are very different from the one that was selected for testing. Further, perceptual studies are costly in terms of time

---

[7] In the Phonetically Balanced Word Lists, the monosyllabic test words are chosen so that they approximate the relative frequency of phoneme occurrence in each language (Goldstein, 1995).

[8] The modified Rhyme Test is a word list for statistical intelligibility testing that uses 50 six-word lists of rhyming or similar-sounding monosyllabic English words. Each word is constructed from a consonant-vowel-consonant sound sequence, and the six words in each list differ only in the initial or final consonant sound. Listeners are shown a six-word list and then asked to identify which of the six is spoken by the talker.

and the number of persons required when obtaining speech intelligibility performance data. Therefore, it is sometimes preferable to estimate the effect of the specific environment on speech intelligibility on the basis of physical measurements of the operational acoustic environment. Such measures do not eliminate the need for final assessment of speech intelligibility using perceptual speech intelligibility tests; however, they are fast and convenient measures for comparing various environments and for making numerous initial predictions regarding speech intelligibility.

Table 11-17.
Speech intelligibility criteria for voice communication systems recommended by MIL-STD 1472F (1999).

Communication Requirement	Score		
	PB	MRT	AI
Exceptionally high intelligibility; separate syllables understood	90%	97%	0.7
Normal acceptable intelligibility; about 98% of sentences correctly heard; single digits understood	75%	91%	0.5
Minimally acceptable intelligibility; limited standardized phrases understood; about 90% sentences correctly heard (not acceptable for operational equipment)	43%	75%	0.3

Speech intelligibility index (SII)

Since speech intelligibility is a function of the SNR and acoustic characteristics of the environment, speech intelligibility in a given environment may be estimated on the basis of some physical data collected in this environment. Such estimations cannot replace completely perceptual tests described in the above section; however, they are much faster and cheaper to conduct, and they have some predictive value.

The standard objective measure of speech intelligibility used in the U.S. is the *speech intelligibility index* (SSI) described in the ANSI S3.5-1997 (R2007) standard. There are two specific speech intelligibility indexes recommended by and described in this standard. The first is a revised version of the AI, and the second is the *speech transmission index* (STI).

AI is a measure of speech intelligibility originally proposed by French and Steinberg (1947) and Beranek (1947) and further developed by Kryter (1962a, 1962b; 1965). The AI concept is based on the relationship between the “standard” speech spectrum and the spectrum of the background noise. The noise spectrum is measured in several frequency bands across the frequency range from 160 Hz to 6300 Hz, which was determined to be critical to the understanding of speech. The AI is calculated by combining the SNRs of all bands weighted by coefficients indicating each band’s relative contribution to speech intelligibility. The overall intelligibility then is expressed on a scale from 0 to 1. Several methods of dividing the speech spectrum into frequency bands are suggested in ANSI S3.5-1997 (R2007), including the twenty-band, one-third octave band, and an octave band method. Each method uses different weighting factors representing the corresponding band’s overall contribution to speech intelligibility. The AI gives an index value that represents the proportion of speech information that is above the noise – not the percentage of speech items that will be recognized. The ANSI S3.5-1997 (R2007) standard provides data relating AI values to speech intelligibility scores obtained for several common perceptual tests (e.g., nonsense syllables, rhyme tests, PB words, sentences, limited vocabularies and known sentences). These relationships are shown in Figure 11-27. The AI version described in the ANSI standard also can be applied to determine the effects of hearing loss on intelligibility. The calculation procedure treats the hearing threshold in the same way as ambient noise and intelligibility is calculated given as the percentage of speech information that is above the hearing threshold.

One of the drawbacks to AI is that it does not account for temporal distortion (e.g., echoes, reverberation, and automatic gain control), and non-linear distortion (e.g., system overload, quantization noise) affecting the speech. To account for these effects Steeneken (1992) and Steeneken and Houtgast (1980, 1999) developed the *speech transmission index* (STI) based on the concept of the modulation transfer function (MTF). The authors assumed that the intelligibility of a transmitted speech is related to the preservation of the original temporal pattern of speech and created a test signal that represents speech as a noise 100% modulated with several modulation frequencies. This modulated test signal is broadcast from a loudspeaker at the talker's location. A microphone is placed at the receiving end of the communication system to capture the broadcasted signal along with effects of reverberation and background noise present in the environment. The residual depth of modulation of the received signal is compared with that of the test signal in a number of frequency bands. Reductions in the modulation depth are associated with loss of intelligibility. These reductions constitute, in part, the *effective* SNR and are calculated in seven relevant frequency bands from 125 Hz to 8 kHz. These weighted values then are combined into a single index having a value between 0 and 1. As with the AI, intelligibility performance on a number of common perceptual measures is given for a number of STI values.

Both the AI and the STI have been implemented in psychoacoustic software programs and commercial room acoustics measurement devices that can be used to measure intelligibility in any operational environment. Although STI accounts fairly well for temporal and nonlinear effects, translation of index values to percent correct scores is only approximate, as in the AI case. Further, neither AI nor STI can account for the spatial arrangement of the sound sources in the operational environment, and they estimate speech intelligibility for the "worst-case scenario," when speech and noise are arriving from the same location in space.

Both AI and STI are based on measurements taken from a single microphone. Although there have been recent efforts to account for binaural effects (Wijngaarden and Drullman, 2008), to date no official binaural version of these tests exist. Many reports have shown an advantage of binaural listening for speech recognition. Two factors seem to contribute to this advantage. First, binaural listening allows the listener to utilize the *better ear advantage*, i.e., the listener can attend to the signal with his/her better ear or with the ear where the SNR is highest and ignore information in the less favorably positioned second ear (Brungart and Simpson, 2002; Culling, Hawley and Litovsky, 2004). Second, the listener can use spatial localization cues (described later in this chapter) to separate the speech information from the noise and can attend to the spatial location of the speech (Hawley, Litovsky and Culling, 2004; Kopčo and Shinn-Cunningham, 2008).

Despite the large number of measures of speech intelligibility, both perceptual and objective, none of these yet are truly able to measure the degree to which speech communication occurs. Beyond recognizing the phonemes and syllables that make up words and sentences, speech communication requires higher order comprehension of the thoughts and ideas that are transmitted along with them. Sometimes communication occurs that is not explicitly contained in the speech. Pragmatic information contained in our schematic knowledge about the world and the meaning of certain words in combination with certain patterns of events is not easily measured by either perceptual or objective speech measures. Nor can these measures fully predict which information will be attended to or processed by the listener or how this will change as a function of the contextual environment. Therefore, the speech tests and intelligibility indexes are best used for the comparison of elements in the channels affecting communication, i.e., they can give relative information (better/worse) but not absolute information about actual speech intelligibility.

Speech quality

The concepts of timbre and sound quality are used mainly in reference to music, sound effects, and virtual auditory environments. Assessment of speech is almost entirely based on speech intelligibility. Speech quality – in its sound quality meaning – is assessed much less frequently. It is also a confusing term because it has a second connotation that is similar to speech intelligibility.

Speech quality in its first, traditional meaning refers to the pleasantness of talker's voice and naturalness of speech flow. To stress this fact, such speech quality is sometimes referred to as vocal quality or voice quality, although such usage is not consistent. A limitation of speech quality defined as above is that it difficult to assess it for speech of poor intelligibility. Similarly, changes in speech quality can only be reliably assessed if they are not accompanied by large changes in speech intelligibility. If they are, speech quality changes are buried deep below speech intelligibility changes and are hard, if even possible, to be evaluated. In such cases the parallel judgments of speech intelligibility and speech quality frequently result in highly correlated scores, especially if the changes in speech intelligibility are fairly large (McBride, Letowski and Tran, 2008; Studebaker and Sherbecoe, 1988; Tran, Letowski, and McBride, 2008). However, when the compared samples of speech have the same, or very similar, and relatively high speech intelligibility, these samples may still greatly differ in speech quality and these differences may be reliably assessed by the listeners. In general, the higher speech quality, the greater satisfaction and listening comfort of the listener. The listener may prefer even a little less intelligible speech if the benefit in speech quality is large. For example, changes of low frequency limit of the channel transmitting speech signal have small if any effect on speech intelligibility but can greatly affect speech quality. Therefore, once the speech is sufficiently intelligible, it is important to assess and maximize its speech quality. If speech quality is relatively low, it may cause listener's annoyance and affect listening comfort and long-term performance of the listener.

The second meaning of speech quality is more technical and encompasses all aspects of transmitted speech. Its function is to represent overall audio quality of transmitted speech. The fact that scores for speech intelligibility and speech quality are highly correlated for imperfect speech led to the concept that speech quality really incorporates speech intelligibility and may be used as a sole criterion for assessment of speech transmission. Such connotation of speech quality is primarily used in telephony, speech synthesis, and digital communication. It encompasses natural and digital (lost packets, properties of speech codecs) causes of degraded speech intelligibility, presence of noise in the transmission channel, transmission reflections (echoes), and channel cross-talk. It is usually assessed by the listeners' ratings on a 5-step quality scale leading to a score called the *mean opinion score* (MOS). The standardized MOS scale used in evaluation of both audio and video transmission quality is shown in Table 11-18.

Table 11-18.
Mean opinion score (MOS) scale used in assessment of audio quality of transmitted speech (ITU-T, 1996).

MOS	Quality rating	Impairment rating
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Auditory Spatial Perception

Spatial perception is the awareness of environment, its boundaries, and internal elements. It allows an observer to determine directions, distances to and between, sizes, and shapes (spatial orientation) as well as realize utilitarian value or emotional impact of the whole environment or its specific part (space quality). If this awareness is focused on acoustic properties of the environment and their auditory image, such spatial perception is called *auditory spatial perception*. Auditory spatial orientation is mostly involved with directions, distances, and characteristic sound reflections that provide information about the size of the environment and materials of its boundaries. The utilitarian and emotional aspects of auditory spatial perception are reflected in the listener satisfaction with the perceived spaciousness of the environment, that is, in perceived spatial sound quality.

Auditory spatial orientation

Auditory spatial orientation is one of the critical abilities of living organisms. In the case of humans, the sense of balance located in the vestibular portion of the inner ear provides information about the position of the human body in reference to the force of gravity. Senses of vision and hearing, and to much lesser extent, the sense of smell, provide information regarding positions of other objects in space in relation to the position of the body. While vision is the primary human sense in providing information about the surrounding world that can be seen, hearing system is the main source of spatial orientation allowing humans to locate objects in space, even if they cannot be seen.

Anatomic structures and physiologic processes of the auditory system have been described in Chapter 8, *Basic Anatomy of the Hearing System,* and Chapter 9, *Auditory Function.* In general, human ability to perceive spatial sound and localize sound sources in space is based on the presence of two auditory sensors: the ears and the presence and elaborate shape of human pinnae.

Several acoustic cues are used by humans for auditory orientation in space. The importance of the specific cues depends on the type of surrounding environment and the specific characteristics of the sound sources present in this environment Thus, in order to understand the mechanics of the spatial auditory perception, it is necessary to outline the primary elements of the space leading to spatial orientation (Scharine and Letowski, 2005). These elements are:

- *Azimuth* – the angle at which the specific sound source is situated in the horizontal plane or the angular spread of the sound sources of interest in the horizontal plane (horizontal spread or panorama; see Figure 11-24),
- *Elevation* (zenith) – the angle at which the specific sound source is situated in the vertical plane or the angular spread of the sound sources of interest in the vertical plane (vertical spread),
- *Distance* – the separation of the listener from the specific sound source or the separation between two sound sources situated in the same direction (perspective or depth; see Figure 11-24), and
- *Volume* - the size and the shape of the acoustic environment in which the observer is situated.

Azimuth, elevation, and distance represent polar coordinates of any point of interest in a Cartesian space having its origin anchored at the listener’s location, and the volume is a global measure of the extent of the space that affects the listener. The set of polar coordinates is shown in Figure 11-28. The awareness of these four elements of space leads to auditory perception of surrounding space. This perception encompasses sensations (perceptions) of directions in horizontal and vertical plane, recognition of auditory distance and auditory depth, and sensation (perception) of ambience (perceived size of space) that together allow us to navigate through the space and feel its spaciousness (see Figure 11-24). They also allow us to distinguish between the specific locations of various sound sources and to describe their relative positions in space. Some of these abilities are the direct result of different auditory stimuli acting on each of the ears of the listener, whereas others result from single ear stimulation. In the latter case, the presence of two ears improves auditory performance, but the perceptual response is not the result of differential processing of two ears’ inputs by the auditory system.

Binaural hearing

The human ability to hear a sound with two ears is called *binaural hearing*. If the same sound is received by both ears such auditory stimulation is called the *diotic presentation* and has been described in Chapter 5, *Audio Helmet Mounted Displays*. Diotic presentation of a stimulus results in the lower binaural threshold of hearing and the higher binaural loudness than the respective monaural (single ear) responses. This process is called *binaural summation* or *binaural advantage* and has been discussed previously in this chapter. Note, however, that when a target sound is presented in noise, the same masked threshold in observed in both the monaural and binaural listening conditions assuming that both the target sound and the noise are the same in both ears (Moore, 1989).

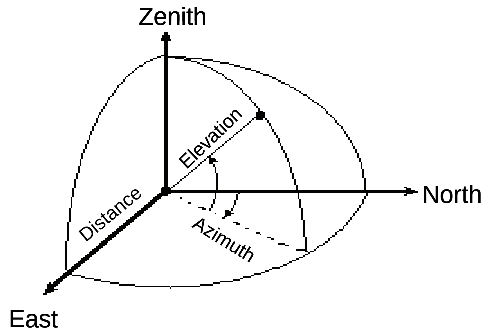


Figure 11-28. Azimuth, elevation, and distance in polar coordinates.

If the same sound is received by both ears but the ears differ in their properties, the binaural advantage is typically less. In addition, the ear disparity may lead to difficulties in pitch perception, called *binaural displacusis* (Van den Brink, 1974; Ward, 1963). The binaural diplacusis is the difference in pitch sensation in the left and right ear in response to the same pure tone stimulus. This difference leads to difficulty in pitch judgments of pure tone stimuli, but it is washed out and not perceived for complex stimuli.

If the sounds received by two ears are not the same in each ear, they can be treated by the auditory system as two independent sounds that contralaterally mask each other or as two slightly different representations of the same stimulus, resulting in a single fused auditory image appearing to exist inside or outside of the listener's head. The actual perceptual response depends on the character and extent of the differences between the sounds, i.e., the degree of correlation between the left and right ear stimuli. Note that the ears being some distance apart allows even the same original sound arriving at the left and right ear to differ to some degree in its spectral content and temporal envelope. Perception of such different but highly correlated stimuli is the basis for sound source localization in space. In contrast, when the left and right ear stimuli are not or poorly correlated with each other, such stimulation is called the *dichotic presentation* (described in Chapter 5, *Audio Helmet Mounted Displays*).

One of the most intriguing phenomena of binaural listening is the *binaural masking level difference* (binaural MLD or BMLD). The binaural MLD is the decrease in the masked threshold of hearing under some binaural listening conditions in comparison to the monaural listening condition. This phenomenon can be observed when a person listens binaurally to a target sound masked by a wideband noise but either target sound or noise differs in phase between the ears.

As mentioned earlier, the binaural masked threshold of hearing is the same as the monaural masked threshold of hearing if both the target sound and the masking noise are identical in both ears. However, when the phase of either the target sound or noise is reversed 180° in phase in one of the ears, the audibility of the target sounds markedly improves (Noffsinger, Martinez and Schaefer, 1985). Even more surprisingly, the monaural masked threshold of hearing improves when the same noise is added to the opposite ear. The improvement is in the order of 9 dB, which is larger than the approximate 6-dB improvement observed when the target sound, rather than the masking noise is added to the opposite ear (Moore, 1989). When both the masking noise and the target sound are added to the opposite ear, the masked threshold increases and becomes again the same as in the monaural listening condition.

The binaural MLD phenomenon was originally reported by Licklider (1948) for speech recognition in noise and by Hirsh (1948) for detection of pure tone signals in noise. Licklider (1948) reported that speech recognition through earphones in a noisy environment was greatly improved when the wires leading to one of the earphones were reversed. Hirsh (1948) and others (e.g., Durlach and Colburn, 1978; Egan, 1965; Roush and Tait, 1984; Schoeny and Carhart, 1971) reported thereafter that when continuous tone and noise are presented in phase in both ears (S_oN_o condition) or are reversed in phase in both ears ($S_\pi N_\pi$ condition), the masked detection threshold for the tone is the same as for the monaural condition. However, when either the tone or noise are reversed in phase, the $S_\pi N_o$ condition or S_oN_π condition, respectively, the detection threshold for the tone improves dramatically and the improvement is as large as 10 to 15 dB. The masking noise can be either a wideband noise or, in the case of a pure tone target sound, a narrowband noise centered on the frequency of the pure tone target sound. The improvement is the greatest for low frequency tones in 100 to 500 Hz range and decreases to 3 dB or less for stimulus frequencies above about 1500 Hz. If the phase shifts are smaller than 180°, a similar, although smaller, binaural MLD effect has been observed. In general, the larger the phase shift the larger the size of the binaural MLD effect. The exact physiologic mechanism of the binaural MLD is still unknown although some MLD results can be explained by the equalization-cancellation (EC) mechanism proposed by Durlach (1963). According to Colburn (1977), the binaural MLD effect also can be accounted for by the response patterns at the outputs of the coincidence detectors in the medial superior olivary (MSO) nucleus (e.g., Colburn, 1977).

Masking by noise aside, if the same sound is received by both ears, the sound source is perceived as located in the median plane of the listener. If the sounds differ in their time of arrival and/or intensity, the sound source is perceived as being located at a certain azimuth angle to the left or to the right of the median plane but not at the median plane. This effect is called *lateralization*, interpreted as "to the left" or "to the right" from the median plane but does not necessarily imply any specific location.

Note that in the case of binaural reception of auditory stimuli a sound source can be perceived as located outside the head (e.g., in natural environments or during loudspeaker-based sound reproduction) or inside the head (e.g., earphone-based sound reproduction). In the former case, the sound source location can be identified relatively precisely in both horizontal and vertical plane. When the sound source is located inside the head, it can only be crudely located on a shallow imaginary arc connecting left and right ear, and the perceived deviation of the auditory image location from the median plane can only be judged as partial of full (toward one of the ears) lateralization. Thus, spatial phenomena inside-the-head is referred to as lateralization while the term localization is reserved for spatial phenomena outside of the head.

Lateralization and binaural cues

Spatial perception of sound is based on two main sets of cues: binaural cues and monaural cues. Binaural cues result from the differences in stimuli received by two ears of the listener and are basic cues facilitating sound lateralization. The binaural cues were described first in 1907 by Lord Rayleigh as foundation of what is often called the Lord Rayleigh's *duplex theory*. There are two binaural cues that facilitate sound localization in the horizontal plane: (a) *interaural level difference* (ILD), also referred to as *interaural intensity difference* (IID), and (b) *interaural time differences* (ITD) or *interaural phase difference* (IPD). Both cues are shown in Figure 11-29.

The IID refers to the difference in the intensity of sound arriving at the two ears caused by the baffling effect of the head. The head is casting an "acoustic shadow" on the ear farther away from the sound source, decreasing the intensity of sound entering that ear. The effectiveness of the baffling effect of the head depends on the relative size of the head (d) and the sound wavelength (λ). The larger the difference between d and λ ($d >> \lambda$), the stronger the baffling effect. Thus, at low frequencies, where the dimensions of the human head are small in comparison to the wavelengths of the sound waves, sound waves diffract around the head, and the difference in sound intensity at the left and right ear is small if any. However, at high frequencies, the intensity differences caused by the acoustic shadow of the head are large and provide effective localization cues. For example, when the sound source is situated in front of one ear of the listener, the IID (or ILD) can be as large as 8 dB at 1 kHz and 30 dB at 10 kHz (Steinberg and Snow, 1934). The effect of sound frequency on the size of the ILD cue is shown in Figure 11-30.

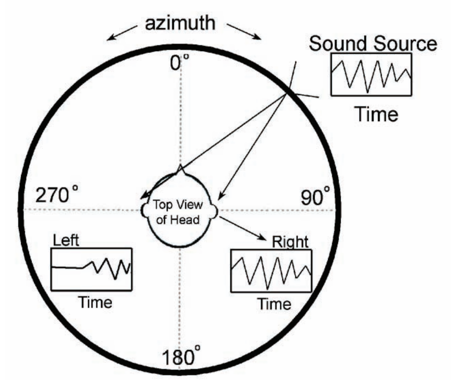


Figure 11-29. The interaural time difference (ITD) and interaural level differences (ILD) created by a sound arriving from 45° azimuth angle. Note that the sound arrives earlier and has more energy at the right ear.

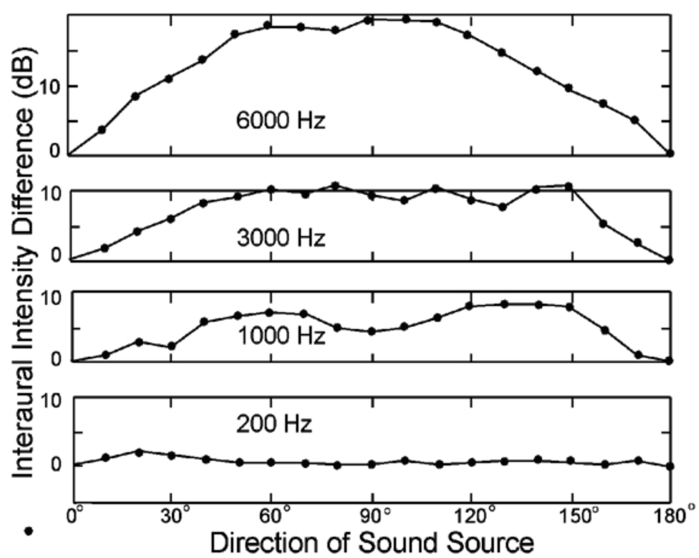


Figure 11-30. Interaural level differences (IID/ILD) for four pure tone signals; 220 Hz, 1000 Hz, 3000 Hz, and 6000 Hz. At 200 Hz, there is no shadowing effect due the sound diffraction around the head (adapted from Feddersen et al., 1957).

The ITD refers to the difference in the time of arrival of the sound wave at the two ears. If a sound source is located in the median plane of the head, there is no difference in the time of sound arrival to the left and right ear. However, if a sound is presented from the side of the head or any other angle off the median plane, the sound reaching the further away ear arrives with a certain time delay. Assuming that the human head can be approximated by a sphere, the resulting time difference can be calculated from the equation:

$$\Delta t = \frac{r}{c}(\alpha + \sin\alpha), \qquad \text{Equation 11-21}$$

where Δt is the ITD in seconds, r is the radius of the sphere (human head) in meters, c is the speed of sound in m/s, and α is the angle (azimuth) of incoming sound in radians (Scharine and Letowski, 2005). The dependence of the ITD on the angular position of the sound source is shown in Figure 11-31. The maximum possible ITD occurs when the sound source is located on the imaginary line connecting both ears of the listener and is dependent on the size of the listener's head, the speed of sound, and to some extent on the distance of the sound source from the listener's head. For example, for a head with the diameter d = 20 cm and a sound wave velocity c = 340 m/s, the maximum achievable ITD is about 0.8 ms. For a given head size, larger ITDs indicate more lateral and smaller ITDs less lateral sound source locations. The smallest perceived ITD is to order of 0.02 to 0.03 ms and is being detected when the sound is arriving from a 0° angle, i.e., from a sound source directly in front of the listener. This difference corresponds to the shift in the horizontal position of the sound source by about a 2° to 3° angle.

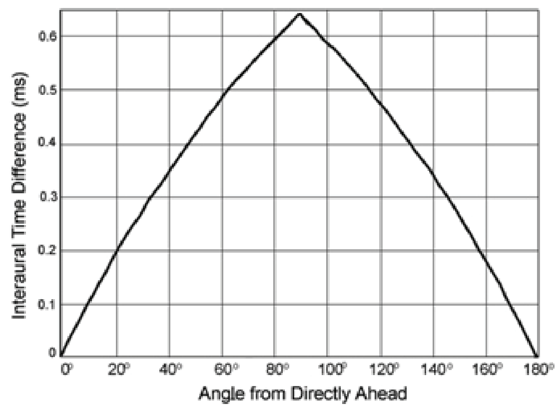


Figure 11-31. Interaural time differences plotted as a function of azimuth (adapted from Feddersen et al., 1957).

The ITD cue works well at lower frequencies but it fails at high frequencies. First, the phase information becomes ambiguous above approximately 1200 to 1500 Hz depending on the size of the head. At this frequency the length of one period of the sine wave corresponds to the maximum time delay of sound traveling around the head. This means that at this and higher frequencies ITD may be larger than duration of a single period of the waveform making time delays ambiguous. This ambiguity is shown in Figure 11-32. Note that the second waveform in the last pair of waveforms is delayed by the whole period regarding the previous waveform, but both of them arrive at the same phase.

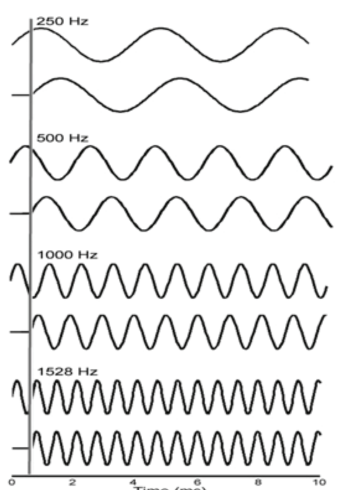


Figure 11-32. Comparison of the interaural phase relationship of various sinusoids.

Second, each auditory neuron fires in synchrony with a particular phase in the auditory waveform. This effect is called *phase locking.* The frequency of neuron firing is limited to about 4 to 5 kHz (Rose et al., 1968; Palmer and Russell, 1986), and this limit is because phase timing variability becomes large with respect to the length of the frequency cycle. This means that any phase synchrony in neuron firing is lost for frequencies above 4 to 5 kHz.

Note that the high frequency limitations discussed above are actually the interaural phase difference (IPD) limitations and apply only to the continuous stimuli. However, in the case of clicks, onset transients, and similar non-periodic sounds, the time difference shorter than 0.6 to 0.8 ms can be used to guide sound localization since this is the difference between two single temporal events that are not repeated periodically (Leakey, Sayers and Cherry, 1958; Henning, 1974).

Both of the above mechanisms limit the use of ITD to localization of only low and mid frequency sounds. However, high frequency sound can be localized using ILDs. The complimentary character of the ITD and ILD cues became the foundation of the Lord Rayleigh's (1907) duplex theory, which states that the lateral position of a sound source in the space is determined by the combination of both cues, ILDs at high frequencies and ITDs at low frequencies. A consequence of the duplex theory is that sounds containing frequencies between 1 to 4 kHz should be difficult to lateralize accurately because neither the ITD nor ILD cue is strong enough in this frequency region. Later studies have largely confirmed this theoretical assumption (Stevens and Newman, 1936; Wightman and Kistler, 1992). However, it should be cautioned that these facts only hold true for pure tones. Most sounds are composed of multiple frequencies and can be lateralized using a combination of both cues for their lower and higher components. Thus, pulses of wideband noise containing both the low-and high-frequency energy are the easiest stimuli to localize (Hartmann and Rakerd, 1989) and are the preferred stimuli for directional beacons (Tran, Letowski and Abouchacra, 2000).

Localization and monaural cues

Binaural cues allow effective left-right lateralization, but they have two major limitations. First, binaural cues do not differentiate between sound arriving from the front or the rear of the head. Relative symmetry between front and back of the head results in confusion as to whether the sound is arriving for example, from the 10° or 170° direction. Some binaural differentiation is possible because the head is not cylindrical and the ears locations are not exactly symmetrical, but the front-back localization is not improved much by binaural cues.

Second, binaural cues do not provide any information about sound source elevation. In fact, if one assumes a spherical head, there is a conical region, called a *cone of confusion*, for which a given set of binaural cues is the same. This means that all sound sources located on the surface of a given cone of confusion generate identical binaural cues. As a result, the relative locations of these sound sources cannot be differentiated by the binaural cues alone (Oldfield and Parker, 1986). The concept of a cone of confusion is shown in Figure 11-33. Numerous studies have demonstrated that the cone of confusion is the source of localization errors in both the vertical and the front-back directions (e.g., Oldfield and Parker, 1984; Makous and Middlebrooks, 1990).

The differences between various sound-source locations within a cone of confusion are resolved by the presence of the spectral localization cues, called also the *monaural cues* since they do not require two ears to operate. Monaural cues are the primary cues allowing sound source localization in the vertical place and along the front-back axis.

Monaural cues are directionally dependent spectral changes that occur when sound is reflected from the folds of the pinnae and the shoulders of the listener. These reflections create peaks and notches in the spectrum of the arriving auditory stimulus, changing the spectral content of the waveform arriving at the tympanic membrane. This effect is described in Chapter 9, *Auditory Function,* and shown in two different ways in Figures 9-2 and 11-34. The locations of peaks and notches in the sound spectrum of the arriving auditory stimulus change as a function of the angle of incidence, thus providing information that can be used to distinguish the front from the rear hemisphere (Musicant and Butler, 1985) and between various elevations (Batteau, 1967; Hebrank and Wright, 1974).

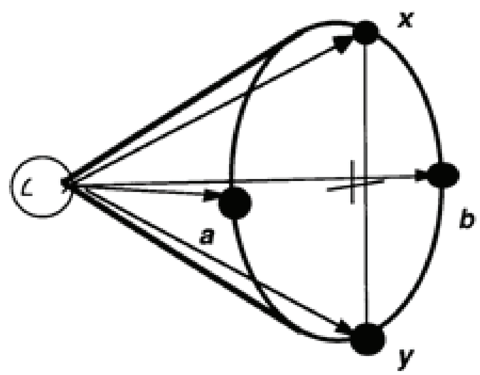


Figure 11-33. The concept of the *Cone of Confusion*. The cone represents a region for which interaural level and phase cues would be the same if a spherical head is assumed (after Mills, 1972).

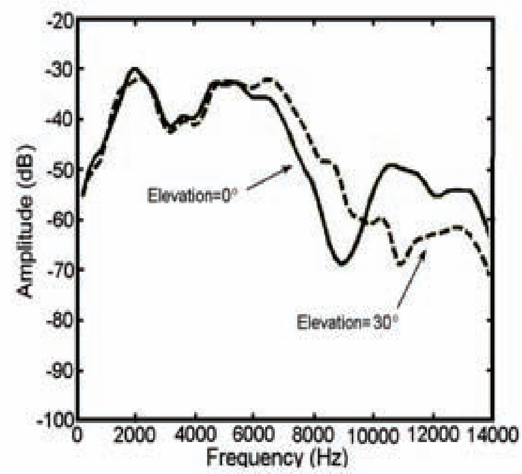


Figure 11-34. Two head-related transfer functions measured for 0° and 30°. This figure illustrates how the frequency notch changes as a function of angular position.

Passive filtering of sound by the concave surfaces and ridges of the pinna is the dominant monaural cue used in sound localization. Gardner and Gardner (1973) observed that localization performance for sound stimuli located on the medial sagittal plane got progressively worse as the pinnae cavities were filled in by using silicon fillers custom-made for each listener. They also observed that the precision of sound source localization in the vertical plane was the best for wideband noises and for narrowband noises with center frequencies in 8 to 10 kHz region. The filtering effect of the shoulders is weaker than that of the concha and pinna ridges, but it is also important for sound localization since it operates in slightly lower frequency range than the others.

Despite their name – monaural cues – these cues are duplicated by the simultaneous monitoring of the sound source location by both ears of the listener. Any asymmetry in the vertical or front-back location of the ears on the surface of the head provides important enhancement of the listener's ability to localize sounds along the respective directions. For some species, like owls, ear asymmetry is the main localization cue in the vertical direction. It also has been mentioned that monaural cues do not only operate in the vertical plane and along front-back axis, but they also operate together with the binaural cues along the left-right axis and enhance human localization precision in the horizontal plane.

The effects of both binaural and monaural cues can be captured by the placement of very small probe microphones in the ear canal of the listener. A sound is presented from various angular locations at some distance from the human head, and the directional effects of the human body and head are captured by the microphone recordings. The difference between the spectrum of the original auditory stimulus and the spectrum of the auditory stimulus recorded in the ear canal is called the *head-related transfer function* (HRTF). The HRTF varies as a function of the angle of incidence of the arriving auditory stimulus and, for small distances between the head and the sound source, also as a function of the distance (Brungart, 1999).

The HRTFs can be recorded for a selection of azimuths and elevations relative to the orientation of the listener's head and in the form of impulse responses convolved with any arbitrary sound to provide arbitrary spatial information about the sound source location. This technique is used to create externalized spatial locations of the sound sources when the auditory stimuli are presented through the earphones. Such spatial reproduction of sound through earphones is often referred to as *3-D audio* when referring to auditory display systems. Additional information about practical applications of the HRTFs may be found in Chapter 5, *Audio Helmet-Mounted Displays.*

It needs to be stressed that the monaural cues are relative cues. Unless a listener is familiar with the original signal and surrounding space, there is no invariant reference to be used to determine what notches and peaks related to sound source location are present in the arriving auditory stimulus. Therefore, sound localization ability, especially in the vertical plane and along the front-back axis, improves with experience and familiarization with both the stimuli and environment (Plenge, 1971). This is also the reason that some authors consider auditory memory as another auditory directional cue (Plenge and Brunschen, 1971). For example, if a listener is familiar with somebody's voice, this familiarity may help the listener to differentiate whether the talker is located in front or behind the listener. The lack of familiarity with specific auditory stimuli is reported frequently in the literature as the secondary reason for front-back confusions and poor localization in vertical plane.

There is one more potential reason for the poor front-back and vertical discrimination of sound source locations. Blauert (2001) observed that narrowband stimuli presented within the medial sagittal plane have the tendency to be associated with the front, rear, or overhead, regardless of the actual position of the sound source. Since this tendency is the same for sounds located in the specific frequency bands, Blauert called these bands the *directional bands*. The concept of the bands is shown in Figure 11-35. In general, the listeners have the tendency to localize stimuli in 125 to 500 Hz and 2 to 6 kHz bands as coming from the front, stimuli in 500 to 2000 Hz and 10 to14 kHz bands as coming from the back, and stimuli in 6 to 10 kHz band as coming from the top if they do not have any other environmental cues. The bands apply to tones and narrowband stimuli but under some condition they may also affect localization of more complex stimuli.

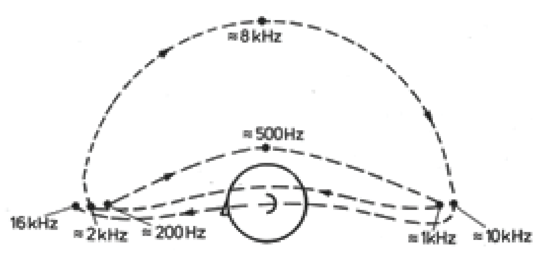


Figure 11-35. The listener's tendency to localize narrowband noises as coming from the front, top, or back if the sound is presented the same number of times from each of these direction (after Blauert, 2001 [Figure 2.6]).

The last but very effective cue that is extremely important in real world environments is that provided by head movement. Even if a sound is unfamiliar, the auditory system can gain disambiguating information if the listener moves his or her head while the sound is present (Perrott and Musicant, 1981; Thurlow, Mangels and Runge, 1967). Small movements of the head from the left to the right will quickly clarify whether the sound is in the front or rear hemisphere. Vertical movements give salient elevation information (Wallach, 1940). Assuming that the sound is long enough to allow for movement, most of the shortcomings of binaural and monaural cues can be overcome (Hirsh, 1971).

All previous discussion has centered on sounds emitted from the stationary sound sources. There are two general measures of sound localization ability that apply to stationary sound sources: *localization acuity* and *localization accuracy*. Localization acuity is a person's ability to discriminate whether the sound source changed its position or not. It is usually described as the *minimum audible angle* (MAA), which is the DL of directional perception. Localization accuracy is a person's ability to localize the sound source in space. It is usually characterized by the standard error (or other measure of dispersion) in the direction recognition task. However, in the real world environments a large proportion of sound sources is not stationary but is moving at various directions and various speeds. Human localization precision of such sound sources is usually measured as the *minimum audible movement angle* (MAMA) for specific direction and speed of the moving sound source (Perrott and Musicant, 1977; 1981). The MAMA is usually larger than the MAA, but they characterize different auditory abilities of the listener. In terms of absolute sound localization, a fast sound source movement makes it more difficult to identify the momentary position of sound source.

The polar characteristic representing a listener's ability to localize sounds in the horizontal plane is frequently referred to as the directional characteristic of the human head. Such a characteristic usually is measured with a narrowband signal for a selection of azimuth angles or by using a turntable and an automatic Bekesy tracing technique (Zera, Boehm and Letowski, 1982). The data usually are displayed as a family of polar patterns obtained for signals of various frequencies in a manner similar to polar patterns of microphones or loudspeakers. Another method of displaying directional characteristics of the listener's head is shown in Figure 11-36 where localization precision for a selected stimulus is shown in numeric form on the imaginary circle around the listener's head.

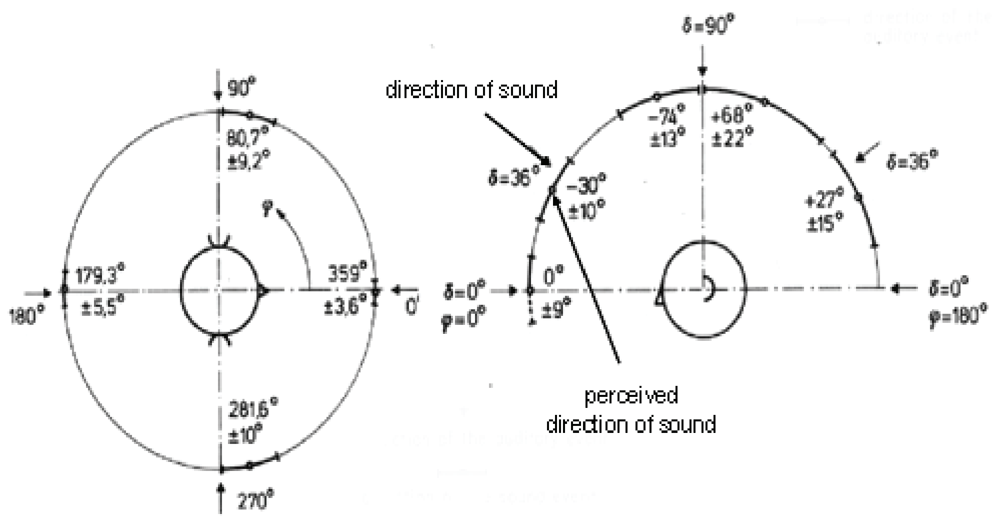


Figure 11-36. Localization uncertainty in the horizontal plane (left panel) and vertical plane (right panel) of white noise pulses of 100 ms duration presented at 70 dB phon level (adapted from Blauert, 2001 [Figures 2.2 and 2.5]).

The precedence effect

Much of the research reported here was conducted in free field environments, e.g., in open outdoor spaces or in anechoic chambers. However, such environments are rare, and it is more common to hear sounds in rooms, near buildings, and in other locations where there are reflective surfaces. The effect of such environments is that one or more reflected sounds arrive at the listener's ears shortly after the original sound. Most of the time, the listener is unaware of this reflected sound, and to some extent continues to be able to localize the sound accurately in spite of the presence of reflected sounds. The mechanism by which this occurs is called the precedence effect.

The *precedence effect* is the phenomenon that the perception of the second of two successively received similar sounds is suppressed if the second sound is delayed up to about 30 to 40 ms, and its intensity does not exceed the intensity of the first sound by more than 10 dB. The precedence effect was discovered originally by Wallach, Newman and Rosenzweig (1949), who conducted a series of experiments testing the effect of a delayed copy of the sound on localization by presenting pairs of clicks over headphones. They observed that if the clicks were less than 5 ms apart, they are fused and are perceived as a single sound image located in the center of the head. However, if the clicks were 5 to 30 ms apart only the first of the two clicks is heard.

The existence of the precedence effect explains the ability of the auditory system to determine the actual position of the sound source without being confused by early sound reflections. If the locations of the two arriving sounds differ, the perceived location of the fused click image is largely determined by the location of the first click. The suppression of the location information carried by the second sounds is known also as he *Haas effect* (1951), after Haas who rediscovered this effect in 1951, and as the *law of the first wavefront*.

In a typical precedence effect demonstration, the same sound is emitted by two loudspeakers separated in space. If the sound coming from one of the loudspeakers is delayed by less than about 1 ms, the fused image is localized somewhere between the locations of both loudspeakers in agreement with the ITD mechanism described previously. Such phantom sound source location resulting from perception of two separate sounds is called *summing location* and the process is called *summing localization* (Blauert, 1999). The range of time in which summing localization operates is approximately equivalent to maximum ITD for a given listener.

When the time delay of the lagging sound is between 1 and 5 ms, the sound appears to be coming from only the lead loudspeaker, but its timbre and depth perception change. If the time delay of the second sound is more than 5 ms but less than 30 ms, depending on the specific environment, only the first sound is heard, and the second sound has no effect. Obviously, if the lagging sound is more than 10 to 15 dB more intense than the leading sound, only the second sound and its direction are heard (Moore, (1997). If the time delay is longer than about 30 ms and the sounds are very similar, the second sound is heard as an echo of the first sound. If the sounds are very different, two separate sounds are heard.

To some degree, the two sounds can be different and the precedence effect may still occur (Divenyi, 1992), but similarity increases the effect. It also has been shown that the precedence effect can take some time to build up (Freyman, Clifton and Litovsky, 1991). The authors described an experiment in which a train of leading clicks with simulated echo clicks delayed by 8 ms was presented. At first, two clicks were clearly heard, but after a few repetitions, the two clicks fused. This fusion is disrupted if the acoustical conditions are changed (Clifton, 1987) For example, Clifton showed that if a train of lead-lag click pairs is presented and then the locations of the leading and lagging click are swapped, fusion is disrupted temporarily and the clicks are once again heard separately. After a few more presentations, they fuse again (see Chapter 13, *Auditory Conflicts and Illusions*, for a more complete description of the Clifton effect and a related effect, the Franssen effect). This can be compared with becoming adapted to a particular acoustic environment. After a few seconds, one begins to ignore the acoustic effects of the room. If the room were suddenly to become drastically altered (an improbable event), the echoes suddenly would become more apparent, only to fade away afterwards.

Auditory distance perception

Auditory distance is the distance between the listener and a sound source, determined on the basis of available auditory cues. If the perception involves an estimation of the distance between two sound sources located along the same imaginary line passing through the head of the listener, such distance is called *auditory depth.* In both cases there are no absolute cues for distance perception; however, there are several relative ones, which combined with non-auditory information, allow individuals to determine the distance from the sound source to the listener. These cues, called distance cues or range cues, depend on the specific environment but in general include: (a) sound loudness, (b) spectral changes, (c) space reverberance (liveness), and (c) motion parallax.

The primary auditory distance cue is *sound loudness*. For familiar sounds, one can compare the loudness of the perceived sound with the knowledge about the natural loudness and intensity of its source (Mershon and King, 1975). In the case of the prerecorded sounds, a critical requirement is that the prerecorded and the reproduced sounds have the same loudness (Brungart and Scott, 2001). Still, the distance to an unfamiliar sound source (or a sound source that may have various sound loudnesses at the source) is difficult to estimate using the loudness cue.

The loudness cue is the most obvious distance estimation cue in the outdoor environments. According to the inverse square law of sound propagation in open space, sound intensity decreases by 6 dB per doubling of the distance from the sound source. However, this rule only holds true for free-field environments. In enclosed spaces, wall reflections reduce this intensity drop associated with the distance from the sound source and at some *critical distance*, which is a function of the sound source distance from the listener and the reflecting walls, obviate the cue altogether.

The second cue is *sound timbre*. Low frequency components are less likely to be obstructed by objects and meteorological conditions than high frequency components of a sound. High frequency components are attenuated by humidity and transmitting matter and absorbed by nearby surfaces. Consequently, distant sounds will have relatively more low frequency energy than the same sounds radiated from proximal (nearby) sound sources and result in different sound timbre. Unfortunately, this cue also requires knowledge of the original sound source in order to be used effectively utilized (Little, Mershon and Cox, 1992; McGregor, Horn, and Todd, 1985).

The third cue, *reverberance* or *liveness*, is a major cue for distance perception in closed spaces or in situations that produce an echo. If the sound source is located close to the listener, the direct-to-reverberant sound ratio is high and sound clarity (e.g., speech intelligibility) is high as well. If the distance is large, the sound becomes less clear and its sound source location less certain (Mershon et al., 1989; Nielsen, 1993). For a given listening space, there may be multiple sound sources each with their own direct-to-reverberant sound intensity ratio. A listener familiar with that space can use this as a source of distance information. However, the specific ratio of these two energies depends on the directivity of the sound source and the location of the listener relative to the sound source and reflective surfaces of the space (Mershon and King. 1975).

Finally, when the listener moves (translates) the head, the change in the azimuth toward the sound source is distance dependent. This cue is called the *motion parallax cue*. If the sound source is located far away from the listener even a relatively large displacement of the head position results only in very small change in the azimuth and relatively small and slow imaginary movement of the sound source. Conversely, if the sound source is located nearby, even a small movement of the head causes a large change in the azimuth that results in a larger imaginary movement of the sound source and in a potential change in the loudness of the sound.

Despite the four auditory cues listed above, the auditory distance estimation is relatively poor, especially if the estimate is expressed in absolute numbers as opposed to a comparative judgment of two distances in space. In general, people underestimate the distance to the sound source in an exponential manner -- the larger the distance to the sound source the larger relative error (Fluitt, Letowski and Mermagen, 2003).

Space perception (Spaciousness)

Spaciousness is a catch-all term describing the human impression made by a surrounding acoustic space. Spaciousness embraces such sensations as ambience (impression that sound is coming from many directions), reverberance or liveness (impression of how much sound is reflected from the walls), warmth (impression of the spectral balance/imbalance in the reflected sound energy), intimacy, panorama, perspective, and many others. Similarly to the timbre domain, some of spaciousness-related terms are highly correlated and there are many that do not have an established meaning.

There are different types of acoustic spaces and therefore different forms of spaciousness. They include such spatial concepts as battlefield, serenity, and soundscape. Each of these terms brings with it a connotation of a specific sonic environment, and often related to it an emotional underpinning reflected in perceived spatial sound quality. For example, Krause (2008) divides all soundscapes into biophony (spaces dominated by sounds produced by non-human living organisms), geophony (spaces dominated by non-biological sounds of nature), and anthrophony (spaces dominated my man-made sounds). Every sound within a particular space brings with it information about the space as well as certain expectations regarding its natural fit within this environment.

In its general meaning, spaciousness is a sensation parallel to the sensation of timbre described previously. It is a perceptual reflection of the size and the character of the area over which a particular sound can be heard and perceived. Therefore, it is a perceptual characteristic of a soundstage, that is, a sound source operating within a particular environment. It needs both the sound and the space to exist. One important element of spaciousness is the size of a personal space in voice communication. Personal space can be generally defined as the area surrounding an individual that the individual considers as personal territory in any human-to-human interaction (Hall, 1966). A personal space is usually highly variable and depends on the personal traits of the individual and the social and cultural upbringing. For example, in the Nordic cultures the radius of personal space is generally larger than in the Southern cultures.

The same general comment about the variability of a personal space applies to auditory personal space defining the minimum acceptable distance between two unrelated people who communicate by voice. However, the radius of the auditory personal space seems to vary between individuals and is less than the radii of social space, aggression space, or work space. It generally is assumed that the distance of one meter (3.28 feet) defines a typical conversational situation and serves as a good estimate of the radius of the auditory personal space.

The concept of the auditory personal space is important for audio communication and this space needs to be preserved in creating phantom sound sources representing real people communicating through audio channels with real or virtual people. If the perceived auditory distance to another talker in an audio channel is quite different from 1 to 1.5 m, such voice communication will distract the operator, increase the workload, and increase the overall level of anxiety. Obviously, the above recommendation has only a general characteristic, and there are specific situations that the communication distance has to be different.

Hearing Loss

Hearing loss is a decreased ability to perceive sound due to an abnormality in the hearing mechanism. According to the American Speech-Language-Hearing Association (ASHA), 28.6 million people in the United States are living with hearing loss (ASHA, 2006). Hearing loss can affect not only the individual’s ability to hear the surrounding environment but also the clarity of speech perception. The functional effects of hearing loss can vary greatly according to the age of onset, period of onset, degree, configuration, etiology, and the individual’s communication environment and needs.

Three main types of hearing loss are labeled as: conductive, sensorineural, and mixed hearing loss. With conductive hearing losses, either the outer ear, middle ear or both are affected. Sensorineural hearing refers to loss that originates in the cochlea, the auditory nerve, or in any portion of the central auditory nervous system

(CANS). A mixed hearing loss is a combination of a sensorineural and a conductive hearing loss, and therefore, can involve many combinations and/or portions of the ear.

The disorders of the outer ear that can lead to a conductive hearing loss can be caused by congenital anomalies or acquired ones. Examples of congenital anomalies include narrowing of the ear canal known as stenosis, an absence of the ear canal known as atresia, a partially formed pinna known as microtia, or an absence of the pinna known as anotia. Examples of acquired anomalies include impacted ear wax or a foreign body in the ear canal. Otitis externa, also known as “swimmer’s ear”, can cause stenosis of the canal, and therefore, lead to a conductive hearing loss. Most conductive hearing losses within the outer ear partially or completely block the transmission of acoustic energy into the middle ear and they are treatable. The resonance of the pinna and the outer ear canal is between 2 to 7 kHz, which enhances the ability of the listener to localize and perceive their acoustic space (Rappaport and Provencal, 2002). Recalling that higher frequency sounds have shorter wavelengths, and therefore, can more easily be deflected, anomalies in the outer can impede the listener’s ability to localize.

Abnormalities in the middle ear causing conductive hearing losses can also be congenital or acquired. Serous otitis media, or inflammation of the middle ear fluid, can cause conductive hearing losses and are temporary in 90% of the cases and usually resolve without treatment within 3 months (Rappaport and Provencal, 2002). Perforations on the tympanic membrane can lead to minimal or maximal hearing loss depending on where the perforation occurs on the membrane. An abnormality in the ossicles caused by chronic ear infections can lead to ossicular erosion, creating a maximum conductive hearing loss of about 60 dB. Conductive hearing losses only affect up to about 60 dB since bone conduction hearing takes place beyond that level. Bone conduction hearing occurs when the sound intensity is strong enough to bypass the middle ear and move the bones in the skull, which moves the cochlear fluids in the inner ear. This stimulates hair cells, which leads to the perception of an auditory signal. Like outer ear conductive hearing losses, losses caused by middle-ear abnormalities are usually treatable.

Sensorineural hearing losses can be caused by one of hundreds of syndromes, a single genetic anomaly, perinatal infection, drugs that are toxic to the auditory system, tumors, idiopathic disease process, aging, or from noise exposure. ASHA reports that 10 million of the 28 million with hearing loss are due to noise exposure (ASHA, 2006). Sensorineural hearing loss is characterized by irreversible damage that distorts auditory perception. Generally, “sensory” hearing loss refers to an abnormality in the cochlea and “neural” refers to an abnormality beyond the cochlea, meaning in the VIIIth nerve or beyond.

Sensorineural loss can be either permanent or temporary. Temporary hearing loss, called also *temporary threshold shift* (TTS), is usually noise-induced and may last from under an hour to several days, and the degree and duration of loss depends upon the duration, intensity, and frequency of the noise exposure (Feuerstein, 2002). Excessive exposure to sounds energy in the frequency range from 2000 to 6000 Hz is most likely to cause permanent changes in hearing. Permanent hearing loss (permanent threshold shift [PTS]), is the residual loss from a noise-induced temporary hearing loss or aging process and results from damaged hair cells. Usually it is a slow process and the individual may not perceive any change in hearing for long time. If the hearing loss is due to acoustic trauma (sudden exposure to very intense noise), the PTS typically will plateau by within the first eight hours. With conductive hearing loss, once sound level is increased to compensate for the PTS, the signal is audible and clear. With sensorineural hearing loss, at frequencies where the PTS occurs, even once audibility is restored, some level of sound distortion usually persists.

The mechanism for cochlear damage may come from a variety of sources including: interruption of blood flow due to ischemic damage, mechanical injury to the hair cells due to the shearing force of the traveling wave, hair cell toxicity caused by a disruption of ionic balances in the cochlear fluids, or hair cell toxicity from an over-active metabolic process caused by an immune response (Lonsbury-Martin and Martin 1993). Age-related hearing loss is most commonly seen in the 7th decade of life, with the basal region of the cochlea being most affected (Weinstein, 2000; Willott, 1991). Both noise-induced and age-related hearing loss (presbycusis) are characterized by high-frequency loss of varying degrees. Therefore, presbycusis, or age related hearing loss, is often difficult to tease out from noise-induced hearing loss. However, age has been positively correlated with a worsening in pure tone thresholds. As women age, on average their thresholds from 3000 Hz and above worsen to a mild hearing loss, whereas men's threshold from 2000 Hz and above worsen to a mild to moderate hearing loss (Weinstein, 2000). Given that background noise is generally low-frequency, this type of noise can exacerbate a high-frequency hearing loss since it can mask the portion of the signal that the listener is able to hear in quiet.

As previously stated, mixed hearing loss is a combination of both sensorineural and conductive hearing loss. Otosclerosis is an example of mixed hearing loss that is caused by a disease process on the ossicular chain, most commonly affecting the footplate of the stapes. Although the cause is unknown, the disease process usually softens the bone, and the bone then hardens and can become fixed to the oval window. Since the resonant frequency of the ossicular chain is around 2 kHz, the hearing loss usually presents as a conductive hearing loss with a sensorineural component at 2 kHz. Another example of a mixed hearing loss could be due to severe acoustic trauma. Trauma could cause a tympanic membrane perforation as well as a noise-induced hearing loss. Mixed losses can be caused by a variety of pathologies or combination of pathologies that can vary greatly in severity.

Hearing loss is described by the degree and type, and configuration. The basic metric used to assess degree of hearing loss is the *pure tone average* (PTA) calculated as an average hearing threshold across a specific range of frequencies. The frequencies considered most important to speech perception are 500, 1000, and 2000 Hz. A loss at these frequencies can more adversely affect speech perception than one that occurs at 3000 Hz and above. Therefore, PTA is most commonly calculated as the average value of the threshold of hearing at 500, 1000, and 2000 Hz expressed in dB. Sometimes the average includes different combination of frequencies, which in such cases should be clearly stated. If they are not, the 500, 1000, and 2000 Hz average needs to be assumed.

The degree of hearing loss based on standard PTA calculation is separated into seven categories listed in Table 11-19. Since this range of frequencies is the most important for speech recognition such defined PTA should be numerically close to the speech reception threshold, which is usually within 5 dB of each other.

Table 11-19.
Classification of hearing loss (Harrell, 2002).

Extent of Hearing Loss (dB HL)	Degree of Hearing Loss
-10 to 15	Normal
16 to 25	Slight
26 to 40	Mild
41 to 55	Moderate
56 to 70	Moderately-severe
71 to 90	Severe
>90	Profound

The PTA metric is generally a monaural metric and defines hearing loss for each ear separately. A symmetric hearing loss is assumed when the PTAs calculated across 500, 1000, and 2000 Hz, and frequently also 3000 Hz or 4000 Hz frequencies for the left and right ear are within 10 dB of each other. Binaural hearing loss may be assessed by the PTA by calculating an arithmetic average of the PTAs obtained separately for left and right ears or by using a better threshold level in either ear at 500, 1000, and 2000 Hz. Each approach leads to a slightly different result but there is yet no strict standard accepted how to calculate the bilateral PTA.

Configuration refers to the hearing thresholds relative to one another. For example, a flat hearing loss means that less than a 5 dB average change exists among octaves. Other categories include gradually sloping, sharply sloping, precipitously sloping, rising, and notch, but their descriptions are beyond the scope of this book.

The U.S. Army classifies hearing loss into four fitness-for-duty categories, H1-H4 (Army Regulation 40-501 [Department of the Army, 2008]). An H-1 designation means that no limitations exist based on the Warfighter's hearing. The determination of fitness for duty is based on threshold levels, measured in dB HL at 500, 1000, 2000, and 4000 Hz. For an H-1 designation, neither ear can have an average threshold at 500, 1000, and 2000 Hz greater than 25 dB HL, with no individual level greater than 30dB. Thresholds at 4000 Hz cannot exceed 45 dB HL. For a list a specific requirements see AR 40-501. An example of a hearing test is given in Figure 11-37. In general, hearing profiles are intended to influence the Warfighter's occupation specialty to ensure that no further loss results to due duty and that no harm results due to the hearing loss.

The Department of Defense and the Occupational Health and Safety Association (OHSA) require Warfighters to have hearing threshold tests prior to hazardous noise exposure. For Warfighters routinely exposed to noise hazards, annual pure tone threshold tests are required (AR 40-501). Not only does the annual hearing test help determine the need for further audiology testing to evaluate fitness for duty status, but also monitors any change in hearing that may occur. Specifically, these hearing tests document if any significant changes that occur at 2000, 3000, and 4000 kHz. A *significant threshold shift* (STS) is defined as a 10 dB or more average shift at the aforementioned frequencies. A STS can be consistent with noise-induced hearing loss and can alert the Warfighter and his/her command to needed improvements in compliance with hearing conservation measures (i.e., hearing protection devices). Permanent noise-induced hearing loss is a pervasive hazard in the military but it is preventable.

REFERENCE AUDIOGRAM — 1. ZIP CODE/APO/FPO/PAS

(This form is subject to the Privacy Act of 1974 - use Blanket PAS - DD Form 2005)

AUDIOMETRY

1 — 1 - REFERENCE ESTABLISHED PRIOR TO INITIAL DUTY IN HAZARDOUS NOISE AREAS; 2 - REFERENCE ESTABLISHED FOLLOWING EXPOSURE IN NOISE DUTIES; 3 - REFERENCE RE-ESTABLISHED AFTER FOLLOW-UP PROGRAM

16. AUDIOMETRIC DATA RE: ANSI S3.6 - 1996	LEFT						RIGHT					
	500	1000	2000	3000	4000	6000	500	1000	2000	3000	4000	6000
17. DATE OF AUDIOGRAM 11-OCT-2006	-5	0	5	65	70	45	5	0	0	25	60	45

18. MEETS REFERRAL CRITERIA	19. MILITARY TIME OF DAY *(Optional)*	20. HOURS SINCE LAST NOISE EXPOSURE	21. EAR, NOSE, AND THROAT PROBLEM AT TIME OF TEST
2 (1 - NO, 2 - YES)	13:55:50	14	2 (1 - NO, 2 - YES, 3 - UNKNOWN)

Figure 11-37. Example of a hearing test from a Warfighter with an H-3 hearing test.

References

Abel, S.M. (1972). Discrimination of temporal gaps. *Journal of the Acoustical Society of America,* 52, 519-524.

Abouchacra, K., and Letowski, T. (1999). Comparison of air-conduction and bone-conduction hearing thresholds for pure tones and octave-band filtered sound effects. *Journal of the American Academy of Audiology,* 10, 422-428.

Abouchacra, K., Letowski, T., and Gothie, J. (2007). Detection and recognition of natural sounds. *Archives of Acoustics,* 32, 603-616.

American Speech-Language-Hearing Association. (2006). The Prevalence and Incidence of Hearing Loss in Adults. Retrieved October 1, 2006 from: http://www.asha.org/public/hearing/disorders/prevalence_adults.htm

Anderson, C.M.B., and Whittle, L.S. (1971). Physiological noise and the missing 6 dB. *Acustica,* 24, 261-272.

American National Standards Institute. (1989). Method for measuring the intelligibility of speech over communication systems. ANSI S3.2-1989. New York: American National Standards Institute (ANSI).

American National Standards Institute. (1994). Acoustical terminology. ANSI S1.1-1994 (R2004). New York: American National Standards Institute (ANSI).

American National Standards Institute. (1997). Methods for calculation of the speech intelligibility index. ANSI S3.2-1997 (R2007). New York: American National Standards Institute (ANSI).

American National Standards Institute. (2004). Specifications for Audiometers. ANSI S3.6-2004. New York: American National Standards Institute (ANSI).

American National Standards Institute. (2005). Measurement of sound pressure levels in air. ANSI S1.13-2005. New York: American National Standards Institute (ANSI).

American National Standards Institute. (2007). Procedure for the computation of loudness of steady sounds. ANSI S3.7-2007. New York: American National Standards Institute (ANSI).

Aures, W. (1984). Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, ein Beitrag zur gehörbezogenen Schallanalyse, Doctoral dissertation; Technische Univesität München.

Aures, W. (1985). Ein Berechnungsverfahren der Rauhigkeit. *Acustica*, 58, 268-281.

Ballas, J.M., and Howard, J.H. (1987). Interpreting the language of environmental sounds. *Environment and Behavior*, 19, 91-114.

Ballas, J.M., Dick, K.N., and Groshek, M.R. (1987). Failure to identify "identifiable" sounds. *Proceedings of the 31st Annual Meeting of the Human Factors Society* (p. 144-146). New York: Human Factors Society.

Ballas, J.M., and Barnes, M.E. (1988). Everyday sound perception and aging. *Proceedings of the 32nd Annual Meeting of the Human Factors Society* (p. 194-197). Anaheim, CA: Human Factors Society.

Ballas, J.M. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance,* 19, 250-267.

Batteau, D.W. (1967). The role of the pinna in human localization. *Proceedings of the Royal Society B.* 168, 158-180.

Beattie, R.C., and Culibrk, J. (1980). Effect of competing message on the speech comfortable loudness level for two instructional sets. *Ear and Hearing,* 1, 242-248.

Bench J., Kowal A., and Bamford J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology,* 13, 108-12.

Beranek, L.L. (1947). The design of speech communication systems. *Proceedings of the Institute of Radio Engineers*, 35, 880-890.

Beranek, L.L., Marshall, J.L., Cudworth, A., and Peterson, A.P.G. (1951). Calculation and measurement of the loudness of sound. *Journal of the Acoustical Society of America*, 23, 161-169.

Berger, K., (1969). A speech discrimination task using multiple-choice key words in sentences. *Journal of Auditory Research*, 9, 247-262.

Berlin, C.I., Bordelon, J., St. John, P., Wilensky, D., Hurley A., Kluka, E., and Hood, L.J. (1998). Reversing click polarity may uncover auditory neuropathy in infants. *Ear and Hearing,* 19, 37-47.

Best, V., Ozmeral, E., Gallun, F.J., Sen, K., and Shinn-Cunningham, B.G. (2005). Spatial unmasking of birdsong in human listeners: Energetic and informational factors. *Journal of the Acoustical Society of America,* 118, 3766-3773.

Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M., and Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research,* 27, 32-48.

Bindra, D., Williams, J.A., and Wise, J.S. (1965). Judgment of sameness and difference: Experiments in decision time. *Science,* 150, 1625-1627.

Bismarck, G. von. (1974a). Timbre of steady sounds, A factorial investigation of its verbal attributes. *Acustica,* 30, 146-158.

Bismarck, G. von. (1974b). Sharpness as an attribute of the timbre of steady sounds. *Acustica*, 30, 159-171.

Black, J.W. (1957). Multiple choice intelligibility tests. *Journal of Speech and Hearing Disorders,* 22(2), 213-235.

Blauert, J. (2001). *Spatial hearing: The psychophysics of human sound localization (3rd Ed.).* Cambridge, MA: MIT Press.

Block, G.B., Killion, M.C., and Tillman, T.W. (2004). The "Missing 6 dB" of Tillman, Johnson, and Olsen was found – 30 years ago. *Seminars in Hearing,* 25, 7-16.

Blue, M., Ntuen, C., and Letowski, T. (2004). Speech Intelligibility of the Callsign Acquisition Test in a Quiet Environment. *International Journal of Occupational Safety and Ergonomics,* 10, 179-189.

Boer, E. de. (1956). On the residue in hearing. Doctoral dissertation. Amsterdam: University of Amsterdam.

Boothroyd, A. (1968). Statistical theory of the speech discrimination score. *Journal of the Acoustical Society of America,* 43, 362-367.

Bond, Z.S., and Moore, T.J. (1994). A note on the acoustic-phonetic characteristics of inadvertently clear speech. *Speech Communication,* 14(4), 325-337.

Bradlow, A.R., Torretta, G.M., and Pisoni, D.B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication,* 20(3-4), 255-272.

Bregman, A.S. (1990). *Auditory Scene Analysis*. Cambridge (MA): MIT Press.

Brighouse, G., and Koh, S.D. (1950). The time error in aesthetic judgment. American Psychologist, 5, 317.

Brodgen, W.J., and Miller, G.A. (1947). Physiological noise generated under earphone cushions. *Journal of the Acoustical Society of America,* 19, 620-623.

Bronkhorst, A. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple talker conditions. *Acustica,* 86, 117-126.

Brungart, D.S. (1999). Auditory parallax effects in the HRTF for nearby sources. Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. New Paltz, NY: IEEE

Brungart, D.S. (2001). Evaluation of speech intelligibility with the coordinate response measure. *Journal of the Acoustical Society of America,* 109, 2276-2279.

Brungart, D.S., and Simpson, B.D. (2002). The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal. *Journal of the Acoustical Society of America,* 112, 664-676.

Brungart, D.S., and Scott, K.R. (2001). The effects of production and presentation level on the auditory distance perception of speech. *Journal of the Acoustical Society of America*, 110, 425-440.

Brunt, M. (1985). Bekesy audiometry and loudness balance testing. In: Katz, J. (Ed.), *Handbook of Clinical Audiology*. Baltimore, MD: Williams and Wilkins.

Bürck, W., Kotowski, P., and Lichte, H. (1936). Frequenzspektrum und Tonnerkennen. *Annalen der Physik,* 25, 433-449.

Buus, S., Schorer, E., Florentine, M., and Zwicker, E. (1986). Decision rules in detection of simple and complex tones. *Journal of the Acoustical Society of America,* 80, 1646-1657.

Canévet, G., Germain, R., Marchioni, A., and Scharf, B. (1981). Adaptation de sonie. *Acustica,* 49, 239-244.

Carney, L.H. (1994). Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness. *Hearing Research,* 76, 31-44.

Cassidy, J., and Ditty, K. (2001). Gender differences among newborns on a transient otoacoustic emissions test for hearing. *Journal of Music Therapy*, 37, 28-35.

Chi, S.-M., and Oh, Y.-H. (1996). *Lombard effect compensation and noise suppression for noisy Lombard speech recognition.* Paper presented at *The Fourth International Conference on Spoken Language Processing*, Philadelphia.

Ching, T.Y.C., Dillon, H., and Byrne, D. (1998). Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. *The Journal of the Acoustical Society of America,* 103(2), 1128-1140.

Choo, T.H. (1954). A study of time error in aesthetic judgment of auditory stimuli. Studies in Psychology EWHA Women's University, 1, 67-83.

Chocholle, R., and Krutel, J. (1968). Les seuils auditifs différentiels d'intensité en fonction de la durée des stimuli. *Comptes Rendus de la Société de Biologie*, 162, 848-885.

Clark, D. (1987). Listening test technology for automotive sound systems. SAE 1987 International Conference and Exhibition, Paper 870145. Detroit, MI: Society of Automotive Engineers (SAE).

Clifton, R.K. (1987) Breakdown of echo suppression in the precedence effect. *Journal of the Acoustical Society of America*, 82, 1834-1835.

Colburn, H.S. (1977). Theory of binaural interaction based on auditory-nerve data. II : Detection of tones in noise. *Journal of the Acoustical Society of America,* 61, 525-533.

Conway, A.R.A., Cowan, N., and Bunting, M.F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review,* 8, 331-335.

Cord, M.T., Walden, B.E., and Atack, R.M. (1992). Speech Recognition in Noise Test (SPRINT) for H-3 Profiles. Unpublished report. (Available from authors, Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, D.C. 20307-5001)

Corso, J. (19573). Age and sex differences in thresholds. *Journal of the Acoustical Society of America,* 31, 498-507.

Corso, J.H. (1963). Aging and auditory thresholds in men and women. *Archives of Environmental Health,* 6, 350-356.

Cox, R.M., Alexander, G.C., and Gilmore, C. (1987). Development of the Connected Speech Test (CST). *Ear and Hearing*, 8, 119S-126S.

Culling, J.F., Hawley, M.L., and Litovsky, R.Y. (2004). The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *Journal of the Acoustical Society of America,* 116, 1057-1065.

Dau, T. (1996). Modeling auditory processing of amplitude modulation. Ph.D. dissertation. Oldenburg (Germany): Carl von Ossietzky Universität Oldenburg.

Deary, I.J., Head, B., and Egan, V. (1989). Auditory inspection time, intelligence, and pitch discrimination. *Intelligence,* 13, 135-147.

Delany, M.E. (1970). *On the stability of auditory threshold. NPL Aero Report Ac44*. London: National Physical Laboratory.

Denenberg, L.J., and Altshuler, M.W. (1976). The clinical relationship between acoustic reflex and loudness perception. *Ear and Hearing,* 2, 79-82.

Department of Defense. (1999). Department of Defense design criteria standard: Human engineering. MIL-STD-1472F. Washington, DC: U.S. Department of Defense.

Department of the Army. (2008). Army Regulation (AR 40-501 - 2008). Standards of Medical Fitness.

Dirks, D., and Bower, D. (1968). Masking effect of speech competing messages. *Journal of Speech and Hearing Research,* 12, 229-245.

Dirks, D., and Kamm, C. (1976). Psychometric functions for loudness discomfort and most comfortable loudness level. *Journal of Speech, Language, and Hearing Research*, 19, 613-627

Divenyi, P.L. (1992). Binaural suppression of nonechoes. *Journal of the Acoustical Society of America*, 91, 1078-1084.

Dobie, R.A., and Van Hemel, S. (2004). *Hearing Loss.* Washington, D.C.: National Academies Press.

Doughty, J.M., and Garner, W.R. (1947) Pitch characteristics of short tones. II: Pitch as a function of duration. *Journal of Experimental Psychology*, 38, 478-494.

Dufour, A. (1999). Importance if atttentional mechanisms in audiovisual links. *Experimental Brain Research,* 126, 215-222.

Durlach, N. (1963) Equalization and cancellation theory of binaural masking level differences. *Journal of the Acoustical Society of America*, 35, 1206-1218.

Durlach, N., and Colburn, H.S. (1978). Binaural phenomena. In: Carterette, E.C., and Freidman, M.P. (Eds.), *Handbook of Perception*. New York: Academic Press.

Durlach, N.I., Mason, C.R., Kidd, G. Jr., Arbogast, T., Colburn, H.S., and Shinn-Cunningham, B.G. (2003). Note on informational masking (L). *Journal of the Acoustical Society of America*, 113, 2984-2987.

Durlach, N.I., Mason, C.R., Gallun, F.J., Shinn-Cunningham, B.G., Colburn, H.S., and Kidd, G., Jr. (2005). Informational masking for simultaneous nonspeech stimuli: Psychometric functions for fixed and randomly mixed maskers. *Journal of the Acoustical Society of America*, 118, 2482-2497.

Egan, J.P. (1948). Articulation testing methods. *Laryngoscope,* 58(9), 955-991.

Egan, J.P. (1965). Masking-level differences as a function of interaural disparity in intensity of signal and of noise. *Journal of the Acoustical Society of America,* 38, 1043-1049.

Egan, J.P., and Hake, H.W. (1950). O the masking pattern of simple auditory stimulus. *Journal of the Acoustical Society of America,* 22, 622-630.

Eisler, H. (1966). Measurements of perceived acoustic quality of sound-reproducing systems by means of factor analysis. *Journal of the Acoustical Society of America,* 39, 484-492.

Elliott, L.L. (1962). Backward and forward masking of probe tomes of different frequencies. *Journal of the Acoustical Society of America,* 34, 1116-1117.

Elliott, L.L., and Katz, D.R. (1980). *Northwestern University children's perception of speech (NU-CHIPS)*. St. Louis, MO: Auditec of St. Louis.

Emanuel, D., Letowski, S., and Letowski, T. (2009). The decibel. In: Emanuel, D., and Letowski, T., *Hearing Science*. Baltimore, MD: Lippincott, Williams, and Wilkins.

Erber, N.P. (1980). Use of the auditory numbers test to evaluate speech perception abilities of hearing-impaired children. *Journal of Speech and Hearing Disorders*, 45, 527-532.

Erber, N.P., and Witt, L. H. (1977). Effects of stimulus intensity on speech perception by deaf children. *Journal of Speech and Hearing Disorders,* 42, 271-278.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. *Pflüngers Archiv für de gesamte Physiologie,* 11, 403-432.

Fairbanks, G. (1958). Test of phonemic differentiation: The Rhyme Test. *Journal of the Acoustical Society of America,* 30, 596-600.

Fan, W.L., Streeter, T.M., and Durlach, N.I. (2008). Effect of spatial uncertainty of masker on masked detection for nonspeech stimuli (L). *Journal of the Acoustical Society of America*, 124, 36-39.

Fastl, H. (1976). Temporal masking effects: I. Broad band noise masker, *Acustica,* 35, 287-302.

Fastl, H., and Bechly, M. (1983). Suppression in simultaneous masking. *Journal of the Acoustical Society of America*, 74, 754-757.

Fastl, H., and Stoll, G. (1979). Scaling of pitch strength. *Hearing Research,* 1, 293-301.

Fay, D.D. (1988). *Hearing in Vertebrates: A Psychophysics Databook*. Winnetka (IL): Hill-Fay Associates.

Fechner, G.T. (1860). *Elemente der Psychophysik.* Leipzig (Germany) Breitkopf und Härtel. [Elements of Psychophysics (Vol. 1, Adler, H.E., translatior). New York: Holt, Rinehart, and Winston (1966)].

Feddersen, W.E., Sandel, T.T., Teas, D.C. and Jeffress, L A. (1957). Localization of high-frequency tones, *Journal of the Acoustical Society of America,* 29, 988-991.

Feeney, M.P., and Franks, J.R. (1982). Test-retest reliability of a distinctive feature difference test for hearing aid evaluation. *Ear and Hearing,* 3, 59-65.

Feldtkeller, J. and Zwicker, E. (1956). *Das Ohr als Nachrichtenempfanger*. Stuttgart, Germany: Hirzel.

Feuerstein, J. (2002). Occupational hearing conservation. In: Katz, J., Brukard, R., and Medwetsky, L. (Eds.), *Handbook of Clinical Audiology* (5th Ed.). Baltimore: Lippincott Williams and Wilkins.

Fidell, S., and Bishop, D.E. (1974). Prediction of acoustic detectability Technical Report 11949. Warren (MI): U.S. Army Tank Automotive Command.

Fifer, R.C., Jerger, J.F., Berlin, C.I., Tobey, E.A., and Campbell, J.C. (1983). Development of a dichotic sentence identification test for hearing-impaired adults. *Ear and Hearing,* 4, 300-305.

Finitzo-Hieber T., Gerling, I.J., Matkin, N.D., and Cherow-Skalka, E. (1980). A sound effects recognition test for the pediatric audiological evaluation. *Ear and Hearing,* 1, 271-276.

Fishman, Y., Bolkov, I.O., Noh, M.D., Garell, P.C., Bakken, H., Arezzo, J.C., Howard, M.A., and Steinschneider, M. (2001). Consonance and dissonance of musical chords: natural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology*, 86, 2761-1788.

Fitzgibons, P.J., and Gordon-Salant, S. (1998). Auditory temporal order perception in younger and older adults. *Journal of Speech and Hearing Science*, 41, 1052-1060.

Flanagan, J.L., and Guttman, N. (1960). On the pitch of periodic pulses. *Journal of the Acoustical Society of America*, 32, 1308-1319.

Fletcher, H. (1934). Loudness, pitch, and the timbre of musical tones and their relation to the intensity, the frequency, and the overtone structure. *Journal of the Acoustical Society of America*, 6, 59-69.

Fletcher, H. (1938). Loudness, masking and their relation to the hearing process and the problem of noise measurement. *Journal of the Acoustical Society of America*, 9, 275-293.

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics*, 12, 47-65.

Fletcher, H., and Munson, A. (1933). Loudness, its definition, measurement and calculation. *Journal of the Acoustical Society of America*, 5, 82-108.

Fletcher, H., and Munson, W.A. (1937). Relation between loudness and masking. *Journal of the Acoustical Society of America*, 19, 90-119.

Florentine, M., and Buus, S. (1984). Temporal gaps detection in sensorineural and simulated hearing impairment. *Journal of Speech and Hearing Research*, 27, 449-455.

Fuitt, K., Letowski, T., and Mermagen, T. (2003). Auditory performance in the open sound field. *Journal of the Acoustical Society of America*, 113, 2286.

Fraisse, P. (1982). Rhythm and tempo. In: Deutsch, D. (Ed.), *The Psychology of Music*. New York: Academic Press.

French, N.R., and Steinberg, J.C. (1947). Factors governing the intelligibility of speech sounds. *Journal of the Acoustical Society of America*, 19, 90-119.

Freyman, R.L., Clifton, R.K., and Litovsky, R.Y. (1991). Dynamic processes in the precedence effect. *Journal of the Acoustical Society of America*, 90, 874-884.

Fry, D.B. (1961). Word and sentence tests for use in speech audiometry. *Lancet*, 2, 197-199.

Fry, D.B., and Kerridge, P.M.T. (1939). *Tests for the hearing of speech by deaf people*. London: H.K. Lewis and Co.

Fryer, P.A., and Lee, R. (1980). Absolute listening tests – further progress. The 65th Audio Engineering Convention, preprint A-2. London: Audio Engineering Society (AES).

Gabrielsson, A. (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology*, 15, 63-72.

Gabrielsson, A., and Lindström, B. (1985). Perceived sound quality of high fidelity loudspeakers. *Journal of Audio Engineering Society*, 33, 33-53.

Gabrielsson, A., and Sjögren H. (1976). Preferred listening levels and perceived sound quality at different sound levels in “high-fidelity” sound reproduction. Report TA-82. Stockholm (Sweden): Karolinska Institutet.

Gabrielsson, A., and Sjögren H. (1979). Perceived sound quality of sound-reproducing systems. *Journal of the Acoustical Society of America*, 65, 1919-1933.

Gaeth, J.H. (1970). *A scale for testing speech discrimination: Final report, project no. RD-2277-S*. Washington: Division of Research and Demonstration Grants, U.S. Social and Rehabilitation Service.

Galilei, G. (1638). Dialogues concerning two new sciences. Translated by H. Crewe and A. de Salvio. New York: McGraw-Hill, 1963]

Galton, F. (1883). *Inquires into human faculty and its development*. London: J.M. Dent and Sons.

Gardner, H.J. (1971). Application of a high-frequency consonant discrimination word list in hearing-aid evaluation. *Journal of Speech and Hearing Disorders*, 36, 354-355.

Gardner, M.B. (1964). Effect of noise on listening levels in conference telephony. *Journal of the Acoustical Society of America*, 36, 2354-2362.

Gardner, M.B. (1964). Study of noise + other factors in conference telephony. *Journal of the Acoustical Society of America,* 36, 1036.

Gardner, M.B., and Gardner, R.S. (1973). Problem of localization in the median plane: effect of pinnae cavity occlusion, *Journal of the Acoustic Society of America,* 53, 400-408.

Garner, W.R., and Miller, G.A. (1947a). The masked threshold for pure tones as a function of duration. *Journal of the Acoustical Society of America,* 37, 293-303.

Garner, W.R., and Miller, G.A. (1947b). Differential sensitivity to intensity as a function of the duration of the comparison tone. *Journal of Experimental Psychology,* 34, 450-463.

Garstecki, D.C., and Bode, D.L. (1976). Aided and unaided narrow band noise thresholds in listeners with sensorineural hearing impairment. *Ear and Hearing,* 1, 258-262.

Gavrenau, V. (1966). Infra sons: Générateurs, détecteurs, propriétés physiques, effects biologiques. *Acustica*, 17, 1-10.

Gavrenau, V. (1968). Infrasound. *Science Journal*, 4, 33.

Gässler, G. (1954). Über die Hörschwelle für Schallereignisse mit verschieden breitem Frequentzspektrum. *Acustica,* 4, 408-414.

Glasberg, B.R., and Moore, B.C.J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research,* 47, 103-138.

Gockel, H., Moore, B.C.J., Carlyon, R.P., and Plack, C.J. (2007). Effect of duration on the frequency discrimination of individual partials in a complex tone and on the discrimination of fundamental frequency. *Journal of the Acoustical Society of America,* 121, 373-382.

Goldstein, J.L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America,* 54, 1496-1516.

Goldstein, M. (1995). Classification of methods used for assessment of text-to-speech systems according to the demands placed on the listener. *Speech Communication*, 16, 225-244.

Grau, J.W., and Kemler Nelson, D.G. (1988). The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness. *Journal of Experimental Psychology: General,* 117, 347-370.

Green, D.M. (1988). *Profile Analysis: Auditory Intensity Discrimination*. Oxford: Oxford University Press.

Greenwood, D.D. (1961a). Auditory masking and the critical band. *Journal of the Acoustical Society of America,* 33, 484-502.

Greenwood, D.D. (1961b) Critical bandwidth and frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America,* 33, 1344-1356.

Greenwood, D.D. (1990). A cochlear frequency-position function for several species – 29 years later. *Journal of the Acoustical Society of America,* 87, 2592-2605.

Griffiths, J.D. (1967). Rhyming minimal contrasts: a simplified diagnostic articulation test. *Journal of the Acoustical Society of America*, 42, 236-241.

Groben, L.M. (1971). Appreciation of short tones. *Proceedings of the 7th ICA Congress on Acoustics, Paper 19-H-6*. Budapest (Hungary): International Commission on Acoustics (ICA).

Gullick, W.L. (1971). *Hearing: Physiology and Psychophysics*. New York: Oxford University Press.

Gullick, W.L., Gescheider, G.A., and Frisina, R.D. (1989). *Hearing*. New York: Oxford University Press.

Guirao, M., and Stevens, S.S. (1964). Measurement of auditory density. *Journal of the Acoustical Society of America,* 36, 1176-1182.

Gygi, B., Kidd, G.R., and Watson, C.S. (2004). Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America*, 115, 1252-1265.

Gygi, B. (2001). Factors in the identification of environmental sounds. Ph.D. dissertation, Department of Psychology, Indiana University.

Gygi, B., and Shafiro, V. (2007). General functions and specific applications of environmental sound research. *Frontiers in Bioscience,* 12, 3152-3166.

Gygi, B., Kidd, G.R., and Watson, C.S. (2007). Similarity and categorization of environmental sounds. *Perception and Psychophysics,* 69, 839-855.

Haas, H. (1951). Über den Einfluss eines Einfachechos an die Hörsamkeit von Sprache. *Acustica,* 1, 49-58.

Hall, E.T. (1966). *Then Hidden Dimension*. New York: Anchor Books.

Hamilton, P.M. (1957). Noise masked threshold as a function of tonal duration and masking tome band width. *Journal of the Acoustical Society of America,* 29, 506-511.

Harrell, R. (2002). Pure tone evaluation. In: Katz, J., Burkard, R., and Medwetsky, L. (Eds.), *Handbook of Clinical Audiology* (5th Ed.). Baltimore, MD: Lippincott Williams and Wilkins.

Hartmann, W.M. (1997). *Signals, Sound, and Sensation*. Woodbury, NY: American Institute of Physics.

Hartmann, W.M., and Rakerd, B. (1989). Localization of sound in rooms. IV: The Franssen effect. *The Journal of the Acoustical Society of America.* 86, 1366-1373.

Haskins, H.L. (1949). A phonetically balanced test of speech discrimination for children. Unpublished master's thesis, Northwestern University, Evanston, IL.

Hawkes, R.J., and Douglas, H. (1971). Subjective acoustic experience in concert auditoria. *Acustica,* 24, 235-240.

Hawkins, J.E., and Stevens, S.S. (1950). The masking of pure tones and of speech by white noise. *Journal of the Acoustical Society of America,* 22, 6-13.

Hawley, M.L., Litovsky, R.Y., and Culling, J.F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *Journal of the Acoustical Society of America,* 115, 833-843.

Hebrank, J., and Wright, D. (1974). Spectral cues used in the localization of sound sources on the median plane, *Journal of the Acoustical Society of America,* 56, 1829-1834.

Heffner, H.E., and Heffner, R.S. (1998). Hearing. In: Greenberg, G., and Haraway, M.M. (Eds.), *A Handbook of Comparative Psychology*. New York: Garland.

Heffner, H.E., and Heffner, R.S. (2003). Audition. In: Davis, S. (Ed.), *Handbook of Research Methods in Experimental Psychology* (pp. 413-440). Hoboken, NJ: Blackwell Publishing.

Heffner, R.S., and Heffner, H.E. (1993). Degenerate hearing and sound localization in naked mole rats (Heterocephalus glaber) with an overview of central auditory structures. *Journal of Comparative Neurology,* 3321, 418-433.

Hellman, R. (1991). Loudness measurements by magnitude scaling: Implication for intensity coding. In: Bolanowski, S.J., and Gescheider, G.A. (Eds.), *Ratio Scaling of Psychological Magnitude*. Hillsdale, NJ: Erlbaum.

Hellman, R., and Zwislocki, J. (1961). Some factors affecting the estimation of loudness. *Journal of the Acoustical Society of America,* 60, 672-679.

Hellman, R., Miśkiewicz, A., and Scharf, B. (1997). Loudness adaptation and excitation patterns: Effects of frequency and level. *Journal of the Acoustical Society of America,* 101, 2176-2185.

Hellström, A. (1977). Time errors are perceptual. *Psychological Research,* 39, 345-388.

Helmholtz, H von. (1863). *Der Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig (Germany):Fredrich Viewwieg und Sohn. [On the Sensations of Tone as a Physiological Bases for the Theory of Music (Edited and translated by A. J. Ellis). New York: Dover, 1954].

Henning, G.B. (1974). Lateralization and binaural masking-level difference. *Journal of the Acoustical Society of America,* 55(6), 1259-1262.

Henning, G.B., and Bleiwas, S.L. (1967). Amplitude discrimination in noise. *Journal of the Acoustical Society of America,* 41, 1365-1366.

Henning, G.B., and Grosberg, S.L. (1968). Effect of harmonic components on frequency discrimination. *Journal of the Acoustical Society of America,* 44, 1386-1389.

Henry, J.A., Flick, C.L., Gilberty, A., Ellingson, R.M., and Fausti, S.A. (2001). Reliability of hearing thresholds: Computer-automated testing with ER-4B Canal Phone™ earphones. *Journal of Rehabilitation Research and Development,* 38, 1-18.

Henry, P., and Letowski, T. (2007). Bone conduction: Anatomy, physiology and communication. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory. ARL Technical Report ARL-TR-4138.

Hirsh, I. (1948). The influence of interaural phase on interaural summation and inhibition. *Journal of the Acoustical Society of America,* 20, 536-544.

Hirsh, I. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America,* 31, 759-767.

Hirsh, I.J. (1971). Masking of speech and auditory localization. *Audiology*, 10, 110-114.

Hirsh, I.J., Davis, H., Silverman, S.R., Reynolds, E.G., Eldert, E., and Benson, R.W. (1952). Development of materials for speech audiometry. *Journal of Speech and Hearing Disorders,* 17, 321-337.

Hirsh. I., and Sherrick, C.E., Jr. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology*, 62, 423-432.

Hochberg, I. (1975). Most comfortable listening for the loudness and intelligibility of speech. *Audiology,* 14, 27-33.

Hood, J.C. (1950). Studies on auditory fatigue and adaptation, *Acta Otolaryngologica*, (Suppl.) 92, 1-57.

Hood, J.D., and Poole, J.P. (1980). Influence of the speaker and other factors affecting speech-intelligibility *Audiology,* 19(5), 434-455.

Hochhaus, L., and Antes, J.R. (1973). Speech identification and knowing that you know. *Perception and Psychophysics*, 13(1A), 131-132.

House, A.S., Williams, C.E., Hecker, M.H.L., and Kryter, K.D. (1965). Articulation-testing methods: Consonantal differentiation with a closed-response set. *Journal of the Acoustical Society of America,* 37, 158-166.

Houtgast, T. (1974). Lateral suppression and loudness reduction of a tone in noise. *Acustica,* 30, 215-221.

Howes, W.L. (1971). Loudness determined by power summation. *Acustica,* 25, 343-248.

Hudgins, C.V., Hawkins, J.E., Karlin, J.E., and Stevens, S.S. (1947). The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope,* 57(1), 57-89.

Illényi, A., and Korpassy, P. (1981). Correlation between loudness and quality of stereophonic loudspeakers. *Acustica,* 49, 334-336.

International Electrotechnical Commission. (1995). IEC 50801, International Electrotechnical Vocabulary – Chapter 801: Acoustics and Electroacoustics. Geneva: International Electrotechnical Commission.

International Organization for Standardization. (1975). ISO 532, Acoustics – Method for calculating loudness level. Geneva: International Organization for Standardization (ISO).

International Organization for Standardization. (2000). ISO 7029, Acoustics – Statistical distribution of hearing thresholds as a function of age. Geneva: International Organization for Standardization (ISO).

International Organization for Standardization. (2003). ISO 226 (R2003), Acoustics - Normal equal-loudness level contours. Geneva: International Organization for Standardization (ISO).

International Organization for Standardization**.** (2005). ISO 389-7, Acoustics - Reference zero for the calibration of audiometric equipment - Part 7: Reference threshold of hearing under free-field and diffuse-field listening conditions. Geneva: International Organization for Standardization (ISO).

International Organization for Standardization. (2006). ISO 16832, Acoustics - Loudness scaling by means of categories. Geneva: International Organization for Standardization (ISO).

International Telecommunication Union - Telecommunication. (1996). ITU-T Recommendation P.800. Methods for subjective determination of transmission quality. Geneva: International Telecommunication Union – Telecommunication (ITU-T).

Jaroszewski, A., and Rakowski, A. (1976). Pitch shifts in post-stimulatory masking. *Acustica,* 34, 220-223.

Jerger, S., Lewis, S., Hawkins, J. and Jerger, J. (1980). Pediatric speech intelligibility test. I. Generation of test materials. *International Journal of Pediatric Otorhinolaryngology,* 2, 217-30.

Jesteadt, W., Bacon, S.P., and Lehman, J.R. (1982). Forward masking as a function of frequency, masker level, and signal delay. *Journal of the Acoustical Society of America,* 71, 950-962.

Junqua, J. (1993). The Lombard reflex and its role on human listeners and automatic speech recognizers, *Journal of the Acoustical Society of America,* 93, 510-524.

Junqua, J. (1996). The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex, *Speech Communication,* 20, 13-22.

Justus, T.C., and Bharucha, J.J. (2002). Music perception and cognition. In: Yantis, S., and Pashler, H. (Eds.), *Stevens' Handbook of Experimental Psychology*, *Sensation and Perception,* I, 453-492.

Kaernbach, C., and Bering, C. (2001). Exploring the temporal mechanism involved in the pitch of unresolved harmonics. *Journal of the Acoustical Society of America,* 110, 1039-1048.

Kalikow, D.N., Stevens, K.N., and Elliott, L.L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America,* 61, 1337-1351.

Kant, I. (1781). Kritik der reinen Vernunft. Riga: J.F. Hartknoch [Critique of Pure Reason. English translation: Werner Pluhar. Indianopolis: Hackett, 1996].

Katz, D.R., Elliott, L.L. (1978). Development of a new children's speech discrimination test. Paper presented at the 1978 *Annual American Speech-Language-Hearing Association Convention*. Chicago, IL.

Keating, P.A., and Manis, F. (11998). The Keating-Manis phoneme deletion test. *UCLA Working Papers in Phonetics,* 96, 162-165.

Keith, R.W. (1977). An evaluation of predicting hearing loss from the acoustic reflex. Archives of Otolaryngology, 103, 419-424.

Keith, R.W., Young, M., and McCroskey, R. (1999). A brief introduction to the Auditory Fusion Test – Revised. *Educational Audiology Review*, 16, 2.

Kidd, G. Jr., Mason, C.R., Deliwala, P.S., Woods, W.S., and Colburn, H.S. (1994). Reducing informational masking by sound segregation. *Journal of the Acoustical Society of America*, 95, 3475-3480.

Kidd, G. Jr., Mason, C.R., Richards, V.M., Gallun, F.J., and Durlach, N.I. (2007). Informational masking. In: Yost, W. (Ed.), *Springer Handbook of Auditory Research*, 29, 143-190. New York: Springer Verlag.

Killion, M.C. (1978). Revised estimate of minimum audible pressure: Where is the "missing 6 dB"? *Journal of the Acoustical Society of America*, 63, 1501-1508.

Kobayashi, M., Morimoto, M., Sato, H., and Sato, H. (2007). Optimum speech level to minimize listening difficulty in public spaces. *Journal of the Acoustical Society of America,* 121, 251-256.

Koester, T. (1945). Time-error and sensitivity in pitch and loudness discrimination as a function of time interval and stimulus level. *Archives of Psychology,* 297, 69.

Koester, T., and Schoenfeld W.N. (1946). The effect of context upon judgment of pitch differences. *Journal of Experimental Psychology,* 36, 417-430.

Koh, S.D. (1962). The Fechnerian time error and Pratt's time error in affective judgments of musical excerpts. *American Psychologist*, 17, 369.

Koh, S.D. (1967). Time-error in comparison of preferences for musical excerpts. *American Journal of Psychology*, 80, 171.

Kohler, W. (1923). Zur theorie des sukzessivvergleichs und der zeitfehler. *Psychologische Forshung,* 4, 115-175.

König, E. (1957). Effect of time on pitch discrimination thresholds under several psychophysical procedures. *Journal of the Acoustical Society of America*, 29, 606-612.

Kopčo, N., and Shinn-Cunningham, B.G. (2008). Influences of modulation and spatial separation on detection of a masked broadband target. Journal *of the Acoustical Society of America,* 124, 2236-2250.

Kopra, L.L., and Blosser. D. (1968). Effects of method of measurement on most comfortable loudness level for speech. *Journal of Speech and Hearing Research*, 11, 497-508.

Krause, B. (2008). Interview: From rock'n'roll to the sounds of nature. *New Scientist,* 2664, 42-43

Kryter, K.D. (1962a). Methods for the calculation and use of the Articulation Index, *Journal of the Acoustical Society of America*, 34, 1689-1697.

Kryter, K.D. (1962b). Validation of the articulation index. *Journal of the Acoustical Society of America,* 34, 1698-1702.

Kryter, K.D. (1965). Some comparisons between rhyme and PB-word intelligibility test. *Journal of the Acoustical Society of America,* 37, 1146.

Kuriyagawa, M., and Kameoka, A. (1966). PSE tracing method for subjective harmonics measurement and monaural phase effect on timbre. *Journal of the Acoustical Society of America,* 39, 1263-1263 (A).

Lamore, J.J. (1975). Perception of two-tone octave complexes. *Acustica*, 34, 1-14.

Lane, H.L., and Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, 14, 677-709.

Leakey, D.M., Sayers, B.M., Cherry, C. (1958). Binaural fusion of low-frequency and high-frequency sounds. *Journal of the Acoustical Society of America*, 30(3), 222-223.

Lenhardt, M., Skellett, R., Wang, P., and Clarke, A. (1991). Human ultrasonic speech perception. *Science*, 253, 82-85.

Leshowitz, B. (1971). Measurement of the two-click threshold. *Journal of the Acoustical Society of America,* 49, 462-466.

Letowski, T. (1982). A note on the differential limen for frequency differentiation, *Journal of Sound and Vibration*, 85, 579-583.

Letowski, T. (1984). Auditory assessment of acoustic signals and audio equipment. *Chopin Academy Scientific Series,* 8, 1-401. (Monograph).

Letowski, T. (1985). Development of technical listening skills: Timbre Solfeggio. *Journal of Audio Engineering Society*, 33, 240-244.

Letowski, T. (1989). Sound quality assessment: Concepts and criteria. Paper D-8 presented at the 87th Audio Engineering Convention (Preprint # 2825). New York: Audio Engineering Society (AES).

Letowski, T. (1992). Timbre, tone color, and sound quality: Concepts and definitions. *Archives of Acoustics*, 17, 17-31.

Letowski, T. (1994). Guidelines for conducting listening tests on sound quality. Proceedings of the Noise-Con 1994 Conference. Fort Lauderdale, FL: International Institute of Sound and Vibration.

Letowski, T. (1995). Sound quality scales and systems. *Proceedings of the VI Symposium on Tonmeistering and Sound Engineering*. Warszawa, Poland: Polskie Radio S.A.

Letowski, T., and Amrein, K. (2005). The effects of auditory training on the listener's ability to detect and recognize sounds in noise. *Proceedings of the 12th International Congress on Sound and Vibration (ICSV).* Lisbon, Portugal: International Institute of Acoustics and Vibration (IIAV).

Letowski, T., and Dreisbach L. (1992). Pleasantness and unpleasantness of speech. Proceedings of the 11th International AES Conference, pp. 159-165. Portland (OR): Audio Engineering Society (AES).

Letowski, T., Frank T., and Caravella J. (1993). Acoustical properties of speech produced in noise. *Ear and Hearing,* 14(5), 332-338.

Letowski, T., and Makowski, W. (1977). Właściwości obrazów słuchowych [Properties of auditory images]. *Technika Radia i TV*, 1, 16-21.

Letowski, T., and Miskiewicz, A. (1995). Development of technical listening skills for sound quality assessment. Proceedings of the Inter-Noise 1995. Newport Beach, CA: Inter-Noise.

Letowski, T., and Rakowski A. (1971). The DL for loudness of musical sounds. *Proceedings of the VII ICA Congress*, 19-H-5, 325-328. Budapest, Hungary: International Commission on Acoustics (ICA).

Letowski, T., and Smurzyński, J. (1980). Time error in perception of sound brightness. *Archives of Acoustics,* 5, 143-146.

Letowski, T., and Smurzynski, J. (1983). Evaluation of electroacoustic devices by equivalent scale methods. *Archives of Acoustics*, 8, 133-154.

Letowski, T., Tornatore, A., Clark, J., and MacBeth, B. (1993). Word discrimination in various multitalker noises. *Journal of the Acoustical Society of America*, 94(4/Pt.2), 1778

Letowski, T., Tornatore, A., MacBeth, B., Salter, S., and Smeal, M. (2001). Acoustic and perceptual properties of several multitalker noises. Proceedings of the 8th International Congress on Sound and Vibration, p. 2467-2474 [Hong Kong: July 2-6, 2001], Hong Kong: The Hong Kong Polytechnic University.

Levitt, H., and Webster, J.C. (1997). Effects of noise and reverberation on speech. In: Harris, C.M. (Ed.), *Handbook of Acoustical Measurements and Noise Control*, New York: Acoustical Society of America.

Liang, C.A., and Christvich, L.A. (1961). Frequency difference limens as a function of tonal duration. *Soviet Physical Acoustics*, 6, 75-80.

Liberman, A., Coper, F.S., Shankweiler, D., and Studert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.

Lichte, W.H. (1941). Attributes of complex tones. *Journal of Experimental Psychology,* 28, 455-481.

Licklider, J.C.R. (1948). The influence of interaural phase relations upon the masking of speech by white noise. *Journal of the Acoustical Society of America*, 20, 150-159

Licklider, J.C.R. (1951). Basic correlates of the auditory stimulus. In: Stevens, S.S. (Ed.), *Handbook of Experimental Psychology*. New York: Wiley and Sons.

Licklider, J.C.R. (1954). "Periodicity" pitch and "place" pitch. *Journal of the Acoustical Society of America*, 26, 945.

Ling, D., and Ling, A. (1978). Aural habilitation: The verbal foundations of learning in hearing-impaired children. Washington, DC: A.G. Bell.

Little, A.D., Mershon, D.H., and Cox, P.H. (1992). Spectral content as a cue to perceived auditory distance. *Perception*, 21, 405-416.

Lombard, E. (1911). Le signe de l'élévation del la voix. *Annales del Maladies de l'Oreille et du Larynx,* 37, 101-119.

Lonsbury-Martin, B., and Martin, K. (1993). Auditory dysfunction from excessive sound stimulation. In: Cummings, C., and Fredrickson, J. (Eds.), *Otolaryngology-Head and Neck Surgery*, 4, 2885-2990. St. Louis: Mosby Year Book.

Lubman, D. (1992). Objective metrics for characterizing automotive interior sound quality. *Proceedings of the Inter-Noise 1992 Conference*, 139, 1067. Toronto: Institute of Noise Control Engineering.

Lufti, R.A. (1986). Two- versus four-tone masking, revisited. *Journal of the Acoustical Society of America,* 80, 422-428.

Lufti, R.A. (1990). How much masking is informational masking? *Journal of the Acoustical Society of America*, 88, 2607-2610

Lundeen, C., and Small, A.M. (1984). Te influence of temporal cues on the strength of periodicity pitch. *Journal of the Acoustical Society of America*, 75, 1578-1587.

Makous, J.C., and Middlebrooks, J.C. (1990). Two-dimensional sound localization by human listeners, *Journal of the Acoustical Society of America*, 87, 2188-2200.

Marks, L.E. (1978). Binaural summation of the loudness of pure tones. *Journal of the Acoustical Society of America*, 64, 107-113.

Martin, M.C., and Grover, B.C. (1976). Current hearing aid performance and the requirements of hearing impaired persons. *Journal of the Audio Engineering Society*, 24, 177-181.

Marozeau, J., Epstein, M., Florentine, M., and Daley, B. (2006). The test of the Binaural Equal-Loudness-Ratio hypothesis for tones. *Journal of the Acoustical Society of America*, 120, 3870-3877.

Massaro, D. (1975). Perceptual processing time in audition. *Journal of the Acoustical Society of America*, 57, S5 (abstract).

Masterton, B., Heffner, H.E., and Ravizza, R. (1969). The evolution of human hearing. *Journal of the Acoustical Society of America*, 45, 966-985.

Mauk, M.D., and Buonomano, D.V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience*, 27, 307-340.

McAdams, S. (1984). Spectral fusion, spectral parsing, and the formation of auditory images. Technical Report STAN-M-22. Stanford, CA: Stanford University.

McBride, M.E, Letowski, T., and Tran, P. (2005). Search for the optimum vibrator location for bone conduction communication. Proceedings of the HFES 49th Annual Meeting, CD ROM, Orlando (FL): September 26-30, 2005.

McBride, M., Letowski, T., and Tran, P. (2008). Bone conduction reception: Head sensitivity mapping. *Ergonomics*, 55, 702-718.

McBride, M., Letowski, T., and Tran, P. (2008). Head mapping: Search for an optimum bone microphone placement. *Proceedings of the 42nd Annual HFES Meeting*. New York (NY): Human Factors Society (HFS).

McDermott, J.B. (1969). Multidimensional analyses of circuit quality judgments. *Journal of the Acoustical Society of America*, 45, 774-781.

McDonald, J.J., Teder-Sälejärvi, W.A., DiRusso, F., and Hillyear, S.A. (2005).Neural basis of auditory-induced shifts in visual time-order perception. *Nature Neuroscience (Online)*, 1-6.

McGill, W.J., and Goldberg, J.P. (1968). Pure-Tone Intensity Discrimination and Energy Detection. *The Journal of the Acoustical Society of America*, 44, 576-581.

McGregor, P., Horn, A.G., and Todd, M.A. (1985). Are familiar sounds ranged more accurately? *Perceptual and Motor Skills*, 61, 1082.

McPherson, D.F., and Pang-Ching, G.K. (1979). Development of a distinctive feature discrimination test. *Journal of Auditory Research*, 19, 235-246.

Mershon, D.H., Ballenger, W.L., Little, A.D., McMurtry, P.L., and Buchanan, J.L. (1989). Effects of room reflectance and background noise on perceived auditory distance. *Perception*, 18, 403-416.

Mershon, D.H., and King, L.E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception and Psychophysics*, 18, 409-415.

Michaels, R.M. (1957). Frequency difference limens for narrow bands of noise. *Journal of the Acoustical Society of America*, 29, 520-522

Micheyl, C., Arthaud, P., Reinhart, C., and Collet, P. (2000). Informational masking in normal-hearing and hearing-impaired listeners. Acta Otolarngologica, 120, 242-246.

Miller, G.A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *Journal of the Acoustical Society of America*, 19, 609-619.

Miller, G.A., and Taylor, G. (1948). The perception of repeated bursts of noise. *Journal of the Acoustical Society of America*, 20, 171-182.

Miller, G.A., Heise, G.A., and Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41(5), 329-335.

Mills, A.W. (1972). Auditory localization. In: Tobias, J.V. (Ed.), *Foundations of Modern Auditory Theory*, II, 301-348. New York: Academic Press.

Miskolczy-Foder, F. (1959). Relation between loudness and duration of tone pulses. *Journal of the Acoustical Society of America*, 31, 1128-1134.

Miśkiewicz, A., Scharf, B., Hellman, R., and Meiselman, C. (1992). Loudness adaptation at high frequencies. *Journal of the Acoustical Society of America*, 94, 1281-1286.

Mitrinowicz, M.A., and Letowski, T. (1966). Auditory threshold measurements with one-third octave noise bands. *Archiwum Akustyki*, 1, 103-112.

Møller, H., and Pedersen, C.S. Hearing at low and infrasonic frequencies. (2004). *Noise Health*, 6, 37-57.

Moore, B.C.J. (1973). Frequency difference limes for short-duration tones. *Journal of the Acoustical Society of America*, 54, 610-619.

Moore, B.C.J. (1977). Effects of relative phase of the component on the pitch of three-component complex tones. In: Evans, E.F., and Wilson, J.P. (Eds.), *Psychophysics and Physiology of Hearing*. London: Academic Press.

Moore, B.C.J. (1997). *An introduction to the psychology of hearing* (4th Ed.). San Diego, CA: Academic Press.

Moore, B.C.J., and Glasberg, B.R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America,* 88, 750-753.

Moore, B.C.J., and Glasberg, B.R. (1996). Revision of Zwicker's Loudness Model. *Acta Acustica*, 82, 335-345.

Moore, B.C.J., and Glasberg, B.R., and Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Acoustical Society of America,* 45, 224-240.

Moore, B.C.J., and Peters, R.W. (1992). Pitch discrimination and phase sensitivity i9n young and elderly subjects and its relationship to frequency selectivity. *Journal of the Acoustical Society of America*, 91, 2881-2893.

Morgan, D.E., Dirks, D., Bower, D., and Kamm, C. (1979). Loudness discomfort level and acoustic reflex threshold. *Journal of Speech and Hearing Research,* 22, 849-861.

Murphy M.P., and Gates, G.A (1997). Hearing loss: Does gender play a role? *Medscape Women's Health eJournal*, 2, 5.

Murphy, W., Themann, C., and Stephenson, M. (2006). Hearing levels in U.S. adults aged 20-69 Years – National Health and Nutrition Examination Survey (NHANES) 1999-2004. *Journal of the Acoustical Society of America,* 119, 3268-3268.

Murray, B.A., Smith, K.A., and Murray, G.G. (2000). The test of phoneme identities: Predicting alphabetic insight in prealphabetic readers. *Journal of Literacy Research*, 32, 421-447.

Musicant, A.D., and Butler, R.A. (1985). Influence of monaural spectral cues on binaural localization, *Journal of the Acoustical Society of America,* 77, 202-208.

Myers, L., Letowski, T., Abouchacra, K., Haas, E., and Kalb, J. (1996). Detection and recognition of filtered sound effects. *Journal of the American Academy of Audiology,* 7, 346-357.

Needham, J.G. (1935). Contrast effect in judgment of auditory intensities. *Journal of Experimental Psychology,* 18, 214-226.

Neuhoff, J.G. (2004). Ecological psychoacoustics: Introduction and history. In: J.G. Neuhoff, J.G. (Ed.), *Ecological Psychoacoustics*. Amsterdam: Elsevier.

Neuhoff, J.G., and McBeath, M.K. (1996). The Doppler illusion: The influence of dynamic intensity change on perceived pitch. *Journal of Experimental Psychology-Human Perception and Performance,* 22, 970-985.

Nielsen, S.H. (1993). Auditory distance perception in different rooms. *Journal of the Audio Engineering Society,* 41, 755-770.

Nilsson, M., Soli, S.D., and Sullivan, J.A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America,* 95, 1085-1099.

Nofsinger, D., Martinez, C.D., and Schaefer, A.B. (1985). Puretone techniques in evaluation pf central auditory function. In: Katz, J. (Ed.), *Handbook of Clinical Audiology,* 18, 337-354. Baltimore, MD: Williams and Wilkins.

Nye, P.W., and Gaitenby, J.H. (1974). The intelligibility of synthetic monosyllabic words in short, syntactically normal sentences. Haskins Labs Status Report on Speech Research, 37/38, 169-190.

Oldfield, S.R., and Parker, S.P.A. (1984). Acuity of sound localisation: A topography of auditory space. I. Normal hearing conditions, *Perception*, 13, 581-600.

Olson, H.F. (1967). *Music, Physics and Engineering*. Dover: New York

Owens, E., and Schubert, E.D. (1977). Development of the California Consonant Test. *Journal of Speech and Hearing Research*, 20, 463-474.

Ostroff, J.M., McDonald, K.L., Schneider, B.S., and Alain, C. (2003). Aging and the processing of sound duration in human auditory cortex. *Hearing Research,* 181, 1-7

Palmer, A.R., and Russell, I.J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells, *Hearing Research,* 24, 1-15.

Patterson, R.D. (1987). A pulse ribbon model of monaural phase perception. *Journal of the Acoustical Society of America,* 82, 1560-1586.

Pearson, K.S., Bennett, R.L., and Fidell, S. (1977). Speech levels in various noise environments. Report No. EPA-600/1-77-025; Washington, DC: U.S. Environmental Protection Agency (EPA).

Pederson, O.T., and Studebaker, G.A. (1972). A new minimal contrasts closed-response-set speech test. *Journal of Auditory Research*, 12, 187-195.

Perrott, D.R., and Musicant, A.D. (1977). Minimum auditory movement angle: Binaural localization of moving sound sources. *The Journal of the Acoustical Society of America,* 62, 1463-1466.

Perrott, D.R., and Musicant, A.D. (1981). Dynamic minimum audible angle: Binaural spatial acuity with moving sound sources, *Journal of Auditory Research,* 21, 287-295.

Peterson, G.E., and Lehiste I. (1962). Revised CNC lists for auditory tests. *Journal of Speech and Hearing Disorders,* 27, 62-70.

Pfafflin, S.M. (1968). Detection of auditory signals in restricted sets of reproducible noise. *Journal of the Acoustical Society of America,* 43, 487-490.

Pfafflin, S.M., and Matthews, M.V. (1966). Detection of auditory signals in reproducible noise. *Journal of the Acoustical Society of America,* 39, 340-345.

Picheny, M.A., Durlach, N.I., and Braida, L.D. (1985). Speaking clearly for the hard of hearing. 1. Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research,* 28(1), 96-103.

Picheny, M.A., Durlach, N.I., and Braida, L.D. (1986). Speaking clearly for the hard of hearing. 2. Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research,* 29(4), 434-446.

Picheny, M.A., Durlach, N.I., and Braida, L.D. (1989). Speaking clearly for the hard of hearing. 3. An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. *Journal of Speech and Hearing Research,* 32(3), 600-603.

Pick, J.H.L., Siegel, G.M., Fox, P.W., Garber, S.R., and Kearney, J.K. (1989). Inhibiting the Lombard effect. *The Journal of the Acoustical Society of America*, 85, 894-900.

Pickles, J. (1988). *An introduction to the physiology of hearing.* San Diego: Academic Press Inc.

Plenge, G. (1982). Über das Problem der Im-Kopf-Localisation. *Acustica,* 26, 241-252.

Plenge, G., and Brunschen, G. (1971) A priori knowledge of the signal when determining the direction of speech in the median plane. Proceedings of the 7th International Congress on Acoustics, (ICA), Paper19-H-10. Budapest, Hungary: ICA.

Plomp, R. (1964). Rate of decay of auditory sensation. *Journal of the Acoustical Society of Americ*a, 36, 277-282.

Plomp, R. (1970). Timbre as a multidimensional attribute of complex tones, In: Plomp, R., and Smoorenburg, G.F. (Eds.), *Frequency Analysis and Periodicity Detection in Hearing,* Leiden: Sijthoff.

Plomp, R. (1976). *Aspects of Tone Sensation: A Psychophysical Study*. London: Academic.

Plomp, R., and Bouman, M.A. (1959). Relation between hearing threshold and duration for tone pulses. *Journal of the Acoustical Society of America,* 31, 749-758.

Plomp, R., and Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America,* 38, 548-560.

Plomp, R., Pols, L.C.W., and van der Geer, J.P. (1967). Dimensional analysis of vowel tones, *Journal of the Acoustical Society of America,* 41, 707-712.

Pollack, I. (1948). Monaural and binaural threshold sensitivity for tones and for white noise. *Journal of the Acoustical Society of America*, 20, 52-57.

Pollack, I. (1951). Sensitivity to differences in intensity between repeated bursts of noise. *Journal of the Acoustical Society of America*, 23, 650-653.

Pollack, I. (1952). Loudness of bands of noise. *Journal of the Acoustical Society of America*, 24, 533-538.

Pollack, I. (1954). Intensity discrimination thresholds under several psychophysical procedures. *Journal of the Acoustical Society of America*, 26, 1056-1059.

Pollack, I. (1975). Auditory informational masking. *Journal of the Acoustical Society of America,* 57(Suppl.1), S5.

Pols, L.C.W., van der Kamp, L.J.T. and Plomp, R. (1969). Perceptual and physical space of vowel sounds, *Journal of the Acoustical Society of America*, 46, 458-467.

Postman, L. (1946). The time-error in auditory perception. *American Journal of Psychology*, 53, 193-219.

Pratt, R., and Doak, P. (1976). A subjective rating scale for timbre. *Journal of Sound and Vibrations*, 45, 317-321.

Price, G.R., and Hodge, D.C. (1976). Combat sound detection: Monaural listening in quiet. Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory. HEL Technical Memorandum HEL-TM-35-76.

Price, G.R., Kalb, J.T., and Garinther, G.R. (1989). Toward a measure of auditory handicap in the Army. *Annals of Otology, Rhinology and Laryngology*, 98, 42-52.

Rabin, J. (1995). Two eyes are better than one: binocular enhancement in the contrast domain. *Ophthalmic and Physiological Optics*, 15(1), 45-48.

Rajan, R., and Cainer, K.E. (2008). Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. *Neuroscience*, 154, 784-795.

Rakowski, A. (1977). Measurements of pitch. *Catgut Acoustical Society Newsletter, 27*, 9-11.

Rakowski, A. (1978). *Kategorialna percepcja wysokości dźwięku w muzyce* [Categorical perception of pitch in music]. Warsaw, Poland: Państwowa Wyższa Szkoła Muzyczna.

Rakowski, A. (1978). Kategoriala percepcja wysokosci dzwieku w muzyce. Habilitation dissertation. Warszawa, Poland: Panstwowa Wyzsza Szkola Muzyczna.

Rammsayer, T., and Lustnauer, S. (1989). Sex differences in time perception. *Perceptual and Motor Skills*, 68, 195-198.

Rao, M., and Letowski, T. (2004). Callsign Acquisition Test (CAT): Speech intelligibility in noise. *Ear and Hearing*, 27, 120-128.

Rappaport, J., and Provencal, C. (2002). Neurotology for audiologists. In: Katz, J., Brukard, R., and Medwetsky, L. (Eds.), *Handbook of Clinical Audiology* (5th Ed.). Baltimore: Lippincott Williams and Wilkins.

Rayleigh, L. (1907). On our perception of sound direction, *Philosophical Magazine*, 13, 214-232.

Relkin, E.M., and Doucet, J.R. (1997). Is loudness simply proportional to the auditory nerve spike count? *Journal of the Acoustical Society of America*, 101, 2735-2740.

Resnick, S.B., Dubno, J.R., Hoffnung, S., and Levitt, H. (1975). Phoneme errors on a nonsense syllable test. *Journal of the Acoustical Society of America*, 58, S114.

Richards, A.M. (1975). Most comfortable loudness for pure tones and speech in the presence of the masking noise. *Journal of the Speech and Hearing Research*, 18, 498-505.

Riesz, R.R. (1928). Differential intensity sensitivity of the ear. *Physics Review*, 31, 867-875.

Ritsma, R.J. (1967). Frequencies dominant in the perception of the pitch of complex sounds. *Journal of the Acoustical Society of America*, 42, 191-198.

Ritsma, R.J., and Bilsen, F.A. (1970). Spectral regions dominant in the perception of repetition pitch. *Acustica*, 23, 334-340.

Robinson, D.W. (1986). Sources of variability in normal hearing sensitivity. *Proceedings of the 12th International Congress on Acoustics (paper B11-1)*. Toronto: ICA.

Robinson, D.W., and Dadson, R.S. (1957). Equal-loudness relations and threshold of hearing for pure tones. *Journal of the Acoustical Society of America*, 29, 1284-1288.

Roederer, J.G. (1974). *Introduction to the Physics and Psychophysics of Music*. New York: Springer-Verlag.

Rose, J.E., Brugge, J.F., Anderson, D.J. and Hind, J.E. (1968). Patterns of activity in single auditory nerve fibres of the squirrel monkey. In: de Reuck, A.V.S., and Knight, J. (Eds.), *Hearing Mechanisms in Vertebrates*, London: Churchill.

Ross, M., and Lerman, J. (1970). A picture identification test for hearing-impaired children. *Journal of Speech and Hearing Research*, 13, 44-53.

Rossing, T. (1989). *The Science of Sound*. Reading, MA: Addison-Wesley Publishing Company.

Rosenzweig, M.R., and Postman, L. (1957). Intelligibility as a function of frequency of usage. *Journal of Experimental Psychology*, 54(6), 412-422.

Roush, J., and Tait, C.A. (1984). Binaural fusion, masking level difference, and auditory brainstem response in children with language-learning disabilities. *Ear and Hearing*, 5, 37-41.

Sachs, M.B., and Kiang, N.Y.S. (1967). Two-tone inhibition in auditory-nerve fibers. *Journal of the Acoustical Society of America,* 43, 1120-1128.

Sakaguchi, S., Arai, T., and Murahara, Y. (2000). The effect of polarity inversion of speech on human perception and data hiding as an application. Proceedings of the IEEE 2000 International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2000), 2, 917-920. Istanbul (Turkey): ICASSP.

Sammeth, C.A., Birman, M., and Hecox, K.E. (1989). Variability of most comfortable and uncomfortable loudness levels to speech stimuli in the hearing impaired. *Ear and Hearing,* 10, 94-100.

Sanders, J.W., and Honig, E.A. (1967). Brief tone audiometry: Results in normal and impaired ears. *Archives of Otolaryngology*, 85, 640-647.

Sanders, J.W., and Josey, A.F. (1970). Narrow-band noise audiometry for hard-to-test patients. *Journal of Speech and Hearing Research*, 13, 74-81.

Saunders, F.A. (1962). Violins old and new – an experimental study. *Sound*, 1, 7-15.

Scharf, B. (1983). Loudness adaptation. In: Tobias, J.V., and Schubert, E.D. (Eds.), *Hearing Research and Theory,* 2, 1-56. New York: Academic Press.

Scharine, A.A. (2002). *Auditory scene analysis: The role of positive correlation of dynamic changes in intensity and frequency*. Unpublished Dissertation. Tempe, AZ: Arizona State University.

Scharine, A.A., and Letowski, T. (2005). Factors affecting auditory localization and situational awareness in the urban battlefield. ARL Technical Report ARL-TR-3474. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Scharf, B. (1978). Loudness. In: Carterette, E.C., and Friedman, M.P. (Eds.), *Handbook of Perceptio*n, 4, 187-242. New York: Academic Press

Scharf, B., and Fishken, D. (1970). Binaural summation of loudness: Reconsidered. *Journal of Experimental Psychology,* 86, 374-379.

Schlauch, R.S., DiGiovanni, J.J., and Ries, D.T. (1998). Basilar membrane nonlinearity and loudness. *Journal of the Acoustical Society of America*, 103, 2010-2020.

Schlauch, R.S., Ries, D.T., and DiGiovanni, J.J. (2001). Duration discrimination and subjective duration for ramped and damped sounds. *Journal of the Acoustical Society of America,* 109, 2880-2887.

Schoeny, Z.G., and Carhart, R. (1971) Effects of unilateral Ménières disease on masking-level difference. *Journal of the Acoustical Society of America,* 50, 1143-1150.

Schorer, E. (1986). Critical modulation frequency based on detection of AM versus FM tones. *Journal of the Acoustical Society of America,* 79, 1054-1057.

Schouten, J.F. (1940). The residue and the mechanism of hearing. *Proceedings of the Koninklije Nederlandse Akademie van Wetenschappen,* 43, 991-999.

Schouten, J.R., Ritsma, R.J., and Cardozo, B.L. (1961). Pitch of the residue. *Journal of the Acoustical Society of America,* 34, 1418-1424,

Schubert, E.D. (1978). History of research on hearing. In: Carterette, E., and Friedman, M. (Eds.), *Handbook of Perception,* 4, 40-80. New York: Academic Press.

Schultz, M.C., and Schubert, E.D. (1969). A multiple choice discrimination test. *The Laryngoscope*, 79, 382.

Seashore, C.E. (1899). Some psychological statistics. *University of Iowa Studies in Psychology*, 2, 1-84.

Sekey, A. (1963). Short-term auditory frequency discrimination. *Journal of the Acoustical Society of America,* 35, 682-690.

Sekuler, R., and Blake, R. (1994). *Perception* (3rd Ed.). New York: McGraw Hill.

Sergeant, L., Atkinson, J.E., and Lacroix, P.G. (1979). *The NSMRL tri-word test of intelligibility*. Springfield, VA: National Technical Information Service.

Shafiro, V., and Gygi, B. (2004). How to select stimuli for environmental sound research and where to find them. *Behavioral Research Methods, Instrumentation, and Computers,* 36, 590-598.

Shailer, M.J., and Moore, B.C.J. (1983). Gap detection as a function of frequency, bandwidth, and level. *Journal of the Acoustical Society of America*, 74, 567-473.

Shanefield, D. (1980). The great ego crunchers: Equalized, double-blind tests. *High Fidelity,* 45, 57-61.

Shannon, R.V. (1976). Two-tone unmasking and suppression in a forward-masking situation. *Journal of the Acoustical Society of America*, 59, 1460-1470.

Shannon, R.V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hearing Research*, 11, 157-189.

Shofner, W.P., and Selas, G. (2002). Pitch strength and Stevens' power law. *Perception and Psychophysics,* 64, 437-450.

Shower, E.G., and Biddulph, R. (1931). Differential pitch sensitivity. *Journal of the Acoustical Society of America*, 3, 275-287.

Siegenthaler, B.M., and Haspiel, G.S. (1966). *Development of two standardized measures of hearing for speech by children*. University Park, PA: Pennsylvania State University.

Silverman, S.R., and Hirsh, I.J. (1955). Problems related to the use of speech in clinical audiometry. *The Annals of Otology, Rhinology, and Laryngology*, 64, 1234-1244.

Slot, G. (1954). *Audio Quality*. Eindhoven, Holland: Philips.

Small, A.M., Bacon, W.E., and Fozard, J.L. (1959). Intensive differential threshold for octave=band noise. *Journal of the Acoustical Society of America*, 31, 508-520.

Smith, F.O. (1914). The effect of training on pitch discrimination. *Psychological Monographs*, 16(69), 67-103.

Smith, J.O., and Abel, J.S. (1999). Bark and ERB bilinear transforms. *IEEE Transactions on Speech and Audio Processing*, 7, 697-708.

Smoorenburg, G.F. (1970). Pitch perception of two-frequency stimuli. *Journal of the Acoustical Society of America*, 48, 924-942.

Solomon, L.N. (1958). Semantic approach to the perception of complex sounds. *Journal of the Acoustical Society of America*, 30, 421-425.

Somerville, T., and Gilford, C.L.S. (1959). Acoustics of large orchestral studios and concert halls. *Journal of the Audio Engineering Society*, 7, 160-172.

Sone, T., Suzuki, Y., Ozawa, K., and Asano, F. (1994).Information of loudness in aural communication. *Interdisciplinary Information Sciences*, 1, 51-66.

Speaks, C., and Jerger, J. (1965). Method for measurement of speech identification. *Journal of the Acoustical Society of America*, 37, 1205.

Spiegel, M.F. (1979). The range of spectral integration. *Journal of the Acoustical Society of America*, 66, 1356-1363.

Staffeldt, H. (1974). Correlation between subjective and objective data for quality loudspeakers. *Journal of the Audio Engineering Society*, 22, 402-415

Stapells, D.R., Picton, T.W., and Smith, A.D. (1982). Normal hearing thresholds for clicks. *The Journal of the Acoustical Society of America*, 72, 74-79.

Steeneken, H.J.M. (1992). On measuring and predicting speech intelligibility. Unpublished Dissertation, University of Amsterdam. Amsterdam (Holland): Institute of Phonetic Sciences,

Steeneken, H.J.M., and Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *Journal of the Acoustical Society of America*, 67, 318-326.

Steeneken, H.J.M., and Houtgast, T. (1999). Mutual dependence of octave-band weights in predicting speech intelligibility. *Speech Communication*, 28, 109-123.

Steinberg, J.C., and Snow, W.B. (1934). Auditory Perspective - Physical Factors. Proceedings of the Symposium on Wire Transmission of Symphonic Music and its Reproduction in Auditory Perspective. *Electrical Engineering*, 1, 12-17.

Steinke, G. (1958). Subjektive Bewertung der Übertragungsqualität. *Technische Mitteilungen aus dem Betriebslaboratorium für Rundfunk und Fernsehen*, 13, 27-30.

Stelmachowicz, P.G., Beauchaine, K.A., Kalberer, A., and Jesteadt, W. (1989). Normative thresholds in the 8- to 20-kHz range as a function of age. *Journal of the Acoustical Society of America,* 86, 1384-1391.

Stevens, S.S. (1934a). The volume and intensity of tones. *American Journal of Psychology*, 46, 397-408.

Stevens, S.S. (1934b). Tonal density. *Journal of Experimental Psychology*, 17, 585-592.

Stevens, S.S. (1935). The relation of pitch to intensity. *Journal of the Acoustical Society of America*, 6, 150-154.

Stevens, S.S. (1955). The measurement of loudness. *Journal of the Acoustical Society of America,* 27, 815-829.

Stevens, S.S. (1956). Calculation of the loudness of complex noise. *Journal of the Acoustical Society of America*, 28, 807-832.

Stevens, S.S. (1957). On the psychophysical law. *Psychological Review*, 64, 153-181.

Stevens, S.S. (1972). Perceived level of noise by Mark VII and decibels. *Journal of the Acoustical Society of America*, 51, 575-601.

Stevens, S.S., and Davis H. (1938). *Hearing.* New York: Wiley and Sons.

Stevens, S.S., and Newman, E.B. (1936). The localization of actual sources of sound, *American Journal of Psychology,* 48, 297-306.

Stevens, S.S., and Volkmann, J. (1940). The relation of pitch to frequency: A revised scale. *The American Journal of Psychology,* 8, 329-353.

Stevens, S.S., Volkmann, J., and Newman, E.B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America,* 8, 185-191.

Studebaker, G.A., and Sherbecoe, R.L. (1988). Magnitude estimations of the intelligibility and quality of speech in noise. *Ear and Hearing*, 9, 259-267.

Stumpf, C. (1911). Konsonanz und Konkordanz. *Beitrage zur Akustik und Musikwissenschaft*, 6, 116-150.

Suetomi, D., and Nakajama, Y. (1998). How stable is time-shrinking? *Journal of Music Perception and Cognition*, 4, 19-25.

Summers, W.V., Pisoni, D.B., Bernacki, R.H., Pedlow, R.I., and Stokes, M.A. (1988). Effects of noise on speech production: acoustic and perceptual analyses. *The Journal of the Acoustical Society of America*, 84, 917-928.

Suzuki, Y., Sone, T., and Kanasashi, K. (1982). The optimum level of music listened in the presence of noise. *Journal of the Acoustical Society of Japan (E),* 3, 55-65.

Szlifirski, K., and Letowski, T. (1981). System podstawowych kryteriów oceny nagrań dźwiękowych [A system of criteria for the quality evaluation of sound recordings]. *Przegląd Techniki Radia i TV,* 1, 36-42.

Szymaszek, A., Szelag, E., and Sliwowska, M. (2006). Auditory perception of temporal order in humans: The effect of age, gender, listener practice and stimulus presentation mode. *Neuroscience Letters*, 403, 190-194.

Tang, Q., Liu, S., and Zeng, F-G. (2006). Loudness adaptation in acoustic and electric hearing. *Journal of the Association for Research in Otolaryngology,* 7, 59-70.

Tang H., and Letowski T. (2007). High-frequency hearing threshold measurements using insert earphones and warble signals. *Proceedings of the 2007 Industrial Engineering Research Conference.* Nashville, TN: Institute of Industrial Engineers (IIE).

Terhardt, E. (1974). On the perception of periodic sound fluctuations (roughness). *Acustica,* 30, 201-213.

Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research,* 1, 155-182.

Terhardt, E., Stroll, G., and Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 71, 679-688.

Thomas, G.J. (1949). Equal-volume judgment of tones. *American Journal of Psychology,* 62, 182-201.

Thwing, E.J. (1955). Spread of perstimulatory fatigue of the pure tone to neighboring frequencies. *Journal of the Acoustical Society of America,* 27, 741-748.

Thurlow, W.R., Mangels, J.W., and Runge, P.S. (1967). Head movements during sound localization, *Journal of the Acoustical Society of America,* 42, 489-493.

Tillman, T.W., Carhart, R.T., and Wilber, L.A. (1963). *A test for speech discrimination composed of CNC monosyllabic words (N.U. Auditory Test No. 4)*. Brooks Air Force Base, TX: USAF School of Aerospace Medicine, Aerospace Medical Division (AFSC).

Toole, F.E. (1982), Listening tests- turning opinion into fact. *Journal of the Acoustical Society of America*, 30, 431-445.

Tran T., Letowski T., and Abouchacra K. (2000). Evaluation of acoustic beacon characteristics for navigation tasks. *Ergonomics*, 43(6), 807-827.

Tran, P., Letowski, T., and McBride, M. (2008). Bone conduction microphone: Head sensitivity mapping for speech intelligibility and sound quality. *Proceedings of the 2008 International Conference on Audio, Language and Image Processing (ICALIP 2008)*, 107-111. Shanghai (China): Institute of Electrical and Electronic Engineers (IEEE).

Truman, S.R., and Wever, E.G. (1928). The judgment of pitch as a function of the series. *University of California Publications in Psychology,* 3, 215-223.

Turnbull, W.W. (1944). Pitch discrimination as a function of tonal duration. *Journal of Experimental Psychology*, 34, 302-316.

Uchanski, R.M., Choi, S.S., Braida, L.D., Reed, C.M., and Durlach, N.I. (1996). Speaking clearly for the hard of hearing .4. Further studies of the role of speaking rate. *Journal of Speech and Hearing Research,* 39(3), 494-509.

van den Brink, G. (1974). Monotic and dichotic pitch matchings with complex sounds. In: Zwicker, E., and Terhardt, E., (Eds.), *Models in Hearing*. Berlin, Germany: Springer-Verlag.

van Wijngaarden, S.J. and Drullman, R. (2008). Binaural intelligibility prediction based on the speech transmission index. *Journal of the Acoustical Society of America,* 123, 4514-4523.

Viemeister, N. (1979). Temporal modulation transfer function based upon modulation thresholds. *Journal of the Acoustical Society of America,* 66, 1364-1380.

Voiers, W.D. (1983). Evaluating processed speech using the Diagnostic Rhyme Test. *Speech Technology,* 1, 338-352.

Voiers, W.D., Sharpley, A. and Hehmsoth, C. (1975). Research on diagnostic evaluation of speech intelligibility. Bedford, MA: U.S. Air Force Cambridge Research Laboratories. Research Report AFCRL-72-0694.

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology,* 27, 339-368.

Wallach, H., Newman, E.B., and Rosenzweig, M.R. (1949). Precedence effect in sound localization. *American Journal of Psychology,* 62, 315-336.

Walliser, K. (1968). Zusammenwirkung von Hüllkurven-periode und Toneheit bei der Bildung der Periodentonhöhe. Doctoral dissertation. München (Germany): München Technische Hochschule.

Walliser, K. (1969). Über die Abhängigkeiten der Tonhöhenempfindung von Sinustönen vom Schallpegel von überlagertem drosselndem Störschall und von der Darbietungsdauer. *Acustica*, 21, 211-221.

Ward, W.D. (1954). Subjective musical pitch. *Journal of the Acoustical Society of America,* 26, 369-380.

Ward, W.D. (1963). Diplacusis and auditory theory. *Journal of the Acoustical Society of America,* 35, 1746-1747.

Warfield, D. (1973). The study of hearing in animals. In: Gay, W. (Ed.), *Methods of Animal Experimentation, IV*. London: Academic Press.

Warren, R.M., and Wrightson, J.M. (1981). Stimuli producing conflicting temporal and spectral cues to frequency. *Journal of the Acoustical Society of America,* 70, 1020-1024.

Watson, B.U. (1991). Some relationship between intelligence and auditory discrimination. *Journal of Speech and Hearing Research*, 34, 621-627.

Watson, C.S., Kelly, W.J., and Wronton, H.W. (1976). Factors in the discrimination of tonal patterns, II: Selective attention and learning under various levels of stimulus uncertainty. *Journal of the Acoustical Society of America,* 60, 1176-1186.

Weber, E.H. (1834). *De Pulsu, Resorpitone, Auditu et Tactu: Annotationes Anatomicae et Physiologica*e. Leipzig, Germany: Koehlor.

Webster, F.A. (1951). The influence of interaural phase on masked thresholds, I: The role of the interaural time-deviation. *Journal of the Acoustical Society of America,* 23, 452-462.

Wegel, R.L., and Lane, C.E. (1924). The auditory masking of one sound by another and its probable relation to the dynamics of the inner ear. *Physical Review*, 23, 266-285.

Weinstein, B. (2000). *Geriatric Audiology*. New York: Thieme.

Whilby, S., Florentine, M., Wagner, E., and Marozeau, J. (2006). Monaural and binaural loudness of 5- and 200-ms tones in normal and impaired hearing. *Journal of the Acoustical Society of America*, 119, 3931-3939.

White, L.J., and Plack, C.J. (2003). Factors affecting the duration effect in pitch perception for unresolved complex tones. *Journal of the Acoustical Society of America*, 114, 3309-3316.

Wier, C.C., Jesteadt, W., and Green, D.M. (1977). Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America*, 61, 178-184.

Wightman, F.L., and Green, D.M. (1974). The perception of pitch. *American Scientist*, 62, 208-215.

Wightman, F.L., and Kistler, D.J. (1992). The dominant role of low-frequency interaural time differences in sound localization, *Journal of the Acoustical Society of America*, 91, 1648-1661.

Willott, J. (1991). *Aging and the Auditory System*. San Diego: Singular Publishing Group.

Wilson, R.H., and Antablin, J. K. (1980). A picture identification task as an estimate of the word-recognition performance of nonverbal adults. *Journal of Speech and Hearing Disorders*, 45, 223-238.

Woodrow, H. (1951). Time perception. In: Stevens, S.S. (Ed.) *Handbook of Experimental Psychology*, Chapter 32. New York: Wiley and Sons.

Yamashita, T., Ishii, Y., Nakamura, M., and Kitamura, O. (1990). A method for evaluating interior sound quality of automobiles. Paper presented at the 119th Acoustical Society of America (ASA) Conference, paper PP7. State College, PA: Acoustical Society of America (ASA).

Zarcoff, M. (1958). Correlation of filtered, band-limited white noise and the speech reception threshold. *Archives of Otolaryngology - Head and Neck Surgery*, 63, 372-381.

Zeng, F-G., and Shannon, R.V. (1994). Loudness-coding mechanisms inferred from electric simulation of the human auditory system. *Science*, 264, 564-566.

Zera J., Boehm T., and Letowski T. (1982). Application of an automatic audiometer in the measurement of the directional characteristic of hearing. *Archives of Acoustics* 7, 197-205.

Zwicker, E. (1952). Die Grenzen der Hörbarkeit der Amplituden-modulation und Frequenz-modulation eines Tones. *Acustica*, 3, 125-133.

Zwicker, E. (1960). Ein Verfahren zur Berechnung der Lautstärke. *Acustica*, 10, 304-308.

Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *Journal of the Acoustical Society of America*, 33, 248-248.

Zwicker, E., and Fastl, H. (1999). *Psycho-acoustics: Facts and Models* (2nd Ed.). Berlin: Springer.

Zwicker, E., and Feldtkeller, R. (1955). Über die Lautstärke von gleichformigen Geräuschen. *Acoustica* 5, 303-316.

Zwicker, E., and Feldtkeller, J. (1967). *Das Ohr als Nachrichtenempfanger* (2nd Ed.). Stuttgart (Germany): Hirzel Verlag.

Zwicker, E., Flottrop, G., and Stevens, S.S. (1957). Critical band width in loudness summation. *Journal of the Acoustical Society of America*, 29, 548-557.

Zwicker, E., and Scharf, B. (1965). A model of loudness summation. *Psychological Review*, 72, 3-26.

Zwicker, E., and Terhardt, E. (1980). Analytical expressions for critical-band rate and critical bandwidth as a function of frequency, *Journal of the Acoustical Society of America*, 68, 1523-1525.

Zwicker, E., and Zwicker, U.T. (1991). Dependence of binaural loudness summation on interaural level differences, spectral distribution, and temporal distribution. *Journal of the Acoustical Society of America*, 89, 756-764.

Zwislocki, J. (1960). Theory of temporal auditory summation. *Journal of the Acoustical Society of America*, 32, 1046-1060.

Zwislocki, J.J. (1965). Analysis of some auditory characteristics. In: Luce, R.D., Bush, B.R., and Galanter, E. (Eds.), *Handbook of Mathematical Psychology*, III. New York: Wiley and Sons.

12 VISUAL PERCEPTUAL CONFLICTS AND ILLUSIONS

Leonard A. Temme
Melvyn E. Kalich
Ian P. Curry
Alan R. Pinkus
H. Lee Task
Clarence E. Rash

Vision is arguably the most important of the human senses for a Warfighter. The purpose of visual processing is to take in information about the world around us and make sense of it (Smith and Kosslyn, 2007); vision involves the sensing and the interpretation of light. The visual sense organs are the eyes, which convert incoming light energy into electrical signals (see Chapter 6, *Basic Anatomy and Structure of the Human Eye*). However, this transformation is not *vision* in its entirety. Vision also involves the interpretation of the visual stimuli and the processes of *perception* and ultimately *cognition* (see Chapters 10, *Visual Perception and Cognitive Performance*, and 15, *Cognitive Factors*).

The visual system has evolved to acquire veridical information from natural scenes. It succeeds very well for most tasks. However, the information in visible light sources is often ambiguous, and to correctly interpret the properties of many scenes, the visual system must make additional assumptions about the scene and the sources of light. A side effect of these assumptions is that our visual perception cannot always be trusted; visually-perceived imagery can be deceptive or misleading, especially when a scene is quite different from those that pushed the evolution of the visual system in the past. As a result, there are situations where "seeing is not believing," i.e., what is perceived is not necessarily real. These misperceptions are often referred to as illusions. Gregory (1997) identifies two classes of illusions: those with a physical cause and those due to the misapplication of knowledge.

Physical illusions are those due to the disturbance of light between objects and the eyes, or due to the disturbance of sensory signals of eye (also known as physiological illusions). Cognitive illusions are due to misapplied knowledge employed by the brain to interpret or read sensory signals. For cognitive illusions, it is useful to distinguish specific knowledge of objects from general knowledge embodied as rules (Gregory, 1997).

An important characteristic of all illusions is that there must be some means for demonstrating that the perceptual system is somehow making a mistake. Usually this implies that some aspect of the scene can be measured in a way that is distinct from visual perception (e.g., can be measured by a photometer, a spectrometer, a ruler, etc.). It is important to recognize that these "mistakes" may actually be useful features of the visual system in other contexts because the same mechanisms underlying an illusion may give rise to a veridical percept for other situations. An illusion is only an illusion if the "mistakes" are detectable by other means.

While illusions may deceive the Warfigher, there are other limits of the visual system that can lead to mistakes during the conduct of a mission. These include visual masking (the reduction or elimination of the visibility of one brief stimulus, called the "target", by the presentation of a second brief stimulus, called the "mask"), binocular rivalry (an unintentional alternation between different images presented to each eye), and spatial disorientation (a condition in which a Warfighter's perception of position and motion does not agree with reality).

Visual Masking

Visual masking usually refers to the influence of one visual stimulus on the appearance of another visual stimulus, with one or the other or both stimuli being transient. Since, as this discussion will make clear, visual masking occurs all the time in the real world, it certainly plays a key role in the use of visual displays and can therefore be expected to affect the use of heads-up (HUDs) and head-/helmet-mounted displays (HMDs). The following discussion uses two classic experiments to describe the general features of visual masking. Following the discussion of these experiments, some implications of visual masking are generalized to new and evolving display technologies.

In the visual masking literature the visual stimulus causing masking is typically referred to as the *masking stimulus* (MS) or by some other term that emphasizes its masking properties. Similarly, the visual stimulus whose appearance the MS alters is typically referred to as the *test stimulus* (TS) or by some other term that similarly emphasizes its susceptibility to the effects of the MS. Most commonly, the MS and TS are flashed on and off with some defined temporal relation between them. If the MS and/or TS do not vary over time but have essentially unlimited exposure durations, the dependence of the TS appearance on the MS is usually considered to be due to a class of visual mechanisms that are different from masking such as simultaneous contrast. Moreover, the MS and the TS invariably have defined spatial characteristics; that is, they are not uniform illuminations of the visual field.

In masking experiments, changes in the appearance of the TS can be used to assess the effects that the MS has on the visual system. To this extent, visual masking experimental methodologies are often indirect; that is, although the experiment records a defined visual characteristic of the TS, what is really of interest is the visual effect of the MS. For example, some studies refer to the MS as a conditioning stimulus or conditioning flash and treat the TS as little more than a probe of the affects of the MS on the sensitivity of the visual system. For such studies a common measure is the minimum TS luminance required for its detection. Such studies usually make the implicit assumption that TS threshold luminance at the retinal location reflects that location's sensitivity[1].

Masking by light – Crawford masking

The first of the two experiments to be discussed is the classic study published in 1947 by Crawford. The MS in this study was a homogeneous, circular 12° diameter light flashed on for 0.524 second every 7.2 seconds. The TS was a circular spot of light with a diameter of 0.5° that was flashed for 0.01 seconds and that was spatially centered in the MS. The task was to measure the minimum amount of light needed in the TS to detect it. These threshold measurements for one subject are in Figure 12-1, which shows TS threshold brightness as a function of time relative to the onset of the MS. The dotted vertical lines at 0 and at 0.524 seconds mark the MS onset and offset, respectively. Positive numbers indicate time after the onset of the MS while negative numbers indicate time before the onset of MS. The three different functions plotted show the results for three different levels of MS brightness. While the three functions look similar, the greater the MS brightness, the more clearly defined is the function.

The results show that TS threshold is a complex function of time. Despite the fact that MS brightness is constant over its duration, the sensitivity of the visual system, as reflected by the threshold of the TS in the center of the homogeneous MS, changes over time. Consider the brightest MS: The most obvious characteristic of this function is the peak of the TS threshold near MS onset, and the rise of the TS threshold near MS offset. Between these two peaks, the TS threshold changes over the MS exposure; falling rapidly over the first 100 ms or so of the MS exposure, then more slowly until it starts to rise again, approximately 50 ms before the MS offset. After MS offset, TS threshold falls, first relatively rapidly then more gradually out past the measured time window.

The most surprising aspect of the data is that TS threshold begins to increase approximately 100 ms before the onset of the MS. From one perspective, these results are very surprising since they clearly show that the MS affects visual sensitivity apparently before MS onset. This effect – that the MS operates backward in time – usually is called backward masking and has been well replicated under many conditions. It should not be over-

[1] Such methodologies typically make several implicit assumptions. An example of another is that the threshold is determined by the most sensitive visual process operative at that retinal location.

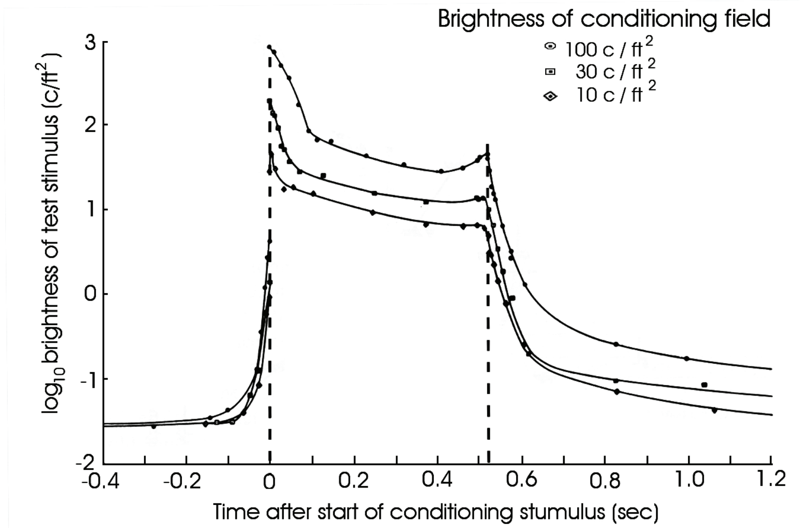


Figure 12-1. Typical Crawford-type masking data (Crawford, 1947). The x-axis is time in seconds, where 0 time is the onset of the conditioning stimulus. The y-axis is the threshold brightness of the test flash. Each of the three different functions shows data for conditioning stimulus of a different brightness. The dotted vertical lines show the onset and offset of the conditioning stimulus.

looked that a similar backward effect on visual sensitivity occurs just before MS offset, but to a lesser degree. Crawford suggested: “There seem to be two possible explanations. Either the relatively strong conditioning stimulus overtakes the weaker test stimulus on its way from retina to brain and interferes with its transmission; or the process of perception of the test stimulus, including the receptive processes in the brain, takes an appreciable time of the order of 0.1 sec., so that the impression of a second (large) stimulus within this time interferes with perception of the first stimulus.” During the more than 60 years since this report, the theoretical bases of such backward masking effects have been elaborated in great detail. But to be fair, similar masking phenomena had been well studied through the early nineteenth century with a very sophisticated understanding of what they imply about visual neural function.

Metacontrast and paracontrast

About six years after Crawford’s classic paper, Alpern (1953) reported a completely different kind of masking that used the visual stimulus arrangement in Figure 12-2. The stimuli were vertically-oriented rectangular bars of light that were each 2.5° by 0.5°. The TS was the bar in the middle. The MS was the pair of flanking bars. These three bars were presented to the right eye of the test subjects. The bar identified as the *standard stimulus* (SS) above the TS was presented to the left eye. Since the components of the stimulus array were distributed between the two eyes in this experiment, it was critical for the study to keep the eyes properly aligned. The little stimulus light, ‘**z**’, which was midway between the SS and the TS, was illuminated all the time and served as the fixation stimulus to help the subject minimize voluntary eye movements. Also, since **z** was seen by both eyes it helped the two eyes stay properly aligned.

The four rectangular bars that formed the stimulus pattern of interest were presented as flashes of 5 ms duration. The variable was the time interval between the flash presentation of the TS and SS pair and the MS. The issue is the effect of the MS on the apparent brightness of the TS as a function of the temporal interval between the TS and the MS. The apparent brightness of the TS is measured by setting TS brightness to be equal to that of the SS, which, as mentioned, was presented in the other eye, simultaneously with the TS.[2] The results are summarized in Figure 12-3.

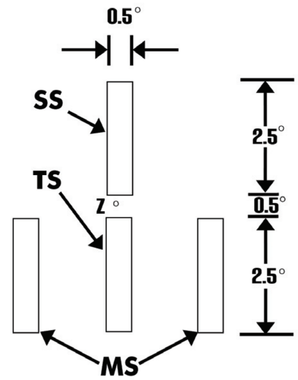


Figure 12-2. The stimulus configuration used in the metacontrast study by Alpern (1953). The SS (standard stimulus) was presented to the left eye, the TS and MS were presented to the right eye. The MS consisted of the pair of rectangular lights that flanked the TS. Stimulus **z** was presented to both eyes to help the two eyes to stay aligned by providing a fixation target. Stimulus **z** was on all the time while all the other stimuli were flashed for 5 ms. The SS and TS were presented at the same time. The task was to set the luminance of the TS to match the SS. The experiment investigated the affect of the interval between the MS and the flash presentation of the TS and SS.

The ordinate scale is the brightness of the TS and the abscissa is the time interval between the onset of the TS and the MS; i.e., the flash onset asynchrony. Note that higher values on the ordinate reflect greater masking effects since the TS must be made brighter to match the constant brightness of the SS. It is important to note that the convention for plotting time on the abscissa in Figure 12-3 is the reverse of the convention used by Crawford, Figure 12-1. In Figure 12-3, positive time indicates that the TS occurred before the MS. Hence, the major effects that Alpern is showing are all backward in time.

This figure contains eight different curves showing the effect of increasing MS brightness on visual masking. With the brightness of the MS set low (either 0.1 or 3.6 foot-Lamberts (ft-L)), the subject set the TS to about 11 foot-Lamberts (ft-L) for all temporal intervals between the MS and the TS. In other words, the TS and SS seemed about equally bright regardless of the flash onset asynchrony. The fact that the data are all about 11 ft-L for all time intervals indicates negligible visual masking when the MS are dim. On the other hand, consider the top curve which shows the TS brightness matches to the 11 ft-L SS when the MS was 3,000 ft-L. When the TS preceded the MS by about 125 ms, the TS had to be set to about 400 ft-L to match the SS of 11 ft-L. This indicates substantial masking since the TS had to be made nearly 40 times brighter to match the constant SS. Furthermore, for these stimuli, masking effects are convincingly evident out to almost 300 ms, which is 60 times longer than the duration of the MS itself. In other words, the brightness of the TS is being modulated by a MS that occurs almost 300 ms later in time and that falls on a completely different part of the retina.

Admittedly, there is a great difference in brightness between the MS and the SS (3000 ft-L vs. 11 ft-L), but Figure 12-3 also plots the masking function when the MS and SS are equally bright, 11 ft-L. For those stimuli, the TS has be to about 100 ft-L to match the SS of 11 ft-L. Under these conditions the masking function peaks about 75 ms, and lasts to something between 150 to 175 ms.

[2] The assumption is that masking between the two eyes is minimal with the presupposition that much of the masking phenomena are retinal. In fact, however, other research has shown that substantial masking does occur between the two eyes.

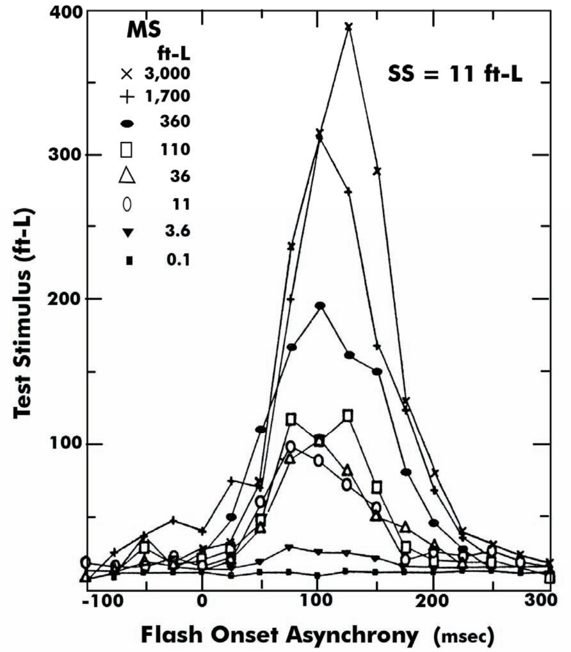


Figure 12-3. The brightness match between the TS and the SS as a function of the temporal interval between the TS and the MS. Note positive time intervals indicate that the TS occurred before the MS whereas negative time intervals indicate that the MS occurred before the TS.

These data illustrate a visual backward masking effect most widely known in the vision science community as *metacontrast*. The typical earmarks of metacontrast are that the MS and TS stimulate non-overlapping parts of the retina, that the peak masking occurs when the MS follows the TS by about 100 ms or so, and that the time course of the masking function is greatly influenced by the specifics of the experiment. Alpern's data in Figure 12-3 also illustrate this last point. The peak of the masking function tends toward larger intervals as the luminance of the MS increases.

Figure 12-3 illustrates another important type of masking, often called *paracontrast*, which in some respects is more akin to the Crawford-type masking. When the MS precedes the TS, the apparent brightness of the TS is also reduced to some extent. This can be seen for the data plotted over the interval from 0 to -100 ms. The magnitude of the paracontrast is a fraction of that of metacontrast and would certainly not be convincingly shown by Alpern's data. But because paracontrast has been so convincingly demonstrated in other research, it is reassuring to see evidence of it in these data.

The data in Figure 12-3 illustrate that the shape of the metacontrast and paracontrast masking functions depend on the luminance of the MS. The literature generalizes this observation but the magnitude and time course of masking depends on more just the MS luminance; they depend on the specifics of the stimulus parameters as well as the response used to measure masking. In general, para- and metacontrast functions tend to be roughly either monotonic (smoothly rising or falling as a function of MS luminance) or U-shaped. Monotonic functions are sometimes called Type A while the U-shaped functions are called Type B (Kahneman, 1968). This means that the plot of TS visibility as a function of the interval between the TS and MS can take on a number of shapes. The metacontrast data in Figure 12-3 clearly show a typical U-shape or Type B function. The small size and variability of the paracontrast data obscure the shape of the functions but there is some indication for both Type A and B functions.

One of the major issues masking research continues to address is the clarification and understanding of the processes that give rise to the one or other type of function. In general, studies that require discrimination or detection of features of the TS show different time functions than studies that record changes in TS appearance, brightness, or contrast, which, in turn, are different from functions determined simply by responses to whether or not the TS occurred. On one hand the plethora of different functions might seem to indicate uncontrolled variables, noise, or other experimental or methodological problems. On the other hand, since the results are orderly, researchers in general consider the spectrum of masking functions that different response criteria produce to be indicative of the kinds of information processing the neural visual system performs. There are now several highly quantitative models of para- and metacontrast published elaborating on known neuroanatomy and physiology (Bachmann, 1994; Breitmeyer, 1984; Breitmeyer and Ogmen, 2006).

Pattern masking

A third, important class of masking studies should be mentioned. These use a MS that incorporates some sort of spatial pattern. The MS structure may be a random noise pattern, an alphanumeric array, or an array of bars or gratings, or some other non-homogeneous spatial distribution appropriate for the purposes of the study. Underlying the use of such structured masking is the notion that the TS contains information and after the TS is turned off, the visual system continues to process the TS information. For example, the visual system can be expected to process a 5-ms long TS of the letter 'D' for longer than the 5-ms duration of the TS. Pattern masking procedures are considered to be a way of blanking or controlling the continued neural trace of the brief TS, following the idea laid out by Crawford's comment quoted earlier (pg 285): "… the process of perception of the test stimulus, including the receptive processes in the brain, takes an appreciable time of the order of 0.1 second ..., so that the impression of a second (large) stimulus within this time interferes with perception of the first stimulus." Based on such pattern masking research, it is now commonly recognized that a visual stimulus produces a neural trace and that this neural trace is available and recognized after the external stimulus has been turned off, as though the trace serves as an input buffer. This visual phenomenon is often referred to as an *iconic memory*.[3]

Masking – A final word

Visual masking may seem to be a rather esoteric concern of vision neuropsychophysiology yet Bachmann surveyed 15 years of the "most authoritative, most cited psychology journals publishing on general problems of psychology, psychophysiology, information processing, and perception. ... among all the articles published within this period masking as a scientific topic was studied almost in 3% of the articles and masking as the method helping to study some scientific problem was employed in 11% of the articles (pg 11)." Clearly visual masking continues to be an important active area of research. Backward masking is intrinsically interesting because it describes how information is conducted through the nervous system. The two book-length reviews of visual masking by Breitmeyer and the book by Bachmann are excellent introductions to the rich literature of this area of research.

To put visual masking into perceptive for its implications in the real world, cockpits, simulations, virtual reality displays, HMDs, and so forth it is helpful to remember that the time course of masking is short, only in the order of hundreds of mss. But the shortness of the masking effects does not make masking unimportant. Instead, masking seems to be fundamental to the way our visual system works. Every time the eye sees something, that something is masking what the eye had just seen the previous instant. The time domain of visual masking is the time domain of eye movements. Every eye movement involves visual masking.

Current display technology enables a form of masking that is essentially new and for which there is little if any research. This masking has to do with the display of information on a transparency or a surface that appears to be

---

[3] *Iconic memory* is a type of very short-term visual (or sensory) memory. An analogous memory for sound is called *echoic memory*, which can be defined as very brief sensory memory of some auditory stimuli.

transparent. An example would be a HUD or HMD. On one hand, the concern historically has been whether the symbology obscures the view of the world behind it. The deeper question here is the extent to which the world behind obscures the symbology. After all, the view of the world, which is visually intricate and complex and full of all the visual information the world contains, is the background on which the symbology is presented. Any motion of the HUD or HMD relative to the world creates transients in the background relative to the superimposed, less transient symbology. The relevant question is the extent to which the visibility of the superimposed symbology is affected by the transients in the background. In this situation, the background is the MS and the foreground is the TS. The issue is more than just the summation of luminance and a reduction of contrast; the issue is the effects on visibility of the continuous presence of transients in the background MS (Harding and Rash, 2004).

The same situation is increasingly common with computer displays. Web pages now display information on textured backgrounds. The information (TS) or the background (MS) may be stationary, may move, or even flash. The TS may be any gradation between an opaque overlay to a transparent one. These display technologies create environments that our visual system has not previously encountered. Since this form of masking derives from new technology, the human visual system may not have evolved biological mechanisms to process these masking effects. Our normal masking functions appear to have evolved confronting opaque surfaces rather than transparencies. It is possible that visual masking mechanisms that underlie our information processing may work in exactly the wrong way for handling (e.g., filtering) the kinds of masking effects that these new technologies create, obscuring rather than enhancing information.

Binocular Rivalry

An HMD can present information to one eye (monocular HMD) or both eyes (biocular or binocular HMD). When using HMDs, it is very common to have dissimilar imagery presented to the two eyes. As a result, there can be a state of competition between the two image representations in the brain. This can result in one representation being suppressed while the other forms a conscious percept (Winterbottom et al., 2007). This selective processing can alternate over time, resulting in a condition referred to as *binocular rivalry*.

There are a number of possible bistable perceptual representations of the visual world, sometimes called bistable stimuli (i.e., having two distinct presentations) (Andrews et al., 2005; Howard, 2002 and 2005; Leopold et al., 2005; Wade, 2005). These include monocular bistable stimuli, such as transparent three-dimensional (3-D) objects, figure ground reversals, ambiguous figures, and images with dissimilar color and orientation (Figure 12-4), and biocular/binocular bistable stimuli, which can lead to binocular rivalry (Figure 12-5).

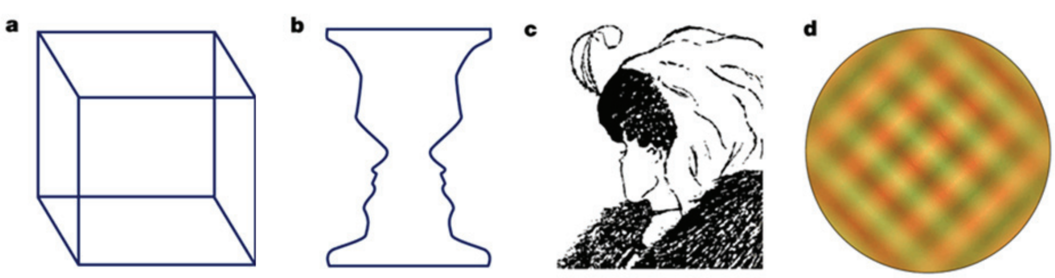


Figure 12-4. Monocular bistable stimuli: (a**)** Necker cube, (b) Rubin's vase versus face figure, (c) Boring's old lady versus young woman figure, and (d) Monocular rivalry, in which two physically superimposed patterns that are dissimilar in color and orientation compete for perceptual dominance (from Blake and Logothetis, 2002).

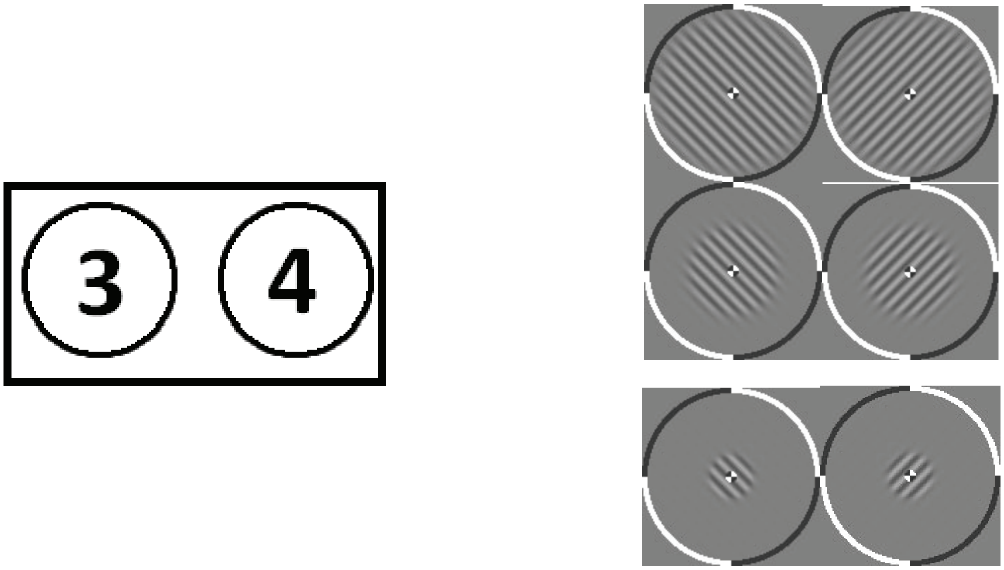


Figure 12-5. These figure pairs can be free-fused by crossing one's eyes and looking at a point between and just in front of the figures to be fused. The left and right figures in each pair will be brought to awareness alternately. Also apparent will be a combination or fragmented patchiness between the formations of stable single figures. The number pair is adapted from Blake (2001). The gratings are from Blake (2008).

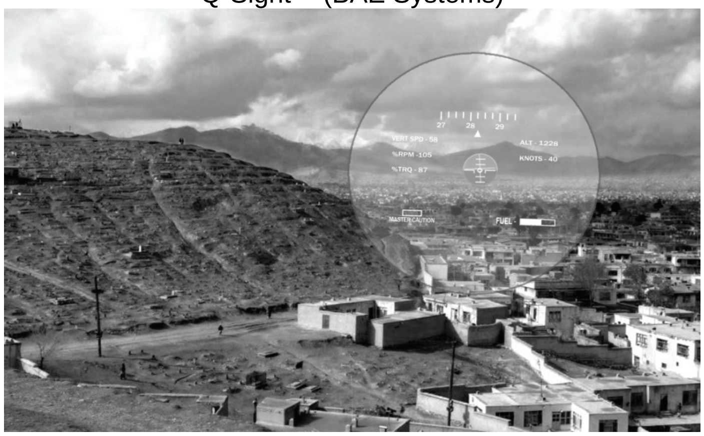
In order to see a single, fused object viewed using two eyes, it is not necessary that the two images be identical. In fact, we use image differences in various ways to enhance perception of the visual world (Cutting and Vishton, 1995; Wagner, 2006). A close object viewed from the different angles provided by two eyes allows objects to be viewed as 3-D, a resolution of the perspective differences. The perception of layout in both personal space and action space is facilitated by binocular disparity (Cutting and Vishton, 1995). A number of studies using both gratings and small letter contrast sensitivity show binocular enhancement of visual acuity (~10%) and contrast (~40%) (Blake and Levinson, 1977; Blake, Sloan and Fox, 1981; Cagenello, Arditti and Halpern, 1993; Campbell and Green, 1965; Rabin, 1995). There is also a significant increase in brightness of objects viewed binocularly (Crozier and Holway, 1938; Lythgoe and Phillips, 1938).

Although there is considerable latitude in our ability to reconcile images that are different in content or at different retinal locations in the two eyes and to capitalize on image differences, there are limits to the degree and kinds of differences that can be resolved into a stable, single, fused percept. The brain devotes significant processing power to avoid seeing double (diplopia).

How the human brain handles these image differences and, in particular, how this relates to HMDs is important to HMD designs and applications. Binocular rivalry is a major concern when using monocular HMDs, particularly when one eye is free to view the user's surrounds (Figure 12-6). However, binocular rivalry can also be a problem when using biocular or binocular HMDs, as when symbology presented to one eye overlays a view of the outside world (through a see-through HMD) seen with both eyes (Figure 12-7), as when symbology presented to one eye overlays an intensified image or forward-looking infrared (FLIR) image of the outside world presented to both eyes, when partially-overlapping images are used to expand the total HMD field-of-view (FOV) (Figures 12-8 and 12-9), or when images to the two eyes are misaligned.

There are two aspects to binocular rivalry. The first is *binocular*. Humans, like other primates and a number of mammalian predators, have eyes in the front. This setup is used to produce a 3-D perception of the world derived from the fusion of two 2-D images. This is important for manipulating near objects and representing action space (Cutting and Vishton, 1995). Herd animals and dolphins have eyes on the sides of their heads, providing a more

Q-Sight™ (BAE Systems)



Simulated view of a scene viewed while using the Q-Sight™ HMD. No suppression is depicted.

Figure 12-6. The monocular Q-Sight™ (top) is a HMD system developed and manufactured by BAE Systems. Pilots using this system at night with image intensification (I^2) or FLIR imagery will suppress an image from one eye, while attending to the image from the other. This ability, however, is not perfect and unexpected alternations do occur. When this type of see-through system is used during the daytime, without FLIR, a complex background with high contrast and high spatial frequencies that can be binocularly fused will tend to decrease rivalry with the symbology, although it will not necessarily eliminate it (Patterson et al., 2007).

global view (perspective) but depend largely on combining monocular cues to represent a 3-D world view. They are often very sensitive to motion and direction of objects. Much of the human brain is tied up in resolving image difference and local ambiguities from the two eyes to produce a remarkably robust, effortless visual representation of the world (Leopold et al., 2005). Andrews, Sengpiel and Blakemore (2005), in paraphrasing Hermann von Helmolz, pointed out that when constructing a perceptual representation of the visual world, the brain has to cope with the fact that any given 2-D retinal image could be the projection of countless object configurations in the 3-D world.

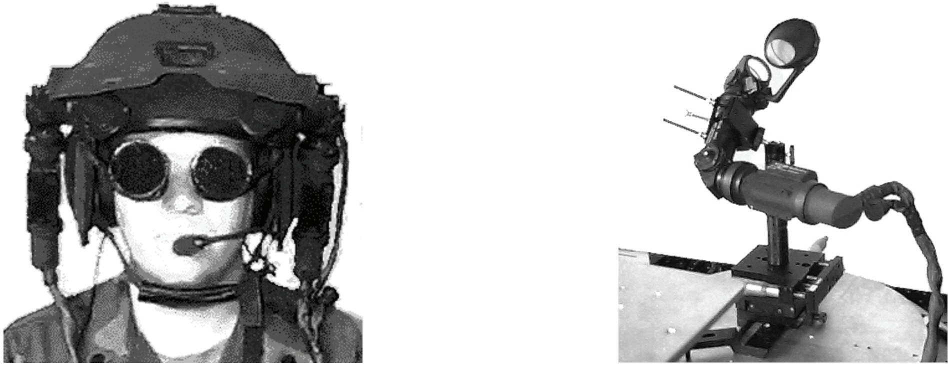
Second, rivalry occurs when the brain cannot resolve the images from the two eyes into a fused, single percept. This is elegantly described by Tong, Meng and Blake (2006):

- "During binocular rivalry, conflicting monocular images compete for access to consciousness in a stochastic, dynamical fashion. Recent human neuroimaging and psychophysical studies suggest that rivalry entails competitive interactions at multiple neural sites, including sites that retain eye-selective information. Rivalry greatly suppresses activity in the ventral pathway and attenuates visual adaptation to form and motion; nonetheless, some information about the suppressed stimulus reaches higher brain areas. Although rivalry depends on low-level inhibitory interactions, high-level excitatory influences promoting perceptual grouping and selective attention can extend the local dominance of a stimulus over space and time. Inhibitory and excitatory circuits considered within a hybrid model might account for the paradoxical properties of binocular rivalry and provide insights into the neural bases of visual awareness itself."

TopOwl™ biocular/binocular HMD with 100% overlap

Left eye view Right eye view

Figure 12-7. The TopOwl™ (top) is a biocular/binocular HMD with 100% image overlap manufactured by Thales. It is currently being used by military helicopters in several countries. The simulated views of the left and right eyes, as shown, are identical and can be free-fused. This is typical when the image is generated by a single sensor (e.g., nose-mounted FLIR), but not when using I^2 tubes mounted on the sides of the helmet, which produce images from different perspectives. Symbology overlay, as depicted, is to one eye only. With the TopOwl™ system monocular or biocular symbology overlays are optional. As with monocular displays, a high contrast background with high spatial frequencies can reduce, but not eliminate binocular rivalry with symbology presented to one eye. As the contrast and higher spatial frequencies of the background lessen, as it can with low contrast I^2 images, the problem of rivalry can increase (Patterson et al., 2007a).



Helmet Integrated Display Sight (HIDSS) with partial-overlap

Combined left and right eye image

Figure 12-8. The HIDSS was a prototype partial-overlap HMD developed by Rockwell-Collins-Kaiser for the Comanche helicopter project. The image depicted here has a simulated 45% overlap that would be biocular or binocular. The symbology is within this area. As with a HMD with full-overlap, the symbology can be presented to one or both eyes. Optically there is no border between the biocular/binocular and monocular portions of the full image. However, a form of rivalry called luning can occur, forming a perceived boundary between the two regions (Klymenko et al., 1994a).

Two of the major parallel visual geniculotstriate neural pathways, from the retina to the lateral geniculate nucleus (LGN) and from the LGN to the striate visual cortex, are the magnocellular (M) pathway and the parvocellular (P) pathway (see Chapter 6, *Basic Anatomy sand Physiology of the Human Eye*). The M and P neural pathways go from retina to the lateral geniculate nucleus and the striate visual cortex in the brain. The M pathway fibers originate from retinal rod photoreceptor cells and the P pathway fibers from cone photoreceptor cells. Information from the M pathway goes to the parietal lobe of the brain involved in processing "where" object-events happen. Neural cells along this pathway are particularly sensitive to movement and lower spatial frequencies associated with overall shapes of objects. Information from the P pathway goes to the inferotemporal lobe of the brain and is involved in the "identification" of objects. The neural cells along this pathway are particularly sensitive to color, the higher spatial frequencies or fine detail of objects, and contrast object contours. There is considerable evidence that image conflicts in the P pathway lead to rivalry, and image conflicts in the M pathway generally do not (He, Carlson and Chen, 2005).

A major debate has emerged in binocular rivalry research community. It is basically about a top-down versus a bottom-up model (Andrews, Sengpiel and Cohen 2005; Blake, Westendorf and Overton, 1980; Crewther et al., 2005). Andrews, Sengpiel and Cohen (2005) represented the debate as follows:

> "Two general theories have emerged. One possibility is that visual information is suppressed by inhibitory interactions prior to or at the stage of monocular confluence. In this concept, changes in perception would be mediated by shifts in the balance of suppression between neurons selective for one or another monocular image. Since these interactions must occur early in the visual pathway (e.g., the lateral geniculate nucleus or layer 4 of primary visual cortex), any changes in the activity of neurons in higher visual areas, would be explained by a loss of input, perhaps equivalent to closing one eye. The alternative hypothesis is that rivalry reflects a competition between different stimulus representations. This would be comparable to the viewing of other bistable stimuli, such as the vase-face stimulus, and as such would be relevant to the resolution of ambiguity in normal viewing."

The general consensus is that binocular rivalry occurs at multiple stages of visual processing (Alais and Blake, 2005; Blake and Logothetis, 2002; Tong et al., 1998).

It should be pointed out that there are many parallels between the study of binocular rivalry (and related ambiguous figures) and attention. There is both evidence and speculation that they all may reflect common, general neural mechanisms that influence the perceptual content of conscious awareness (Freeman, Nguyen and Alais, 2005).

There is an extensive literature on binocular rivalry. Currently, a comprehensive bibliography is maintained by Robert O'Shea (2009) of the University of Otago, New Zealand. Equally informative is a binocular rivalry demonstration website maintained by Randolph Blake (2008) of Vanderbilt University, Nashville, TN.

There also are a number of excellent reviews of binocular rivalry and its impact with the use of HMDs (see Alais and Blake, 2005; Blake, 2001; Hershberger et al., 1975; Howard, 2002; Laramee and Ware, 2002; Klymenko et al., 1994a,b; Patterson et al., 2007; Winterbottom et al., 2007).

The most familiar result of binocular rivalry is the alternation in consciousness of competing images from the two eyes. The dominant image cannot be held indefinitely (Blake, 2008). Binocular rivalry suppression takes time to develop, on the order of 200 ms; a single fused image containing both fusable and rivalrous features can form, too quickly be followed by rivalry of the incongruous features. Eye sighting dominance probably has little impact on the length of time an image is retained in consciousness (Howard, 2002; Rash, Verona and Crowley, 1990; Rash et al., 2002). Over an extended viewing time, the rate of alternations generally slows.

Image alternation is not strictly periodic, with durations generally following a Gamma distribution (Blake, 2008). There are many factors that influence alternation, but average duration of dominance generally remains constant, whereas the average duration of suppression varies inversely with stimulus strength; weak patterns tend to remain suppressed longer, increasing overall predominance of the stronger image, i.e., percentage of total viewing time (Blake 2008). The depth of suppression, loss of visual sensitivity, is on the order of 0.3 to 0.5 log units.

Hershberg et al. (1975) reviewed the then-current literature on binocular rivalry and HMDs. They also performed a number of studies, including a determination of ambient scene predominance using a HMD based on contour strength variables. They defined predominance as the percentage of total viewing time during which a rivalrous image was perceived at a visibility of 90% or more. They found significant predominance effects for: 1) ambient scene complexity, 2) HMD resolution, 3) HMD luminance, 4) ambient scene luminance, 5) HMD FOV, and 6) HMD contrast. These remain key binocular rivalry variables in HMD design and use. As relative strength of competing patterns are determined by variables such as contour density, pattern contrast, spatial frequency content, and motion (Blake, 2008), it is clear that the quality of images, as defined by blur, contrast, and high spatial frequency content have an impact on the occurrence and duration of the dominant percept.

There is evidence that cognitive factors can influence binocular rivalry alternation (Leopold et al., 2005; Freeman, Nguyen, and Alais, 2005; Patterson et al., 2007). This has fueled the bottom-up versus top-down debate regarding the neural mechanisms in binocular rivalry. Clearly, figure identity, a high order of visual processing, can be a significant factor in alternation of ambiguous figures. However, the issues in binocular rivalry are not clear-cut, where possible feed-forward and feed-back mechanisms (retroinjection to striate cortex) produce complexity (Blake, 2008; Crewther et al., 2005; de Weert, Snoeren and Konig, 2005). Blake (1988) used dichoptic presentation of meaningful and nonmeaningful text. He found no special effect of meaningful text on rivalry. As Helmholz observed, intentional effort to maintain a dominant stimulus was effective but did not prevent alternation (Patterson et al., 2007; Blake, 2005). Chong and Blake (2006) demonstrated that both exogenous and endogenous attention could increase stimulus strength of a dominant stimulus, thereby increasing its predominance. Patterson et al. (2007), in reviewing the impact of cognitive factors on binocular rivalry, concluded that attention, while having an effect on alternation, did not have a large effect, only by as much as 50%.

The impact of Gestalt grouping, particularly when the features of the rival stimulus and the neighboring features form a coherent, global pattern, can increase predominance (Alais, and Blake, 1999; de Weert, Snoeren and Konig, 2005; Engel, 1956; Kovacs et al., 1997; Lee and Blake, 1999). Papathomas, Kovacs and Conway (2005) suggested that a model for Gestalt organization factors may be somewhere between top-down and bottom-up.

Fusion of an image is both independent of binocular rivalry and tends to counter its occurrence (Blake and Boothroyd, 1985; Patterson et al., 2007). Fusion takes precedence over rivalry, a particularly important factor in see-through monocular HMDs with symbology superimposed on an outside scene. However, this consideration interacts with many other variables, including contrast, contour density, color, and spatial frequency content of the competing images.

Another aspect of binocular rivalry is seen with partial-overlap binocular HMDs. Partial-overlapping is a technique used to expand the total FOV of binocular HMDs (see Chapter 3, *Introduction to Helmet-Mounted Displays*). A common portion of the angular regions seen by each eye is fused into a single percept, i. e., viewed binocularly. The right and left portions of the total FOV, flanking the fused binocular region, are viewed monocularly. Luning can develop at the transition between the two, probably a rivalrous, subjective darkening crescent area at the binocular-monocular border (Figure 12-9). This border region can be an area of reduced visibility for visually foveated objects. This is more pronounced with divergent than convergent overlap and with smaller angular regions of binocular overlap. Divergent overlap is when the monocular area imaged on each retina is from the same side as the imaging eye, whereas convergent overlap is where the monocular area imaged on each retina is from the side opposite from the imaging eye.

Klymenko et al. (1994a, b) performed a series of experiments to determine factors affecting the visual fragmentation, phenomenal segregation of the total FOV into two distinct monocular areas and a binocular area. They concluded, along with other researchers, that luning is more pronounced with the divergent mode than the convergent mode. They confirmed that luning could be reduced by placing a "competing edge in the monocular field of the informational eye in order to strengthen it relative to the monocular field border of the noninformational eye. They speculated that blurring the border with the 'noninformational eye would also weaken luning.

As stated by Patterson et al. (2007), binocular rivalry does have a negative impact on visual performance, including increased reaction time and missed information/signals. They went on to say that observers (using HMDs) are not always aware of these decrements in performance.

Much of the information regarding visual performance with HMDs has been gained through pilot surveys (Patterson et al., 2007; Heinecke et al., 2008; Hiatt et al., 2004; Rash, and Martin, 1988; Rash et al., 2004; Rash et al., 2002). While this literature has detailed and high-lighted problems users have with the monocular design of

Figure 12-9. The border between the biocular/binocular central portion of a partial-overlap HMD and the flanking monocular sections can have a crescent shaped area of diminished visibility called luning, probably a variation of binocular rivalry. This area of reduced visibility can obscure objects in the field-of-view (Klymenko et al., 1994a, b).

the Integrated Helmet and Display Sighting System (IHADSS) deployed in the AH-64 Apache helicopter (Figure 12-10), surveys cannot separate out causes of visually related performance issues like undetected drift, estimates of rate of closure, slant detection, nor is it an effective medium for separating out factors like attention, monocularity, and poor image quality that can confound the relationship between reported performance issues and binocular rivalry. Despite these factors, Rash et al. (2002) found 64.4% of AH-44 Apache aviators using the IHADSS system reported unintentional alternations during flight. Most of the aviators surveyed (74.5%) reported being able to switch their attention easily from one eye to the other, and 44.9% reported having developed a strategy to aid switching such as closing one eye, glancing away, or blinking both eyes. One pilot reported “retinal rivalry when there is too much ambient light”; another reported that "If a bright light suddenly comes into view your unaided eye will dominate;” and yet another pilot reported “Binocular rivalry can occur at any time. We just deal with it (e.g. momentarily close one eye).”

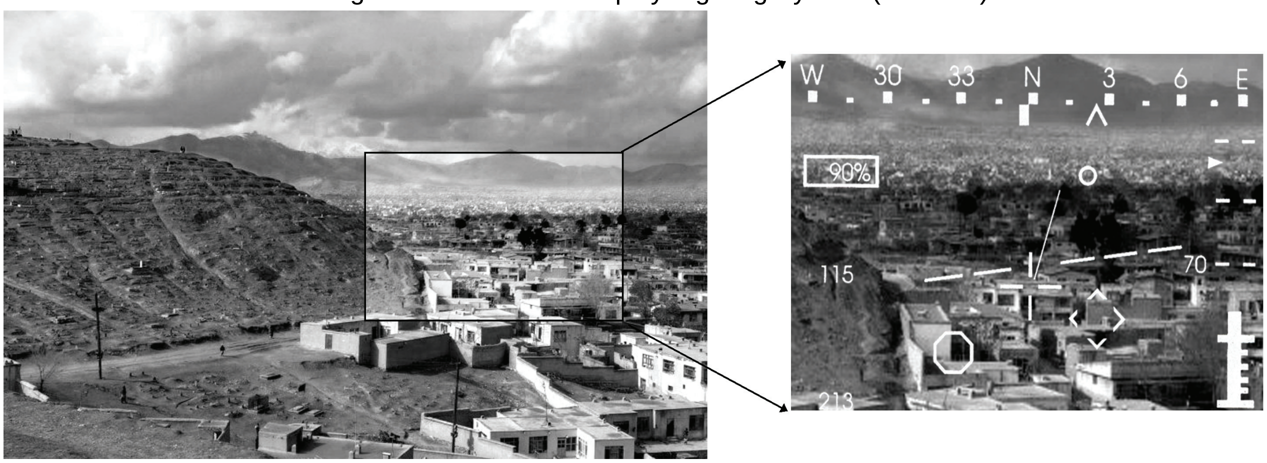
Most surveys of visual issues with the IHADSS were conducted during peace time. However, Heinecke et al. (2008) surveyed Apache aviators using the IHADSS during urban combat in Operation Iraqi Freedom. Their results generally paralleled those from other surveys. There was, however, on striking result that was unexpected. The incidence of problem reports was down. It would seem that the stress of combat directed attention away from equipment-user-problems and towards the task of simply making the equipment they had perform as well as possible.

The Heinecke et al. (2008) report and others importantly demonstrate that humans are adaptable and find ways to make things work and new ways to apply technology. Making a monocular HMD work, with all its problems, binocular rivalry being one, is a case in point. The IHADSS has had a long, successful history and weathered several changes in military mission. The history of the individuals who have used it should provide encouragement for individuals trying to design new HMD systems with fewer problems that help users perform better and with greater transparency.

Hyperstereopsis

The human visual system is based on two visual detectors (the eyes), slightly separated in location on the front of the face. The distance between the pupils of the two eyes is known as the intraocular, and more commonly, the interpupillary distance (IPD). Each eye's retina captures a separate and slightly different image of the external scene. The differences in the two retinal images are called horizontal disparity, retinal disparity, or binocular

Integrated Helmet and Display Sighting System (IHADSS)



Simulated view of a scene viewed while using the IHADSS HMD. The left (unaided eye) sees the cockpit and outside world; the right eye views a 30° (V) by 40° (H) portion of the outside world overlaid with symbology.

Figure 12-10. The IHADSS (top) is a monocular HMD system first developed by Honeywell and currently manufactured by Elbit. This system is used on the U.S. Army Apache AH-64 attack helicopter. The images above represent a daytime application of IHADSS. However, it should be noted that this system is usually used at night with a FLIR image in the HMD and a nighttime view of the cockpit and outside world available to the other eye. Under these conditions pilots learn to suppress vision in one eye, while attending to the image in the other. This ability, however, is not perfect and unexpected alternations do occur.

disparity. When processed by the brain, the result is a perception known as *stereopsis*, which is a binocular cue to depth perception (see Chapter 7, *Visual Function*). Humans generally do not notice depth in objects that are more than a few hundred feet away. This is because at this distance and beyond, the rays arriving at the eyes are essentially parallel, and the retinal disparity and binocular object perspective cues become too small to resolve.

Stereopsis assists in the ability to estimate absolute distances between ourselves and an object, as well as the relative distances between two objects, i.e., which is closer. However, depth perception does not depend on stereopsis alone. Multiple visual cues are used to define our sense of depth. Both differences and similarities between two retinal images are fused and compared within the brain to produce depth perception (Hill, 2004). The cues for depth perception also may be monocular. Monocular cues include:

- Relative size
- Interposition
- Geometric perspective
- Contours
- Shading and shadows
- Monocular motion parallax

For the civilian community, the IPD, defining the separation between the two retinal images, ranges from 57 to 72 mm (1st to 99th percentile male) and has an average of 64 mm. The 95th percentile of U.S. military personnel falls within the 55 to 72mm range of IPD. The average IPD for U. S. Army males is 64 mm and 61 mm for females (Donelson and Gordon, 1991).

In artificial situations where the input sources are located at greater than normal IPD, a condition called hyperstereo exists. A number of terms have also been applied to this visual condition, e.g., hyperstereopsis, tele-stereo, enhanced-stereo, etc. In many such hyperstereoscopic contexts, the separation between the (sources of the) inputs to the two eyes is referred to as the stereo baseline (distance). See Chapter 15, *Cognitive Factors*, for a case study discussion of an example hyperstereo HMD designs.

The effect of greater-than-normal separation of the inputs to the two eyes produces very complicated and varied results that depend on the amount of separation and the point of fixation. For example, a pilot usually will perceive the near ground as if rising up to him/her. When a helicopter pilot is sitting on the ground, it may seem that ground level outside the cockpit is at chest level, causing some pilots to say it looks like they are sitting in a hole. However, distant objects may look natural.

When this greater-than-normal separation of inputs to the two eyes exists, the convergence angle to an object being viewed is increased as compared to the convergence angle that exists for a "normal" IPD. This can cause the distance to a viewed object to appear shorter and the object to appear closer. This difference in perceived distance due to a change of convergence angle is depicted in Figure 12-11. For a normal interocular separation distance (i.e., IPD), the target point located at distance D subtends an angle of α**.** For the increased separation distance depicted for the I^2 tubes in this diagram, the convergence angle (for this configuration) increases to β (top of diagram). However, the human visual system is still operating from the "assumption" of a normal IPD. As a consequence, the apparent convergence angle of β (bottom of diagram) causes the target object's distance to be perceived as D'; D' < D, hence, the target object appears closer. The object size will appear to be approximately the same at both D and D', giving the impression that the object is smaller.

In addition to objects appearing closer, another manifestation of hyperstereo is the ground appearing to slope upward, toward the observer, creating what is often described as a "bowl" or "dish" effect. While it is a commonly used analogy, it is slightly erroneous one. Figure 12-12 attempts to better render the illusion and presents it more as a "mountain top crater" effect. The observer describes the ground nearest to him as appearing closer (higher), with this exaggerated depth effect (the closer than effect) decreasing with distance away from the observer. When the helicopter is on the ground, the pilot perceives the near ground as being at chest level, while distant objects may look natural, a result of the non-linearity of the exaggerated depth perception with increasing distance from the observer.

This hyperstereo effect results from an increased IPD and not from a proportional increase in the vertical dimension subtended by an object. The proportional angular impact of convergence decreases with distance, consequently making the apparent relative horizontal and vertical dimension of objects appear more and more normal. The hyperstereoscopic distortion is largely, although not entirely, a near effect that is usually manifested within a few hundred feet. A good-rule-of-thumb is that when the perspective differences of an object falls below one minute of arc, the impact of hyperstereo becomes negligible, and competing monocular depth cues become dominant.

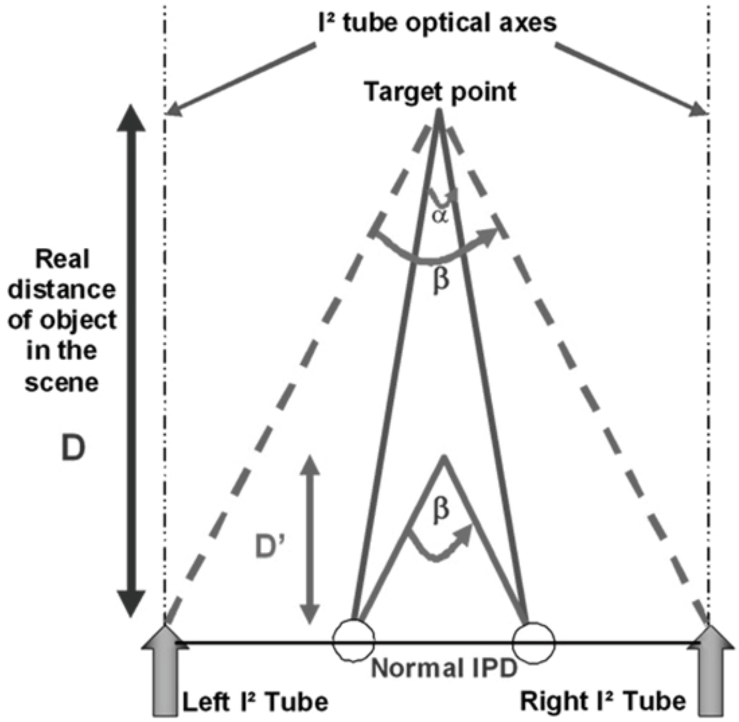


Figure 12-11. Diagram depicting change in perceived distance due to hyperstereo (Kalich et al., 2007).

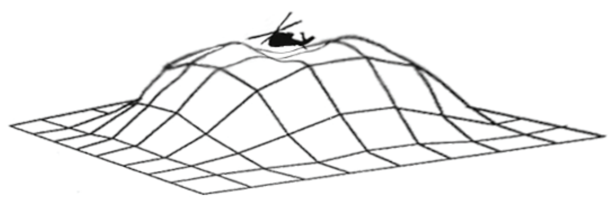


Figure 12-12. Depiction of "mountain top crater" illusion due to hyperstereo.

The preceding narrative is a superficial description of stereo vision and the special condition of hyperstereo. It is intended only to provide the background necessary to understand the impact on this phenomenon on HMD design. The concept of hyperstereo from a vision science perspective is a significantly more complicated topic. A more in-depth discussion would include rivalry of the retinal images and the potential impact of optical differences on hyperstereo effects (e.g., prism, binocular parallax, optical distortion, velocity and acceleration effects, etc.). Priot et al. (2006) provide an excellent review of the hyperstereo (hyperstereopsis) literature from an operational perspective.

Thus far, hyperstereo has been described as a potentially problematic attribute. However, some atypical hyperstereo configurations (based on camera pairs with extremely wide baselines or temporal delays with a single camera) have been investigated for their possible use in aerial search and rescue, target detection, and traversing drop-off terrain tasks (e.g., Cheung and Milgram, 2000; Schneider and Moraglia, 1994; Watkins 1997).

Studies evaluating hyperstereo vision

HMD designs with hyperstereo are not new and date at least to the mid-1980s. The U.S. military has evaluated and conducted studies on several proposed designs. Additional studies have investigated the potential advantages of hyperstereo. The following is a synopsis of the more relevant studies and papers pertinent to this discussion:

- In 1990, the National Aeronautics and Space Administration (NASA) investigated hyperstereo for its potential use in improving hover-in-turbulence performance in rotorcraft (Parrish and Williams, 1990). While objective measures demonstrated some improvement in situation awareness, control activity, and hover stability, pilots reported a subjective dislike because of the exaggerated visual cues experienced.
- In 1992, the Night Vision Laboratory (currently Night Vision and Electronic Sensor Directorate), Fort Belvoir, Virginia, conducted an evaluation of the potential use of the Honeywell INVS/MONARC HMD in helicopters. The INVS was being developed in an attempt to design a night vision I^2 system with lower weight and improved center of mass for fixed-wing aircraft. The objective lenses and intensifier tubes were placed on the side of the helmet with a separation approximately 4X that of normal IPD, introducing the condition of hyperstereo. The study's objective was to compare aviator performance with INVS to performance with ANVIS. On initial concept flights in a TH-1 helicopter (modified AH-1S Surrogate trainer), pilots found the hyperstereopsis and sensor placement on the sides of the helmet to be major deficiencies during terrain flight. The vertical supports in the canopy always seemed to be within the FOV with any head movement, and under starlight conditions, the pilots rated the hyperstereo system unsafe and terminated the study except for demonstration rides (Kimberly and Mueck, 1992). The reported hyperstereo effects were characterized by intermediate and near objects appearing distorted and closer than normal. The ground was reported as appearing to slope upwards toward the observer and regions beneath the aircraft appearing closer than normal. Safety pilots noted a tendency to fly higher than normal during terrain flight.
- In 1992, the U.S. Air Force also conducted testing on potential ejection-safe HMD designs that demonstrated the hyperstereo effect under the Interim-Night Integrated Goggle Head Tracking System (I-NIGHTS) program (Grove, 1992; Gunderman and Stiffler, 1992). I-NIGHTS began as a joint Air Force/Navy development with the Navy as the designated lead. Candidate systems were designed by Kaiser Electronics, Honeywell (same as MONARC) and GEC Avionics). All three designs placed the I^2 tubes at greater than normal IPD. Flights were conducted in the HC-130 (fixed-wing) and MH-53 and MH-60 helicopters. Interestingly, the final reports do not provide either the I^2 separation distances for the HMDs or subject pilot IPDs. The hyperstereo effect apparently was not anticipated, as the flight performance evaluation questionnaire did not specifically ask about this effect, asking only one generalized question regarding image distortions. However, within individual comments, the helicopter pilots reported that the Kaiser HMD "slightly magnified images, creating the illusion of being lower than actual altitude. This became very apparent during landing where the pilot anticipated touchdown at the any moment while he was actually still 3-4 feet in the air."
- In 1993, in support of the development of the Helmet Integrated Display Sight System (HIDSS) HMD for the U.S. Army's RAH-66 Comanche helicopter, the USAARL and the U.S. Army Aviation and Technical Test Center (ATTC), Fort Rucker, Alabama, conducted a flight study which included an investigation of the effects of hyperstereopsis on aviator performance (Armbrust et al., 1993). Eight subject aviators flew 150.5 flight hours in an AH-64 Apache. Subjects performed a series of six modified ADS-33C (U.S. Army, 1989) maneuvers while wearing the ANVIS, Eagle Eye, and MONARC HMDs. These three systems represented IPD ratios (to normal) of 1X, 2X, and 4X, respectively. The effect of hyperstereo viewing on aviator performance was evaluated through the collection of quantitative (i.e., accuracy of hover, drift and heading) and subjective measures (i.e.,

Subjective Workload Assessment Technique [SWAT], Perceptual Task Rating Scale [PTRS], and Subjective Performance Rating Scale [SPRS]). The study concluded that the effects of hyperstereo were minimal. It was stated that aviators "learned compensation strategies quickly." However, it was noted that performance involving altitude estimation was affected to a greater extent. Overall, none of the subjective measures showed any difference in workload associated with the three systems. However, for low level tasks, data did show that the two hyperstereo HMDs were more difficult to fly than ANVIS.[4]

- In 1995-1996, Leger et al. (1998) conducted a two-phase flight test of an earlier configuration of the current TopOwl™ HMD, i.e., visor projection and 40-degree, fully-overlapped FOV. Sixty-six hours were flown in Phase One (40 hours at night; 77 flight hours were accumulated in Phase Two (45 hours at night). While various platforms were used, most of the evaluation was conducted on a SA 330 (Puma) test-bed platform developed for the TIGER program. The interocular separation was 240 mm, 46 mm less than that of the current TopOwl™ version, and was approximately 4X normal IPD. The independent variables in the study were distance and height above the ground. The study reported "a systematic under-estimation of distance and height, (with) pilots feeling closer and lower than they really were." Pilots were reported to have "returned to nominal performance" after 5 to 10 hours of flight.
- In 1998, two German test reports documented flight experience with two hyperstereo HMD designs, the Knighthelm and the TopOwl™ (Hohne, 1998; German Air Force Test Center [WTD], 1998; in Priot et al., 2006). Both evaluations reported altitude evaluation errors. A later German evaluation of just the TopOwl™ concluded that: "The approximately double base distance of the objective lens[es] in relation to the eye creates a false range feeling during hover flight when evaluating the aircraft altitude. The impression gained is one of a low hovering altitude" (Krass and Kolletzki, 2001). In all three evaluations, pilots reported the ability to compensate after relatively few flight hours.
- In 2001, the U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland, conducted a study on the effects of hyperstereo viewpoint offsets of NVGs on accuracy in a simulated grenade-throwing ground task (CuQlock-Knopp et al., 2001). In the study, 32 National Guardsmen were tasked with throwing simulated grenades onto a trap-door target located 20 feet away. The measured data were the radial direction and distance from the target for each toss. Three viewpoint (hyperstereo) configurations (Figure 12-13) were compared to the normal IPD ANVIS. Only two of the three configurations presented a horizontal displacement; the third presented a vertical displacement only. The two horizontal hyperstereo distances were approximately 6.7 and 8.5 inches (170 and 216 mm), both equating to approximately 3X normal IPD. The results of the study showed that the hyperstereo resulted in a statistically significant increase in the magnitude and direction of the throwing errors.
- In 2005, the USAARL conducted a flight investigation in the UH-60 where aviators serving as the co-pilot (but not on the controls) wore the TopOwl HMD. Subjects reported an approximate 6-8 hours acclimation period that is consistent with manufacturer's claims. However, evaluations of standard

[4] One of the authors was a participant in the joint ATTC/USAARL study summarized herein. In his opinion, the reported findings did not fully capture the impact of hyperstereo on aviator performance. First, due to logistical issues, the flights were conducted under extremely benign conditions and at locations that provided too many overriding cues. Second, the AH-64 aircraft provides the least forward looking vision of any U.S. Army aircraft. This inability to look forward circumvented the potential of the pilots to accurately assess the hyperstereo effects. Third, a through review of recorded pilot comments frequently included the perception of "landing in a hole" and having to "feel for the ground." In addition, safety pilots noted that subjects were consistently flying higher than required during terrain flight and had greater difficulty with aircraft drift. These issues were noted in the original report, but were not fully presented in the summary findings.

flight maneuvers identified height estimation, slope estimation and dynamic performance (e.g., rate of closure) as issues requiring addition study (Kalich et al., 2007).

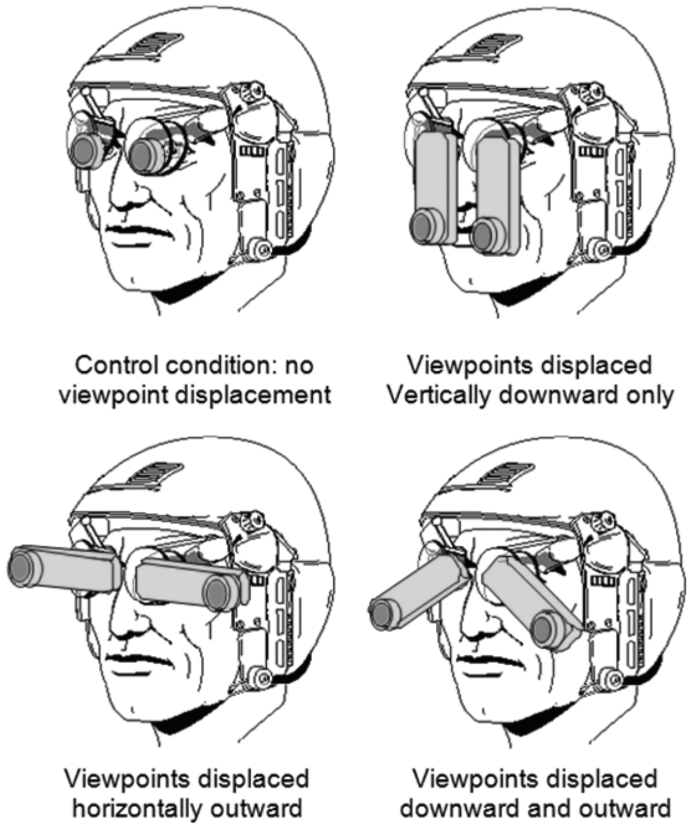


Figure 12-13. An artist's rendition of the four viewpoints used in a simulated grenade-throwing task study (CuQlock-Knopp et al. 2001).

- The most recent studies have been simulation studies conducted by Australian researchers to investigate time to contact, slope, and absolute distance estimation. In the first study (Flanagan, Stuart and Gibbs, 2007), the increased apparent distance created by hyperstereopsis was investigated for moving surfaces approaching observers (as in shipboard operations). There is concern that the hyperstereo display will result in a greater apparent speed of approach towards the surface, and operators will have the impression they have reached the surface before contact actually occurs. Motion towards a surface with hyperstereopsis present was simulated and judgments of time to contact were compared with those under normal stereopsis as well as under binocular viewing without stereopsis. Approaches to a large, random-textured field were simulated. It was found that time to contact estimates were shorter under the hyperstereoscopic condition than those under normal stereo and no stereo, indicating that hyperstereopsis may cause observers to underestimate time to contact leading operators to undershoot the ground plane when landing.
- Stuart, Flanagan and Gibbs (2007a) looked at the potential of the presence of hyperstereopsis to distort the perception of slope in depth (an important cue to landing), because the slope cue provided by binocular disparity conflicts with veridical cues to slope, such as texture gradients and motion parallax. In the experiments, eight observers viewed sparse and dense textured surfaces tilted in depth

under three viewing conditions: normal stereo, hyper-stereo (4X magnification), and hypostereo (1/4X magnification). The surfaces were either stationary, or rotated slowly around a central vertical axis. Stimuli were projected at 6 meters (19.7 feet) to minimize conflict between accommodation and convergence, and stereo viewing was provided by a Z-Screen™ and passive polarized glasses. Observers matched perceived visual slope using a small tilt table set by hand. Slope estimates were found to be distorted by the presence of hyperstereopsis, but to a much lesser degree than predicted by disparity magnification. The distortion was almost completely eliminated when motion parallax was present.

- The final study cited here (Stuart, Flanagan and Gibbs, 2007b) investigated the potential of increased camera separation (hyperstereo) to affect absolute depth perception, because it increases the amount of vergence (crossing) of the eyes required for binocular fusion, and because the differential perspective from the viewpoints of the two eyes is increased. The effect of hyperstereopsis on the perception of absolute distance was investigated using a large-scale stereoscopic display system. A fronto-parallel textured surface was projected at a distance of 6 meters (19.7 feet). Three stereoscopic viewing conditions were simulated – hyperstereopsis (4X magnification), normal stereopsis, and hypostereopsis (1/4X magnification). The apparent distance of the surface was measured relative to a grid placed in a virtual "leaf room" that provided rich monocular cues, such as texture gradients and linear perspective, to absolute distance as well as veridical sterescopic disparity cues. The different stereoscopic viewing conditions had no differential effect on the apparent distance of the textured surface at this viewing distance.

In a joint flight study between Canada, Australia and the United States, conducted in August 2008, but not yet reported, pilot interviews following an average cumulative flight time of 9 hours using the Thales Avionics TopOwl™ HMD, indicated that some level of adaptation to the hyperstereo effect may be achievable. With the exception of within 2-3 feet of the aircraft, the previously described “bowl” or “dish” effect seemed to no longer be experienced. This is a promising finding, but final analysis of the data has not been completed.

The Concept of Illusions

The premise underlying this section is that the phenomena usually classified as visual illusions are an essential part of normal daily vision. They are integral to what and how humans see. In fact, some vision scientists argue that much of what our visual system does under normal conditions, with all its neural machinery, may be devoted to overriding the myriad illusions that are experienced on a routine basis. The following discussion argues that visual illusions are constant, though usually unnoticed, companions to the human visual system. Operationally, vulnerability to visual illusions sets up conditions that are important for the design and use of HMDs.

Just because many illusions normally go unnoticed does not mean that they are all so well-behaved in a visual world that includes HMDs. Many display technologies and strategies specifically capitalize on the propensity of the visual system to be fooled.

Defining visual illusions

A formal definition of visual illusions would be a logical way to start this section; but as Boring (1942) noted: “[s]trictly speaking, the concept of illusion has no place in psychology, because no experience actually copies reality…. In the sense that perception is normally dependent upon subjective factors as well as upon the stimulus, all perception is ‘illusory’ in so far as it does not precisely mirror the stimulus. In this broad sense, the term *illusion* becomes practically meaningless.” This point is important because the word *illusion* should denote more than just a failure to mirror precisely the stimulus. Gregory (1996) makes a similar point, noting that it is a lot easier to provide examples of different illusions and fit them into different categories than it is to provide a good definition.

Nevertheless, there are at least two broad types of definitions for illusions. One type of definition notes the differences between some aspect of reality and the perception of that aspect. This type of definition emphasizes the disparity between the perception and the reality, an emphasis that seems to presuppose the existence of perceptions without such disparities, which, as Boring pointed out above, is not all that sound a premise. This type of definition also invites considerable philosophical speculation about *reality* and *truth*. Another type of definition seems to carry with it some implied explanation or mechanism, such as misperception of size, distance, shape, lighting, or color. The result is that a definition of illusions, like the illusions themselves, is a surprisingly elusive.

The study of illusions[5]

Scientists have been systematically studying illusions since at least the middle of the 19th century.[6] These scientists have argued that illusions reveal something about how the visual system goes actually functions. At the very least, illusions may be tools for understanding the normal workings of the visual system. Like any other tool, its usefulness depends on how it is used. Certainly, in the military battlespace that includes HMDs and their symbology, visual illusions have a pragmatic importance. In essence, our interpretations of synthetic vision displays, virtual reality displays, and conformal displays are at their core visual illusions, albeit controlled ones.

Proximal vs. distal stimuli

In a discussion of illusions, it is important to distinguish between the physical object and the image of that object on the retina. The retinal image is sometimes called the *proximal* image, because that is the stimulus that is close, directly landing on the sensory receptor system and directly affecting it. The physical objects that exist in the distance, sometimes called the *distal* stimuli, really have no direct impact on the receptor system itself. All the visual information about the physical world and all the objects that it contains depend on the proximal, retinal image. It is on the retina that the light energy has its biological effects on the retinal photoreceptors. The visual system constructs the distal world from the retinal image, in a sense back-projecting the proximal stimulus to the world. The problem is that "… since a given state of retinal stimulation is compatible with a countless number of distal arrangements, there is necessarily an irreducible equivocally in optical stimulation that makes going from optical input to distal arrangement impossible" (Epstein, 1995).

Distance perception cues

A previously-made statement is that many visual illusions generally go unnoticed in daily life. It is not that they are ignored, they're just not "seen." For example, consider the perception of depth. The importance of having two eyes for the perception of depth is considered absolute; that each of the two eyes has a slightly different view of the world, and the disparity between the two eyes is important for seeing depth. But many of the cues for depth are actually monocular. However, humans with only one eye away can see and judge depth quite well. But the retina, upon which the optics of the eye projects that rich ménage of everything we see, is really a two dimensional surface stretched on the inside of the rear wall of the eye. There is no depth information within that image; or more precisely, there is no more depth information there than can be found on a printed page. It is true that the world is 3-D; it has depth. It is also true that our perception of the world is 3-D, i.e., containing depth

[5] This section will discuss only visual illusions although all the major sensory systems – hearing, vestibular, kinesthetic, somatosensory, etc. – have demonstrated illusory phenomena.

[6] J.J. Oppel is usually credited with the first systematic study of what he referred to as optical geometric illusions in: Uber geometrisch-optische Tauschungen. *Jahresbericht des Frankfurter Vereins,* 55, 37-47, (1854-1855).

information. But, the interface, the retina, is flat containing no depth but just a pattern of light. Whatever depth we appreciate with one eye depends as much on illusion as does any impression of depth conveyed by relatively poorly-printed graphics on a sheet of paper. Therefore, exploring the monocular cues to depth will be instructive in understanding illusions.

Monocular depth cues

Interestingly, the first understanding of monocular depth cues was discovered and mastered over the centuries by artists with scientists following (surprisingly far) behind, performing the more mundane work of cataloguing, classifying, analyzing, and possibly even explaining these cues. In our discussion we will introduce briefly some of the more obvious of these monocular cues with further exploration being left to the reader via any of the standard texts on visual perception (e.g., Sekuler and Blake, 2005; Wolfe et al., 2005).

Monocular depth cues can be organized in three general categories: (a) cues derived from pictorial renderings of an image on a surface like the retina, (b) cues derived from the physiological responses of the eye, and (c) cues derived from the motion of the eye.

Pictorial depth cues

Pictorial cues are probably the most obvious and are described by many visual perception text books and include the following:

- *Linear perspective* refers to the compelling impression that a pair of straight, parallel lines (like railroad tracks or highway lanes) seem to get closer together the further they are in the distance. In other words, the size of the retinal image of an object gets smaller as the object gets further away. See Figure 12-14.

Figure 12-14. Linear perspective as a monocular cue. The highway lanes appear to get closer together the further away they are.

- The *relative size* of known objects provides some distance information. Up to a point, one can estimate how far away someone is by how big the person's image is. Since people are generally between five and six feet tall, seeing someone smaller than one's thumb induces the belief that the person appears to be far away, not just tiny. See Figure 12-15.

Figure 12-15. Relative size as a monocular cue. In this painting, by making the size of the people smaller, they are perceived as being further away.

- *Detail perspective (texture gradient)* is closely related to linear perspective. Since the surface of most objects has textural detail, the amount of textural detail that can be seen depends on distance. The person may be too far away to recognize the person's face or even whether the person is a man or a woman. The facial features are one textural cue; there is also the textural gradient of the terrain between the observer and that distant person. The gradient of texture visible in the intervening terrain also provides distance information. See Figure 12-16.

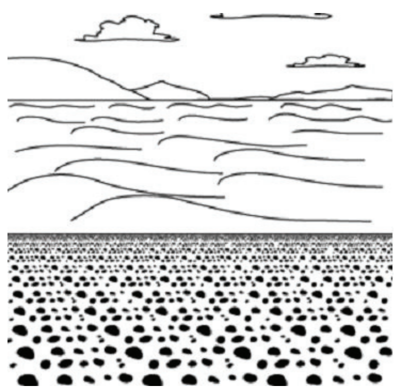


Figure 12-16. Texture gradient as a monocular cue. Pebbles on a beach or waves on the sea look rougher closer up than from a distance; also note the cobblestones of Figure 12-15.

- *Aerial perspective* becomes important if the distances involved are great enough. The atmosphere scatters light; and the more the scatter, the greater is the distance. Furthermore, the amount of scatter depends upon the wavelength (color) of the light; the more the scatter, the shorter is the wavelength.[7] Leonardo da Vinci noted: "There is another kind of perspective which I call aerial perspective because by the atmosphere we are able to distinguish the variations in distance of different buildings which appear placed on a single line; as, for instance, when we see several building beyond a wall, all of which, as they

[7] The components of light can be laid out with a prism to produce the spectrum of light, with its components sorted according to wavelength; the short wavelength components on one end and the long wavelengths at the other. These wavelengths appear as color, the short wavelengths appear as blue and the long wavelengths appear as red. Since, with all things equal, the shorter the wavelength, the greater the scatter; the further an object is, the more bluish is the haze around it.

appear above the top of the wall, look of the same size, while you wish to represent them in a picture as more remote one from another and to give the effect of a somewhat dense atmosphere … Hence you must make the nearest building above the wall of its real color, but make the more distant ones less defined and bluer …. If one is to be five times as distant, make it five times bluer" (Boring, 1942) (Figure 12-17).

Figure 12-17. Aerial perspective as a monocular cue. Image contrast declines with distance as the color shifts to the bluer part of the spectrum.

- The *relative brightness* of objects is a cue to their relative distance. Other things being equal, the closer an object is to the source of the light, the more bright the object looks. For example, a piece of paper lying on a desk under a light looks brighter than an identical sheet of paper laying further way from the light. If the light source is unseen, the visual system extrapolates (unconsciously and automatically) a light source using some simplifying assumptions. Among the cues these calculations incorporate are relative brightness, distance, and size.
- *Light and shade* provide subtle yet surprisingly powerful depth cues. Objects may shade other objects, contributing relative size and depth information about the objects, the light source(s), and the viewer. Objects can cast shadows on parts of themselves. Elements of the surface texture can cast shadows, and the gradient of the shadows may provide more information than the texture itself. Shadows also help differentiate hills from valleys in the same light. Such shadowing effects are illustrated in Figure 12-18.

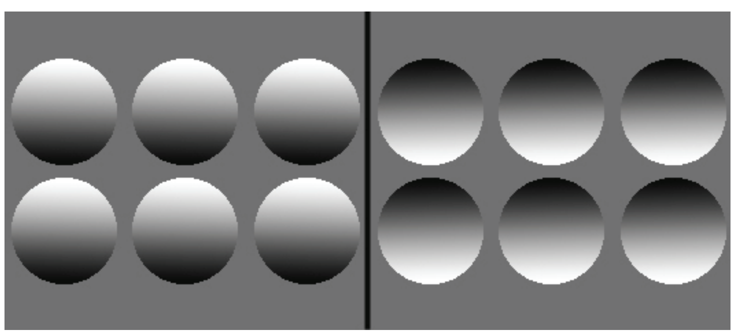


Figure 12-18. Light and shade (shadowing effect) as a monocular cue. The shadows indicate that the left-hand circles are convex and the right-hand circles are concave.

- *Interposition* is a strong cue. Although it is ambiguous, it is far easier to see Figure 12-19 as the King of Clubs lying on top of the King of Spades than it is to see the King of Spades missing a lower left part that is just the right size for it exactly fit next to the complete King of Clubs. In either case, the King of Clubs appears closer.

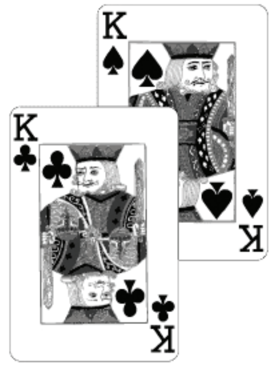


Figure 12-19. Interposition as a monocular cue. The card that appears to be on top also appears closer.

Most of these monocular depth cues are rather easy to appreciate; yet some of these, such as texture gradient, light, and shade may pose more of a challenge than other cues to incorporate in HMDs, virtual reality, synthetic vision or other displays. The way these cues are implemented almost certainly affects the perception and judgment of size and distance of objects in the visual scene, how they are laid out and their relative positions (Rogers, 1995).

Physiological depth cues

Physiological depth cues are less apparent than the monocular depth cues, and their impact on depth perception more difficult to assess. These physiological cues depend on the muscular activity or motion of the eyes:

- *Accommodation* refers to the change of the focusing power of the eye as the visual system shifts attention between objects that are at different distances. The physiology of accommodation is extraordinarily elegant; coordinating lens and iris/pupil diameter changes for the two eyes. The neuromuscular system controlling these binocularly coordinated responses is driven by distance cues which most evidence suggests are calculated from the types of blur in the retinal image of the specific objects on which the eye is focused at the moment. The eye's objective target is under conscious, higher order control, shifting from moment to moment, but the calculation of the retinal image blur in the image of the object of regard is unconscious and automatic. This rapid automatic analysis of the nature of the image blur is a depth cue to which the observer is oblivious.
- *Convergence* refers to the pointing of the eyes' line of sight. Each eye's line of sight must be coordinated so the eyes are looking at the same thing. That way, the two eyes triangulate. When the eyes look at something that is close, the lines of sight converge. When the eyes look at something far away, the lines of sight are less convergent. Convergence is frequently associated with eye elevation; things that are near tend to be lower in the visual world; things that are far tend to be higher. Since accommodation, convergence, and the pupil's response are all highly coordinated, their neuromuscular systems are closely coordinated but not identical since there is more voluntary control of the six extra-ocular muscles of each eye that determine eye pointing than there is of focusing and pupil constriction.

When the systems work, these depth cues function without being consciously noticed, but they are nonetheless important. In the context of the use of HMDs, these ocular-motor depth cues should be considered. A frequently-suggested strategy is to arrange the optical elements of the HMD so that the display approximates optical infinity. Thus, there is no relative motion between the HMD symbology and the objects seen through the HMD in the far distance, since they are superimposed on each other and are at the same optical distance. This strategy is based on the idea that the accommodation of the eye is at rest when the eye is focused at its far point. But some individuals apparently do not tolerate this strategy well, and its implementation becomes complicated when the users have refractive errors that need to be individually corrected. This is particularly challenging for individuals who are farsighted. It should also be noted that setting the optics of the HMD at infinity does not address whether accommodation, convergence, and pupil/iris constriction play a role in the size and distance perception of objects nor does it address any additional effects these ocular-motor responses may have on other visual illusions.

Boring (1942) makes the important point that in the history of experimental psychology the monocular depth cues, although obvious, were considered secondary and less important than the binocular cues of convergence and accommodation, which were considered the primary cues for depth. The fact that accommodation and convergence require a motor response contributed to the idea that they are the primary depth cues. The motor responses provide sensory motor information about the distances of the viewed objects whereas the monocular depth cues provide information with which *judgments* are made. Accordingly, painting of perspective was based on the unreliability of the depth cues. Distances of remote objects depend upon the monocular depth cues; these seen distances are the result of cognitive judgments whereas the binocular depth cues, which after Wheatstone's 1838 stereoscope, included retinal disparity (see below), were immediate and sensory. This distinction of primary and secondary depth cues may be seen as historically quaint on one hand; but on the other hand, it is reminiscent of today's language of bottom-up and top-down distinctions. In this context, HMDs and related display compromise the primary depth cues as well as the secondary ones.

Kinetic depth cues

Kinetic (motion-based) depth cues are those that derive from movements an eye makes as it views the world or by objects as they approach or recede. For eye movements, these are not the accommodation and convergence motions; rather, these motions involve the translation and rotation of the eye as the head and body move through the environment. As objects in motion become smaller, they appear to recede into the distance or move farther away; objects in motion that appear to be getting larger seem to be coming closer. Using kinetic depth perception, the brain calculates a *time to contact distance* at a particular velocity. For example, automobile driving requires constantly judging the dynamically changing headway by kinetic depth perception. At the heart of kinetic depth is motion parallax.

Motion parallax is relatively easy to demonstrate. Look at a distant object, something like a picture on the wall, close one eye, and hold up an index finger. Move your head sideways a couple of inches, slowly back and forth, so you can see the picture changing sides behind your finger. The distant picture seems to move in the same direction as your eye relative to your finger whereas your finger seems to move in the opposite direction of your eye relative to the picture. You have just demonstrated to yourself the difference between with and against motion, one of the most fundamental and important principles of geometric optics.[8] This is also the basis of motion parallax. When you move through a world that contains things at a variety of distances from you, their relative with and against motions signals their relative distances. This occurs even without you consciously being aware of these relative motions. In fact, when you move around, the objects in the world look stationary, these things don't look as though they are moving at all. It is hard to incorporate this important cue in a non-see through HMDs without incorporating some signal about the motion of the eye or head and using that signal to drive the display. If the display is in a vehicle, it may be necessary to integrate eye and/or head motion with vehicular motion to fabricate such parallax cues in a HMD.[9]

---

[8] Remember how the optics of the HMD sets its symbology at optical infinity so that there is no relative motion between the symbology and the distant world on which the symbology is superimposed? There is no relative motion because they are both at the same optical distance. This is exactly the same principle.

[9] Do kinetic depth cues contribute to simulator sickness in visual simulators?

Binocular depth cues

The binocular cues for depth perception have been extensively studied. These cues derive from the basic idea that at any one time an individual's two eyes have slightly different views of the world since they are displaced horizontally from each other.

Stereopsis

Stereopsis is the perception of depth specifically due to the relative spatial disparity, or difference, between the simultaneous images formed on each retina. The disparity between the two retinal images has a host of consequences that has been studied since well before stereopsis was formally discovered by Wheatstone (1838). The long history of active investigation has elaborated the basic notion of stereopsis in many important ways, making the idea proportionally complex, nonetheless a few points will be briefly mentioned here.

Stereopsis is an emergent property of the nervous system. It does not exist in the distant object being looked at, nor is it in either eye alone; it is created by the nervous system out of the information available to it from both eyes. It is something that the binocular visual system fabricates in its opportunistic use of available information.

Retinal disparity

Retinal disparity is key to understanding stereopsis; but disparity itself is a slippery concept that seems to acquire more definitions the more it is examined. When the two eyes are looking at the same object, each eye's line of sight is on that object. Normally, this means that two eyes are turned so that the image of the object is on each eye's fovea, which is an anatomical structure and a landmark on the retina. Simultaneously, every other object in the visual field is also imaged on each retina as well, but the distances between the fixated object and any of these other objects is different on the two retinas. In general, the further away an object's image is from the fovea, the greater is this difference on the two retinas. Yet when a single object's image falls on the two foveae, it is not surrounded by multiple images of its neighbors in the visual field. The two images on the separate retinas all merge into a unitary percept.

The basic idea is that there are corresponding locations on the two retinas. These retinal locations produce a single image when stimulated by the same object. But these corresponding locations are not points; they are areas and the size of these areas increase the further away they are from the fovea. Consequently, these areas of correspondence are defined by the singleness of vision rather than by any anatomical definition. In other words, corresponding retinal areas are not anatomically but functionally defined. The area around corresponding retinal points that produce a single image is called Panum's retinal area. Whenever an image falls on different points on the two retinas but are close enough to be fused into a single percept, those points are within Panum's area.

Corresponding retinal points, Panum's areas, can be back projected to map locations in visual space, which is one way of defining the so-called horopter: "The locus of object points in space simultaneously stimulating corresponding retinal points under given conditions of binocular fixation" (Cline, Hofstetter and Griffin 1980). The horopter is anchored on the fixation point, which is the projection of the fovea. Four typical horopters measured in a single observer are illustrated in Figure 12-20. The horopter is a curved line, defined by the overlapping visual fields of the two eyes and the corresponding points of the two retinas at an instant in time. An object falling on the horopter is seen as fused into a single image.

The horopter and the zone of single vision is one way of mapping or transforming the physical world into a visual space. This is one approach to this type of transformation that the visual system continually performs. The only component of this transformation that really is anchored to an anatomical landmark is fixation, defined normally by the fovea. All the other scaling factors used to define Panum's area, or a horopter, or a zone of binocular vision surrounding the horopter, depend on the details of the measurements. This means that the specific mapping of the physical to the visual depends on the types of psychophysical procedures used, the con-

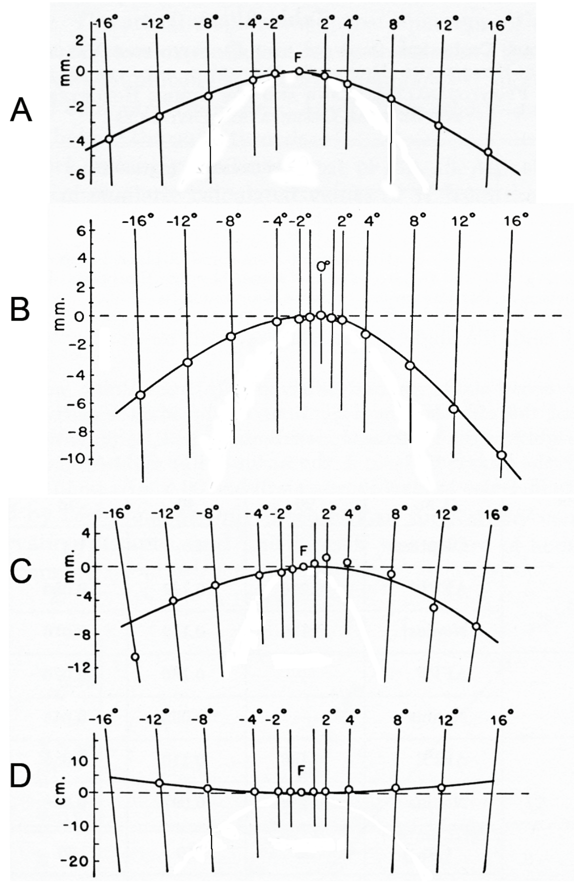


Figure 12-20. Four different horopters measured in the same individual. The difference among the four horopters is the distance at which the individual is focused. For horopters A through D, the focus distance is 20 cm, 40 cm, 76 cm, and 6 meters, respectively. The abscissa is the distance, in degrees, that the target is presented away from fixation. The ordinate is the perceived frontal plane.

figuration of the stimuli, and the distance between the eye and fixated target, to point out just a few such variables. Figure 12-20 shows that horoptors measured at different distances are different from each other. In other words, how we map the physical world to visual space depends on the methods we use to do the mapping.

The horopter is one way of illustrating a basic truth that should be fundamental to the design of displays, heads-up and otherwise. Physical space is not visual space, and visual space need not be Euclidian.[10] Yet deviations from Euclidian geometric mapping of the physical to the visual may contribute to several visual illusions.

Size perception and the constancies

Unless the visual conditions are arranged just right, we generally don't even notice the illusions. This is because we don't see retinal images; we see the objects that generate those retinal images. The person approaching us does not grow in size and get bigger despite the fact that the retinal image is growing. The person remains the same size; he/she is just getting closer. The window looks like a rectangular opening with straight edges and right-

---

[10] For a recent introduction to the literature on this important and complex issue, see Wagner, M., *The Geometries of Visual Space*, Lawrence Erlbaum Associates, Mahwah, NJ.

angled corners, regardless of where we stand when we look through it. The things we look at stay constant; they don't change. That is obvious; so obvious, in fact, that it is hard to understand why such constancy in perception is worth discussing in the first place. Obviously, it would make no sense for things to be organized any other way. After all, if we were bound to see only the retinal image rather than the object in the world, our visual world would change with every motion of the eye since every motion of the eye changes the retina image of the world. Visual perception doesn't work that way, fortunately. The open book lying on the table slightly to one side does not look like a misshapen parallelogram; it looks, exactly like what it is, an open book and to see it as a twisted parallelogram-type figure requires an immensely artificial mental act that is more analysis than perception. This constancy of an object's shape illustrates the constancies at work. Objects keep their size, shape, brightness, color, and so forth as we move around them, or they around us.

Which is the illusion, the distal objects we see in the world, which frankly look nothing like their images on the retina, or the retinal image that is all but invisible to us, completely obscured by the constancy of objects? The retinal image is fundamentally different from the distal image; but that difference is invisible to us who sense only the ever-changing retinal image but who see only the constant distal object. Which one is real and which the illusion?[11]

Size constancy

Size constancy is discussed briefly in Chapter 10 (*Visual Perception and Cognitive performance*) but is worth revisiting in the context of illusions. Certainly one of the most important studies of size constancy is that of Holway and Boring (1941), which has been discussed, replicated, analyzed, and argued about since it was first reported. For that study, a subject sat at the right angle juncture of two corridors so that the subject could look down only one corridor at a time (Figure 12-21). In one corridor, ten feet away from the subject, a white disk of light was projected. The diameter of this disk, the *response disk*, was adjustable. Along the other corridor other disks were presented at various distances out to 120 feet (36.6 meters), denoted *stimulus disks*. Each of these stimulus disks had a different, though constant, diameter. In fact, their diameters were directly proportional to their distance from the subject so that the diameters of all of the stimulus disks produced a constant 1° diameter size image on the retina. The task of the subject was to set the diameter of the response disk, at the constant 10 feet (3 meters), so that it matched the diameter of each of the stimulus disks along the corridor, out to 120 feet (3636 meters). The intensity of the light from the disks was adjusted so that the light was constant and equal at the eyes of the subject for all disks.

In general, one of two types of results can be predicted. On one hand, subjects might see the different diameters of the stimulus disks along the corridor exactly as they actually were; the further away the stimulus disk, the bigger is its diameter. In this case, the large-diameter stimulus disks at greater distance along the corridor would be matched by setting the diameter of the response disk to be large. Similarly, the smaller-diameter stimulus disks at the closer distances would be matched by setting the diameter of the response disk to be small. In this case, it would be as though the subjects were using a tape measure to set the diameters. This is evidence of size constancy. On the other hand, subjects might recognize the retinal image of each of these circles is the same size, 1°, and try to set them all to that size. In this case, the subjects would be responding to the retinal image rather than to the actual physical dimensions of the stimulus disks. This would be a matter of simple trigonometry, keeping the angle constant.

---

[11] The situation is more complicated than that because the object's image exists independently of the retinal image, not the other way around. Consequently, the more inclusive analysis should distinguish the object from its image and from the retinal projection of that image.

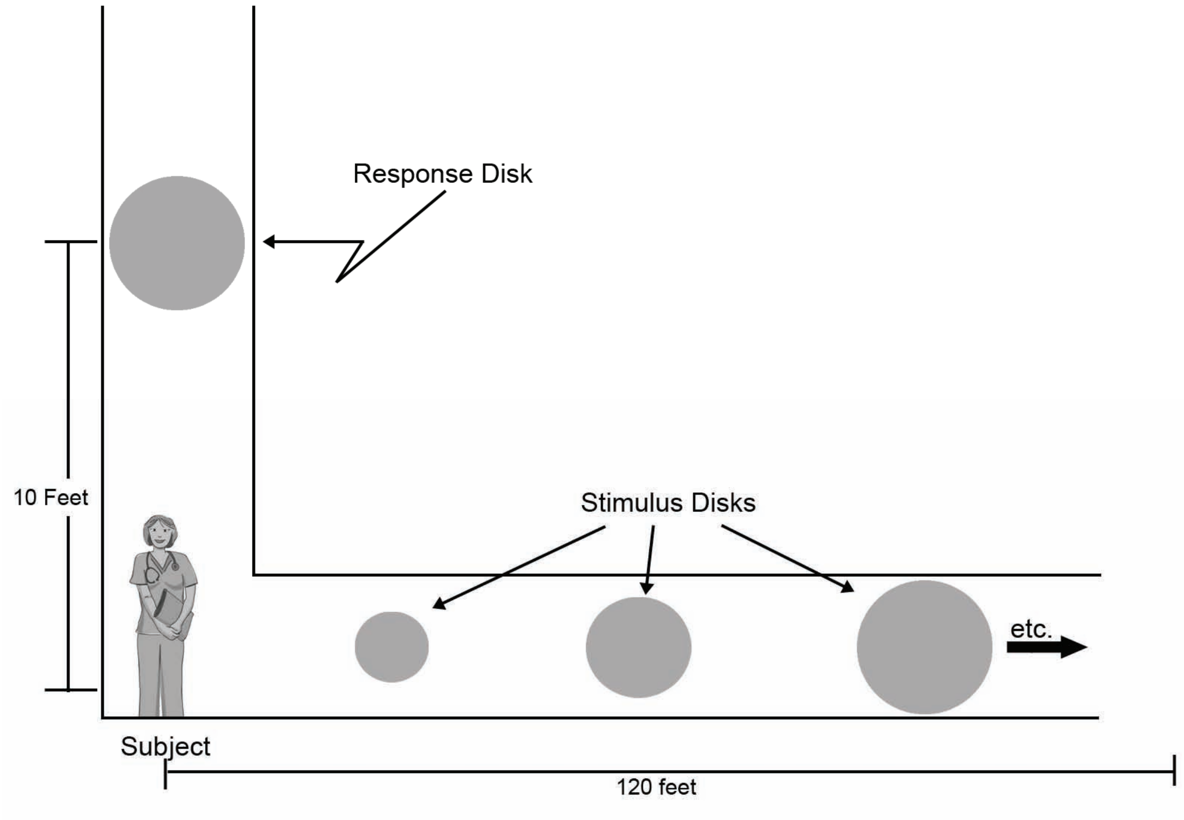


Figure 12-21. Setup for the Holway-Boring (1941) experiment in size constancy. The response disk was a constant 10 feet from the subject. There was a set of stimulus disks. These stimulus disks were placed at distances out to 120 feet from the subject. These stimulus disks were arranged so that they all had the same size diameter, 1.0°, on the subject's retina. The subject adjusted the size of the response disk to match the apparent size of the different stimulus disks under different conditions, as described.

Holway and Boring (1941) incorporated another factor in this study – the viewing conditions. Subjects viewed the disks under four conditions. For condition A, the subjects used both eyes to view the stimuli. For condition B, they used one eye. For condition C, the subject's one eye views the stimuli through a small hole, referred to as an artificial pupil. For condition D, the eye viewed the stimulus disk through the artificial pupil down a long black reduction tunnel that eliminated essentially most frames of reference as well as stray or ambient light. In other words, these four conditions produced progressively sparse visual environments.

The results are simultaneously straightforward yet profound. In condition A, the subjects matched the response disk diameter to the physical diameter of stimulus disk. The subjects, of which there were five, saw the diameter and distance of the stimulus disk and adjusted the response disk diameter on the basis of the stimulus disk's physical dimensions, as though they were using a ruler. In fact, the response diameters were a little larger, as if the subjects were cognitively trying to compensate for what they knew to be the influence of distance on perceived size. They saw the size of the stimulus disk and its distance, and did some sort of mental calculation. In condition D, on the other hand, all the disks were adjusted to approach the same retinal size, the constant 1°. In fact, a graph of response disk diameter as a function of stimulus disk distance showed a line with a very slight positive slope; so that the matches were not completely determined by the retinal image size alone; but, this deviation was small, suggesting that not all of the distance cues were completely eliminated. In other words, under the sparsest visual conditions, the size matching approached the size of the image projected onto the retina, but not totally. Conditions B, and C produced results intermediate between the two extremes of A and D; the more rich the visual configuration, the more the subject is able to recognize the diameter and distance of the stimulus disk. The richer the stimulus environment, the more the matches were determined by the physical diameter of the stimulus disk and less by the retinal projection. Conversely, the poorer the stimulus environment, the more the matches were determined by the retinal projections of the stimulus disk and less by the physical diameter.

This leads to several interesting questions regarding HMDs. What kind of visual environment confronts an operator using a HMD? How rich are the size and distance cues? Are they sufficient to support the size or distance constancies? For example, the Holway-Boring experiment was partly replicated with night vision goggles (NVGs) (Zalevski, Meehan and Hughes, 2001). The visual environment NVGs provide is not that of full daytime, but neither is the environment completely sparse. Its somewhere in between and the size judgments were consistent with that. The response disk diameters were not completely determined by the retinal image size or by the diameters of the stimulus disks; but were closer to the physical stimulus disk diameter than they were to the retinal size. The perceived sizes of an object seen with and without NVGs need not be the same. From this alone it can be expected that NVGs can affect judgments about apparent size, distance, or both. These results further suggest that it may be possible to develop a metric for the evaluation of different visual display technologies based on the extent to which they enable or degrade size constancy.

Shape, brightness, and color constancies

Shape, brightness, and color are some of the other constancies that contribute to our inability to see the retinal image. As demonstrated by the Holway-Boring experiment, the extent to which these constancies hold depends on the specifics of the visual stimuli. In other words, for any specific situation, the extent to which these constancies actually hold depends on the visual conditions produced by an HMD.

Underlying the logic of the Holway-Boring experiment is the simple geometry of the retinal image, which may be described as the size-distance invariance hypothesis. The ratio of an object's size to its distance defines geometrically the retinal image size of that object. The geometry of this relationship is not hypothetical; it is trigonometric. But the dependence of the perceived size of the object on this size/distance ratio is hypothetical. The whole point of the Holway-Boring experiment and its many subsequent replications (and precursors) is that the perceived size of an object need not be determined solely by the geometry that defines the retinal image size. The point is that the perceived size, as well as the perceived distance, of an object is only partly determined by the retinal image. In fact, the conditions in which retinal image size is the determining factor are extremely artificial and difficult to set up. Consequently, the importance of the retinal image size in determining the perception of an object's size is rather small.

The logic and the shortcomings of the size/distance hypothesis illustrated by the Holway-Boring experiment is analogous to the shape/slant invariance hypothesis; that a retinal projection of a given form and size determines a unique relation of apparent shape to apparent slant.[12] Again, the relationship between the slant and shape depends upon the specifics of the stimulus field. At night, with little or no moon, the landing field looks like a trapezoid or parallelogram; during the day, it looks like a landing field rather than a geometric figure.

The same logic applies to the color of an object, which is another of an object's constancies. Severe disorientation would ensue if an object radically changed its color every time the lighting conditions change. Lighting changes commonly are used in theatre for dramatic effect but the colors of the objects usually do not appear to change; they appear to remain the same.

---

[12] For an excellent review see Sedgwick, H. A.: (1988) Space perception. In: Boff, K., Kaufman, L., and Thomas, J. (Eds.), *Handbook of Perception and Human Performance,* (Chapter 21, pp 1-57). New York: Wiley.

Static Illusions

Visual illusions are more typically associated with geometric illusions of form or shape than with the various constancies described above. Common visual illusions typically refer to geometric illusions, ambiguous figures, illusory contours, and impossible figures. Such illusions collectively are referred to as *static* illusions.

Geometric illusions

The term "geometric illusions" refers usually to any of a class of illusions that occurs in line drawings (Robinson, 1998). These geometric illusions may be among the most commonly discussed visual illusions possibly because they are so easily illustrated. Figure 12-22 shows one version of the classic Muller-Lyer illusion. Figure 12-23 shows a few other less common geometric illusions, including the Oppel-Kundt illusion (also referred to as the *filled-space* illusion), which is particularly important historically. Oppel reported this illusion in 1855, in the first formal scientific investigation of this class of visual phenomena, coining the phrase "geometrisch-optische Tauschung" (translated as "geometrical-optical exchanges") (Coren and Girgus, 1978). Since then, thousands of such graphic illustrations may have been created, with possibly nearly as many scientific papers and reports discussing them. The hope of discovering some parsimonious organization for the large universe of fascinating graphics along with simplifying or unifying explanations has been behind much of the interest and research in these geometric optical illusions. Much of the current research in this area is informed by contemporary neurophysiology and electrophysiology of the visual system and by the initiative of artificial vision.[13] According to Robinson (1998), who has provided an excellent review of the field from Oppel through to the early 1970s, "It would not be too bold to claim that stimuli in the visual field almost always interact, especially if they are close together or concurrent. Thus, judgments of the degree of separation and the orientation of lines or areas are influenced by the degree of separation and orientation of other lines or areas in the visual field, especially if they are close by. This makes it easy to invent variations of illusion figures once one has appreciated the essential configuration that gives rise to the illusions." This suggests not only that there may be an unbounded set of such illusions, but that there are certain common themes or methods by which they function. Robinson suggests three general factors, the specifics of which may differ from instance to instance.

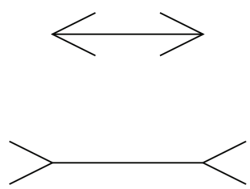


Figure 12-22. The classic Muller-Lyer geometric static illusion in which the distance between the left and right tips of arrows is identical between the upper figure and the lower one, but appears to be longer in the latter than in the former.

The first, and possibly the most important, factor is the role of ambiguity in the illusion. The information in the graphic is just not adequate. This causes perception to vacillate among the possibilities, unable to settle on the real situation. Line drawings specifically work because they evoke rather than delimit. The second factor Robinson proposes is that the illusions evoke processes that normally lead to definitive perceptions, but the illusions fail to provide the closure necessary for a definitive percept. For example, Gregory (1996) has argued that these illusions, like the Miller-Lyer illusion, may engage perceptual processes that encode size and distance but with inadequate and indefinite stimuli. The third factor is what Robinson refers to as the visual system's inability to

[13] The theme of much of this research is to look for physiological functions that seem to mirror the perceptual phenomena. The temptation is to interpret the correlation as an explanation, an approach which has well known pitfalls.

cope with certain input. He uses blurring as an example. To my way of thinking, these all come down to the ambiguity resulting from the sparseness of the stimulus conditions. These are the same factors or influences that underlie the various constancies discussed above; the sparseness or inadequacy of the distal stimulus or visual field result in increased perceptual ambiguities. These are the conditions that confront real operators controlling real vehicles in the real world under conditions of poor visibility.

Figure 12-23. Three different versions of the Oppel-Kundt illusion. For each version, the filled and the unfilled spaces are the same size.

Ambiguous or reversible figures

Ambiguous or reversible figures traditionally have been differentiated from "geometrisch-optische Tauschungen." The most famous of these is the Necker cube, which is illustrated in Figure 12-24 (with two additional reversible figure illusions shown in Figure 12-25 – the Mach's book and Rubin's vase-face illusions).

According to Boring (1942), in 1832, L. A. Necker a Swiss naturalist studying crystals, noted the ambiguous reversible nature of the two-dimension drawing that bears his name. These figures, like the optical illusions described above, have been extensively studied. Some of this work has been reviewed recently (Long and Toppino, 2004). A couple of points should be made about these ambiguous figures.

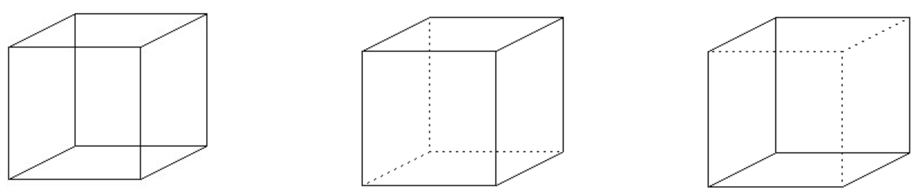


Figure 12-24. The Necker cube (left) and two possible interpretations (middle and right),

Figure 12-25. Mach's book (left) which may be seen as an open book with pages facing you, or as the covers of a book, with the spine facing you, and Rubin's vase-face illusion (right), which may be perceived as a white goblet in front of the background or the two black profiles in front of the white background.

The ambiguity contained in the reversible figures causes the perception to vacillate among well-defined alternatives, usually just two. This is a different situation than the one posed by the geometric illusions where the perception is not set by alternatives but involves a range of indecision. Underlying the continued interest in these reversible figures is an assumption that is often not made explicit. The alternation in perception among the well-defined alternatives reflects the activity of some set of neural processes operative within the viewer. This in turn implies a form of psychophysical isomorphism, which presupposes a view of the relation between the physical and the psychological. A more general model of psychophysical isomorphism assumes that point-to-point relationships in the neuro-sensory systems are preserved in the sensory-psychophysical systems. The analogous model invoked by ambiguous figures can be identified as phenomenal isomorphism, the alternative perceptions of the ambiguous figure are different sensory-psychophysical phenomena that derive from different neuro-sensory systems. Since there are two percepts with common elements, there are two neuro-sensory systems that also have common elements.

Since the Necker cube is a flat projection of a 3-D object, depth is implied in the image. Some researchers, like Gregory, point out that with depth, come expectations of distance and size relations; an object should look bigger when it is near than when it is far; but the opposing faces of the Necker cube are the same size, a violation of these non-conscious expectations, which, in turn contributes to the instability of the perception.

The perceptual ambiguity of some of these figures results in two different views of the same object, for example, the Necker cube remains a cube and the Mach book remains a book. That is different from the vase/face (Figure 12-25) or the duck/rabbit figures (Figure 12-26). These structures alternate between two completely different percepts; each of which is itself complete and unambiguous. In the face/vase illusion, for example, there is really nothing missing in either of the profiles facing each other; each is a complete profile. Viewing the two profiles relegates the space between them to the background. The background too is unambiguous and complete; it does not contain anything. Similarly, when the vase is in view, it is complete; there is nothing ambiguous about it. Nor is there anything ambiguous about the space around it; the space is merely the background. The profiles cease to exist as profiles to become the background. There is depth implied in this figure, to be sure; but depth seems to play less of a role in this face/vase illusion than it does in the Necker cube. Depth cues may play an even smaller role in the wife/mother-in-law (Figure 12-27) or the duck/rabbit reversals. In other words, ambiguities about depth and other cues seem to play different roles in these different figures; these ambiguous figures do not all work the same way. Just because the phenomenology seems similar among these reversible figures, an alternation between two unstable percepts, does not necessarily mean the same neural systems are responsible for the phenomena. This suggests further that the visual processes involved with these different figures may well be different, and conversely, different neural systems may result in apparently similar phenomenology. Some of these systems may be very early in the visual processing while other may be very late, the former having very little cognitive contributions while the latter may be more cognitive and less bound by the stimulus. This distinction is usually referred to at bottom-up or top-down, respectively.

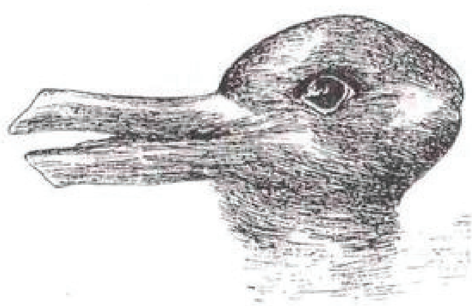


Figure 12 -26. Duck Rabbit illusion.

Figure 12-27. The young/old woman illusion.

These illusions have another important characteristic. The figures support alternative perceptions, each of which is well-structured and complete. The visual system does not need to fill-in or supply missing graphical elements in order to complete the picture. There are no missing graphical elements in these percepts. The ambiguity in the figures does not reside in the different percepts; the ambiguity resides in the figure itself; it supports at least one too many complete percepts. This is a completely different type of ambiguity than that contained in the geometrical-optical illusions, which hardly support any, and the illusory figures generated by the illusory lines and edges discussed next.

Illusory contours

Illusory figures built on illusory contours are another class of illusions. Since the publication of a collection of essays by Gaetano Kanizsa in English in 1979, there has been an increasing amount of interest and research in this family of illusions. There is nothing ambiguous about these compelling illusions. The images are clearly visible; it is impossible not to see them; they just are not actually there.

Figure 12-28 (left) is a common example (Kanizsa triangle), a white triangle, whose apices partly obscure the black disks. Well, the disks are not really disks; they are disks missing a wedge, reminiscent of Packman figures from an early video game. The illusion is that they are disks, or more precisely, that is a part of the illusion. The other part is that the obscuration is caused by a triangle, a unitary, compete, easily seen, simple, geometric shape. The edges of that triangle appear to extend out beyond the missing pie-segment of the three Packmen. The side of the triangle is compellingly but only apparently defined by an edge or line, at least for part of the distance between the Packmen. But the edge is not there. More than that, the white triangle appears brighter than the white page outside the triangle, on the other side of the illusory line. But that’s not true either. The brightness is the same. There is no difference. In addition, the illusion conjures the perception of depth or at least multiple layers. The triangle is superimposed on black circles that are themselves placed or printed on the page. So, there are several components to this illusion. (1) There is the sense of boundary or edge where there is none; (2) There is an impression of a surface or geometric figure where there is none; (3) There is the impression of a difference in brightness between the inside and outside of the figure where there is none; (4) The inducing elements, the Packmen figures, are seen as something they are not, circles; (5) there is a sense of depth stratification, and that too is an illusion.

The black and white elements of Figure 12-28 (left) are reversed in Figure 12-28 (right), which reverses the contrast relationship in the illusion, creating a sharply stark, blacker-than-black triangle. All the relationships and illusory elements described in the previous paragraph apply to this figure but in reverse contrast.

Figure 12-29 (right) is historically noteworthy; Schumann reported it in the first scientific paper to consider such figures. “… one can see that in the middle, a white rectangle with sharply defined contours appears, which objectively are not there. However, under appropriate conditions, I have only succeeded in inducing straight lines and never regularly curved ones” (Schumann, 1987). However, as demonstrated in Figure 12-29 (left), there is really no particular difficulty generating curved illusory figures.

Figure 12-28. Two Kanizsa triangles; the one on the right is a contrast reversal of the one on the left.

Figure 12-29. (Left) Illusory figures built on curved illusory contours; (Right) The first of this class of illusory figures to be reported in the literature (Schumann, 1987).

There exists currently a large literature on this type of illusions, which includes a great deal of discussion concerning the necessary and sufficient conditions for these illusions and their underling causes. The issues raised are really quite complicated, far exceeding the scope of the present review; but there is one more point that needs to be made about them, particularly in the context of see-through HMDs and superimposed, transparent HUD displays. Can these illusory edges and the figures occur with HMDs, either intentionally as design elements, or accidentally, interacting in the see-through fashion, superimposing symbology on the scene? These illusory contours and figures are not far removed from perceptions involving transparency, as Figure 12-30 suggests. It may be suspected that they will be increasingly important as HMD technology develops.

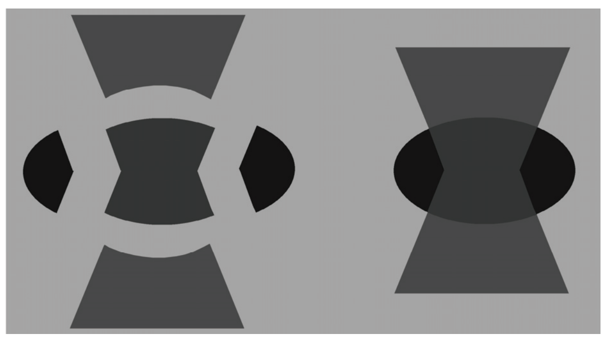


Figure 12-30. The elements on the left can combine to produce the illusion of transparency (Kanizsa, 1979).

Impossible figures

Impossible figures (objects) belong to a final class of static visual illusions to be discussed here and are distinct from those discussed above. These optical illusions, e.g., the Devil's tuning fork (Figure 12-31, left) or the Freemish crate (Figure 12-31, right), are not so much a visual illusion as they are unambiguous, explicit depictions of physical impossibilities. They live in a middle ground between perception and logic, taxing both. Many of the figures and illustrations by Escher work on this principle. These may be better described as illusions of higher order cognition than of perception.

This distinction is not to minimize their importance by any means. It may be that some episodes of spatial disorientation (SD) in aircraft are analogous to these cognitive illusions. One classification of SD distinguishes between instances when individuals recognize that they are disoriented from instances in which the SD goes unrecognized. When the individual recognizes the SD, the challenge is to reconcile two different and mutually exclusive visions of reality. The aviator struggles to figure out how to accomplish this. It is a cognitive problem; the two sources of information just do not fit together. This is very much the experience of fitting the Devil's Tuning Fork into a single percept; it just doesn't fit.

Figure 12-31. Examples of impossible figures: Devil's tuning fork (left) and the Freemish crate (right),

Locus of the illusions

While this is not the place to join in the ongoing academic discussions aimed at clarifying the various visual or cognitive processes underlying the different static visual illusions discussed above, a generalization does seem clear; no one illusion depends on only a single mechanism. Every one of them seems to be multi-determined. Coren and Girgus (1978) convincingly summarize evidence that any one visual illusion involve multiple cascaded processes. For example, the physical principles of the geometric optics describing image formation with light can contribute to the Miller-Lyer illusion. This has nothing to do with any neural processes and all to do with the way an optical system bends light when forming optical image. This occurs as the optics of the eye forms the image on the retina. Then come the processes involving neural crosstalk within the retina before any information leaves the eye. Then there is the analysis added when the information from the two eyes come together, which occurs at various levels through the central nervous system. In addition there are the higher order cognitive effects. This describes a bottom-up version of the system. The top-down version emphasizes the importance of expectation, set, reason, and other cognitive functions on the illusion. This dichotomy is simplistic; regardless of which direction is selected, bottom-up or top-down, recursive or feedback loops appear very quickly.

Space perception

Some may question the practical importance or relevance of these visual illusions; are they anything more than mere curiosities? The position taken here is that these visual illusions are central to the depiction of space and the perception of the relative position of objects that populate the navigable space. Illusions are endemic in the experience of the real three dimensional (3-D) world because the geometric optics of each eye projects onto its retina a planar rendering of the 3-D distal stimulus field. Humans, with two eyes, have a pair of simultaneous, correlated, two-dimensional (2-D) representations of the world – one in each eye. Most of the time, the human visual system successfully isolates the individual from these illusions. Occasionally, they occur in daily life, particularly when the stimulus field becomes sparse; but for the most part, humans don't have to cope with these retinal images.

The ability to represent the three dimensions of the physical world onto two dimensions, the goal of all virtual reality, synthetic vision, and conformal displays, depends totally on the judicious use of the types of visual illusions described above. Such fabrications of three dimensions on a 2-D surface involve some mapping algorithms or transformations as well as assumptions that are either implicitly or explicitly made, but made nonetheless. For example the size/distance or slant/shape invariance hypotheses may be assumed naively without question, simply because the geometry is so appealingly simple. But these assumptions need to be tempered by the various perceptual constancies and the situations that provoke their breakdown. It is becoming increasingly clear that the rendering of three dimensions of reality onto a two dimension surface, even with the tricks of pseudo-depth, will invariably involve confusions and ambiguities. At a very minimum, they will incorporate the confusions and ambiguities that are inherent in the real world. This is evidenced by the growing volume of human factors research on the relative strengths and weaknesses of various perspective or 3-D displays, some of which is discussed by St. John et al. (2001), as well as in Chapters 2, *The Human-Machine Interface Challenge*, and Chapter 10, *Visual Perception and Cognitive Performance.*

One of the deepest conundrums of visual perception is mapping the dimensions of physical space into a spatial representation of visual spatial experience. The 3-D geometry of the visual world is transformed into the planar geometry of the retina. One of the themes of this discussion is that at any one moment the image on the retina of the distal visual stimulus field is highly confusing, ambiguous, and complicated. Our neural-visual machinery is designed to take apart and analyze that confusing, ambiguous, and complicated surface rendering. Separate, simultaneous systems tuned to specific aspects of the image perform these multiple simultaneous analyses, splaying the image apart in different regions of the brain. Color information processed in regions A, edge information processed in regions B, occulo-motor information in regions C, visual disparities in regions D, and so on, all at roughly the same time, and all these regions overlapping or sharing information to some extent. With such a complicated ensemble of cross-talking (leaky) parallel systems, each analyzing particular pieces of the visual puzzle, why should there be one mapping of physical into visual space? And, if there is more than one, how many are there? Are they all equally important? What is their relative importance to a specific task, be it perceptual or motor? And, how would these different mappings be accommodated to HMDs, either see-through or otherwise?

Some vision researchers have explicitly argued that the visual system incorporates multiple simultaneous mappings of the space around us. For example, the earlier discussion of the horoptor described one approach to mapping equal perceptual distances based on fusional areas, the regions that produce single vision. Another approach is mapping regions of the visual field that have equal sensitivity to such specified stimulus parameters as luminance, color or motion. This technique, *perimetry*, is common to the eye clinic (Aulhorn and Harms, 1972). Yet another approach is to equalize or re-scale the visual field in terms of acuity (Anstis, 1974) or cortical magnification (Crowey and Rolls, 1974), size and distance judgments (Wagner, 2004), or any of a number of other specific visual functions (MacLeod and Willen, 1995). The number of different approaches to providing a visual representation of the physical is large. Mapping for one dimension may violate mappings in other dimensions, which could produce confusions and misjudgments along these dimensions.

In a 2-D representation of 3-D space, distance perception of necessity is confused and confusing. Objects get smaller as they get further away but there is a catch, perspective. The effect is not obvious, even textbooks on perception, in illustrations and discussions of depth perception have gotten it wrong (Gillam, 1981). Smallman, Manes and Cowen (2003) state, “It is not widely appreciated, even among vision researchers, that projected width across a scene (X) and projected depth into a scene (Y) taper differently with distance in to the scene. Projected width is inversely proportional to distance. Projected depth (Y) on the other hand is inversely proportional to the square of distance because of foreshortening.” These different relationships certainly complicate the perception and understanding of size and distance information represented in graphical displays. They also complicate the perception and understanding of size and distance information represented in the real world; these are not noticed, because, perception is itself an illusion. Creating displays that mimic or emulate the real world may build these illusions into the display and produce much the same effects.

Dynamic Illusions

As with static visual illusions, dynamic illusions are constantly present. The success of visual information display technologies, including head-mounted and virtual ones, will be better served by understanding that visual perceptions inevitably involve some form of illusions and raises challenges for defining the criteria and desiderata for such displays. A survey of U.S. Army AH-64 Apache helicopter accidents, reported for the period from 1985 to 2002, concluded that dynamic illusions are particularly important when using the Apache’s HMD, the Integrated Helmet and Display Sighting System (IHADSS) (Rash et al., 2003). Of the 228 reported accidents, approximately 93 (41%) involved the HMD in some way, and for 21 of these, the HMD and pilotage night vision sensor system played a role in the accident sequence itself. Furthermore, the most frequent causal factor in all of the accidents studied was the presence of dynamic (motion-based) illusions, which were identified as *disorientation* (14%), *illusory dri*ft (24%), *faulty closure judgment* (10%), and *undetected drift* (24%). The relatively important role of dynamic illusions reported in these accidents suggests that the illusions associated with motion perception warrant special attention. The survey reported that the second most frequent cause was *degraded vision* (i.e., reduced resolution and contrast). This is consistent with the arguments made in the earlier discussion of static illusions, i.e., the more sparse or degraded the visual stimulus field, the more pronounced are the illusory percepts. Nevertheless, the absence of an accident does not mean that the pilot had no visual illusions, since pilots routinely and successfully control aircraft in the presence of multiple visual illusions.

What is motion perception?

Understanding dynamic illusions requires some understanding of motion perception[14] and, of course, at least some of the more common illusions associated with motion. An interesting question is why the perception of motion poses any special perceptual problems. When an object moves in the visual field, such as an automobile passing in front of the eye, the image of the object flows over the retina. It might seem obvious that the movement of the car's image over the retina should produce the perception of the car moving through a static environment. Or, consider a situation in which an individual tracks a passing car with their gaze. In this case, eye movement keeps the image of the car relatively stable on the retina, but the image of the rest of the world around the car is moving. The image of the static environment moves over the retina as the retina moves to keep the image of the car relatively stable. In both cases, there is still differential motion between the images of the moving car and the static environment. Since humans have the perception of motion when an image of an object courses over the retina, it may not only seem a strange but even an unjustified violation of parsimony to ague that there may exist a special system responsible for motion perception.

Consider the passenger in the car, such that the image of the world through which the passenger is passing is visible around him through the windows. As the passenger looks through the windshield, vehicular structures, such as its hood, dashboard, the spots on the transparent glass of the windshield, are all approximately stationary relative to him and the moving world outside. As the car travels, the visual system effortlessly disambiguates the complex patterns of differential motions on the retina. But, as impressive as this accomplishment is, since all these motions are associated with streaming objects that have specific and constant identities, one is tempted still to not be totally convinced that it is necessary to postulate a special system responsible for the perception of motion. Let's say that the driver stops the car at a red light, and the passenger turns his head to look at the driver. Another car pulls to a stop the adjacent lane, on the diver's side, filling the passenger's view of the world behind the diver. As the passenger views the driver, he suddenly perceives the car he is in begin to roll slowly backwards and in a reflective reaction turns quickly to look out the back window to check that the car he is in is not going to roll backward into a car to the rear. However, he quickly ascertain see that his car is not moving at all, and the driver still has a foot on the brake. What really happened is that the light has turned green and the car, which had had seen as stationary behind the driver, in fact had begun gradually to pull forward while his car remained stationary. As the passenger was attending to the driver, some part of the visual system registered the motion of the car visible behind him. That situation produced a strong, compelling sense of motion, even though the passenger and the car he is in were not moving.[15]

Let's examine another situation. An individual arrive at a movie house and spend the next two hours watching a film. Several hundred million dollars were spent and thousands of people worked on the creation of this extended, two-hour illusion. There were no real actors moving in front of the moviegoer, as actually occurs in a stage-

[14] See Chapter 10, *Visual Perception and Cognitive Performance*, for an expanded discussion on motion perception.
[15] This sensation of movement of the self in space produced purely by visual stimulation is called *vection*.

theater. Rather, he just spent two hours watching a sequence of two-dimensional distributions of variously-colored light. The whole experience is conjured.

At the end of the show the credits appear to "roll" by. They are read as if they "scroll upward."[16] The movie often ends with a final stationary image, e.g., production company logo or a final message. For a moment, it may be perceived as a curious fact that this stationary image seems to move in the opposite ("downward") direction.

Motion perception and Gestalt psychology

Boring (1942) noted that Perkinje's studies of motion sickness and vertigo from 1820 through 1827 may mark the beginning of the scientific study of motion perception and illusions. By 1831, the physicist Michael Faraday had described stroboscopic motion, and the word stroboscope had come into use by 1833. Interestingly, many of the devices and effects used to generate such motion perceptions were created for parlor amusement and entertainment. In other words, entertainment drove much of these 19th century technological innovations in much the same way that entertainment is driving display technologies today.

During this same period, Addams (1834) published, "An Account of a Peculiar Optical Phaenomenon Seen after Having Looked at a Moving Body." He was reporting a motion aftereffect sometimes called the *waterfall illusion*. After watching the water coursing downward in a waterfall for a few minutes, such nearby stationary objects as rocks, grass, trees, etc., appear to move in the opposite direction. This illusion is what occurs with the previously described credits at the end of the film. Over the years, many researchers (including Helmholtz [1909]) have argued incorrectly that such motion aftereffects (MAEs) were due to eye movements. Others, including Ernst Mach, have argued that eye movements cannot account for MAEs. For example, if one eye looks at a pair of spirals simultaneously rotating in opposite directions, the eye will have simultaneous MAEs in different circular directions; this cannot be accounted for by eye movements. Mach and others argued that MAEs must reflect the operation of some kind of neural-retinal mechanism(s). Certainly, by the 1870s some vision scientists had begun to argue that motion *per se* was a basic visual, if not retinal, process. Exner (1875) studied the perception of two electric sparks separated in time and distance. He found that when the pair of sparks was flashed with a delay of 45 ms or longer, the order of illumination could be detected. When the sparks were moved closer together, their sequential illumination provoked a perception of motion that was seen even with a delay as brief as 14 ms. In other words, motion was seen even when the time interval between the two sparks was too brief to determine the order in which they were illuminated. The argument developed that motion does not depend upon an object changing its location over time. Exner (1875) concluded that motion was a sensation rather than a perception.[17]

This argument of motion has special importance historically; it is the very root of Gestalt psychology. In Wertheimer's 1912 paper on his studies of apparent motion, specifically on something called *phi* motion, he extended the arguments that Mach and Exner made almost thirty years earlier. He emphasized that such motion perception does not depend on an object or its identity, and that motion perception is not a derived reality but it is a basic phenomenon that reflects cortical processes. He further argued that the veracity of the perception of motion is not dependent on the real motion of a real object but depends on what happens in the brain. These arguments quickly led to a number of studies of differing types of motion, as the Gestaltists differentiated a number of motion types. Table 12-1 is based on Boring's summary and lists a number of these motions and their characteristics.

---

[16] Of course, the credits aren't really scrolling, as that too is an illusion resulting from the rapid sequential presentation of static images, just like the rest of the film.

[17] Although the distinction between a sensation and perception is important historically but less clear today, it is still embodied in the dichotomy between top-down and bottom-up processing. Sensations reflect the atomistic, basic, raw function of a sensory system; while perception is the resultant organization of the basic building blocks into a neural event that is more than the sum of the sensations. A percept is the organization of the set of sensations. But, these distinctions are all just words that beg careful definitions.

Table 12-1.
Types of apparent motion identified by the early Gestaltists
(adapted from Boring, 1942)

Motion type	Year	Author	Characteristics
Alpha	1913	Kenkel (1913)	The apparent expansion and contraction of the Muller-Lyer central line with the sequential presentation of the illusion's components
Beta	1913	Kenkel (1913)	The apparent motion of an object with its sequential presentation at two different locations
Gamma	1913	Kenkel (1913)	The apparent expansion and contraction of an object when rapidly made brighter or dimmer
Delta	1915	Korte (1915)	The apparent motion when the second stimulus is made brighter than the first and the second is seen as occurring first
Phi	1912	Wertheimer (1912)	Stroboscopic motion: the appearance of motion without a moving object
Bow	1916	Benussi (1916)	The impression that a stimulus follows a curved path can occur when a pair of successive flashes presents a stimulus on either side of an obstacle.
Split	1926	DeSilva (1926)	The apparent splitting of a vertical line into the left and right components of its perpendicular when presented sequentially

It should be noted that at about this same time, Korte (1915) described some qualitative relationships among the luminance (*l*), duration (*d*), inter-stimulus interval (*isi*), and space (*s*) that specifically apply to beta motion. These generalizations are sometimes referred to as Korte's Laws and are stated as:

- With constant *isi* and *d*, optimal apparent movement can be maintained with an increase in both *s* and *l*.
- With constant *s* and *d*, optimal apparent motion can be maintained with a decrease in *l* and an increase in *isi*.
- With constant *l* and *d*, optimal apparent motion can be maintained with an increase in both *s* and *isi*.
- With constant *l* and *s*, optimal apparent motion can be maintained with a decrease in *d* and an increase in *isi*.

Illusory motion

The above discussion of apparent motion emphasized the development of the idea that motion perception is not a derived secondary stimulus characteristic but a basic dimension of the visual system. This idea originally rested on evidence from phenomenology and psychophysics but continues to find corroboration in contemporary electrophysiology and neuroanatomy. It is unfortunate that the term 'apparent' motion has been so widely used, because it is imprecise. All perceived motion, whether real or illusory, is apparent. Illusory motion would have been a more accurate term, referring to a situation in which motion is seen but is not actually present

Regardless of the word employed, apparent or illusory, the perception of motion and the processing of motion information now are known to depend upon the function of basic processes in the visual system. Furthermore, the study of the dependence of the percept on stimulus conditions reveals how these systems work. The emphasis on apparent motion begs the further question of whether there are differences between illusory and real motion. This broad question probably only really makes sense in the context of specific cases, since there are so many types of such motion percepts.

The present discussion will address five different types of illusory motion: (1) stroboscopic motion, (2) the autokinetic effect, (3) induced motion, (4) the Pulfrich phenomenon, and (5) motion aftereffects. Each of these types illustrates the operation of different processes involved with motion perception. Together, they illustrate the range of phenomena that can provoke illusory motion. Furthermore, not only do these five phenomena occur in increasingly complex stimulus fields; but each occurs in the real world and affects visually-dependent performance, at least in the context of aviation in the real world.

Stroboscopic motion

As Wertheimer and his colleagues were studying illusory motion and building their arguments about Gestalt psychology, more pragmatically-oriented individuals were developing techniques to use the same phenomena to make moving pictures. The rapid sequential presentation of stationary images produces stroboscopic motion, also called the phi phenomenon (or beta motion).[18] It is extraordinary how endemic this type of illusory stroboscopic motion is in modern society, whereas it was a curiosity in the mid-19th century.

Possibly the simplest example of the basic experimental paradigm may involve a pair of easily observed, identical, short flashes of lights, F1 and F2, presented one right after another, each with a relatively short, well-defined, constant duration (e.g., 250 ms). The interval between the offset of F1 and the onset of F2 is typically referred to as the inter-stimulus interval (ISI) and in this simple example may be set equal to the flash duration (i.e., 250 ms). When F1 and F2 are presented at the same location but separated in time by the 250 ms ISI, the percept is relatively simple: A light comes on, goes off, comes on, goes off, etc., all in the same location. Since F1 and F2 are identical and in the same location, the observer actually perceives only one flashing light.

The situation becomes more interesting when F1 and F2 are spatially separated (e.g., a few degrees) (Figure 12-32), so that one is to the side of the other, e.g., F2 is to the right of F1. With the ISI still at 250 ms, the percept is unambiguously that of two different lights presented in succession at two different locations. When the ISI is very short (e.g., 30 ms), F1 and F2 appear as two simultaneous flashes next to each other; the ISI is too short for the visual system to detect. Clearly, the percept depends on the duration of the ISI. The interesting issue is how the percept changes as a function of the ISI. Specifically, in the ISI range between the 30 ms and 250 ms, there is a range of ISIs that produces a very strong perception of motion between F1 and F2. It is as though F1 jumps clearly and unambiguously to F2. F1 and F2 are not seen as two separate flashes, either simultaneously or successively presented; rather, they are seen as a single flash that moves from one place to another. The ISI that gives the strongest perception of such motion depends upon the spatial distance between the flashes, their luminance, duration, and size, exactly as described by Korte's laws. For example, in the situation described here, with F1 and F2 separated by a few degrees, an ISI of about 60 ms would a produce strong definitive, unambiguous perception of motion.

[18] The *phi phenomenon* is a perceptual illusion described by Max Wertheimer by which a perception of motion is produced by a succession of still images.

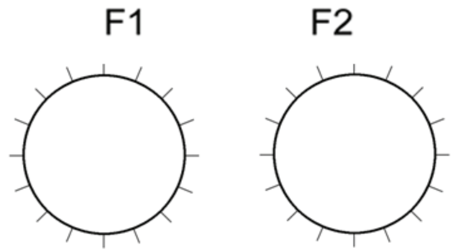


Figure 12-32. A pair of flashing lights (F1 and F2) separated by a few degrees of visual angles.

Ternus motion

Under some conditions, F1 and F2 are seen as different flashes, while under other conditions, they are seen as the same flash. One very important and influential elaboration of it, sometimes referred to as the Ternus display paradigm (Ternus, 1926), has been used for nearly a century to study ‘phenomenal identity.’

Ternus’ simplest case also uses two flashes but with the difference that each flash presents three dots. Figure 12-33 (left) shows F1 with its dots A, B, and C; on the right is F2 with its dots B', C', and D'[19]. The important issue is the spatial relationship among the dots between the two flashes. Dots B and B' are presented in the same retinal position and are indistinguishable. Similarly, dots C and C' also share the same retinal location and are indistinguishable. Therefore, the difference between F1 and F2 is that F1 contains dot A and F2 contains dot D'. However, this difference is not as simple as it may appear. F1 contains a dot to the left of dot B and no dot to the right of dot C; whereas, F2 contains no dot to the left of dot B' but does contain a dot to the right of dot C '.

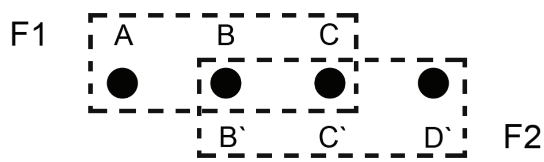


Figure 12-33. The Ternus display. Flash 1 contains dots A, B, C. Flash 2 contains B’, C’, D’. Note that dots B and B’ are in the same location as are dots C and C’.

When F1 and F2 are presented in succession, dots B and C of F1 are unambiguously in the same locations as are dots B’ and C’ of F2. The important question concerns the perception of dots A and D’. The answer depends on the multiple factors described by Korte’s laws, but the present discussion will address only ISI with the other factors remaining constant. For a longer ISI (typically 250 ms) the dots in the successive presentation of F1 and F2 will appear to shift to the right as a single group of three dots. In other words, A becomes B', B becomes C', and C becomes D'; the individual identities of the dots are lost despite the fact that B and B' are identical and fall on the retinal area they have in common; C and C' are also identical and fall on the retinal area they have in common. The visual system ignores their individual identity, submerging it into the group of three, and moving the group as a single unit.

For shorter ISIs, the result is completely different. B and B', as well as C and C', retain their individual identities but at the cost of the individuality of A and D'. In this situation, B and B' seem to be the same dot, just blinking on and off; also C and C' appear to be the same dot blinking on and off. Furthermore, dots A and D' seem to be the same dot jumping back and forth, over two intermediate flashing dots. The percept is clear and unambiguous. Consequently, the perception depends on the temporal interval separating F1 and F2. The

[19] It makes no difference whether these dots are dark dots against a light background or light dots against a dark background.

perception is dichotomous; it is either the motion of the group of dots as a whole or the motion of one dot hopping over a pair of stationary dots that seem to be simultaneously flashing on and off. Korte's laws determine which of these two perceptions is seen.

Ternus extended this paradigm to assess the extent to which an item (dot) retains its identity in the context of ensembles of alternating structures. He illustrated and studied increasingly ambiguous situations. But possibly random dot stereograms (RDSs)[20] have provided the most interesting elaboration of this paradigm. The use of RDS has generated many new insights into vision science and visual perception. They have greatly influenced the understanding of depth, form, shape, and spatial perception and also have had an important role in understanding motion and illusory motion perception.

In its simplest form, a RDS typically consists of a pair of frames, F1 and F2 (Figure 12-34). Both frames consist of a random distribution of homogenous picture elements (e.g., dots, as used in the Ternus display paradigm). These elements in the two frames are both random and uncorrelated. The method for turning this pair of random dot displays into a stereogram is to copy a section of one frame into the other frame, but with a slight lateral displacement. For example, a square patch is copied from F1 and pasted into F2. Now, if one eye sees F1 and the other eye sees the (altered) F2, a normal binocular visual system fuses the two random displays into a single image. The image is a RDS; the square patch of random dots that is common to both F1 and F2 stands out and is evident as a square patch consisting of random dots. Whether the square patch was pasted into F2 slightly to the left or to the right determines whether the patch stands out in front or behind the rest of the image. The image around the pasted central square also consists of random dots and appears as either the background or foreground, depending on whether the central square patch stands out in front or behind.

A key point is that there really is no structural information in either F1 or F2 by themselves; the information arises from the relationship between F1 and F2. To find the image, the visual system performs some process of computational comparison between the two frames that simply cannot depend on any kind of cognitive, one-to-one comparison. Such a cognitive comparison is far too complicated, involving far too many individual elements. The computations must be done automatically, the result of some low-level, pre-conscious visual process that operates on the two retinal images before they reach consciousness.

:

Figure 12-34. A random dot stereogram (RDS) as a pair of 2-D images.

Now, consider the situation in which the two components of the RDS, F1 and F2, are shown to the same eye; but in succession, alternating in time. The central square patch common to F1 and F2 will appear as a patch of random elements apparently moving alternately slightly to the left and right. As in the RDS, there is no square form in either F1 or F2, individually. The square is invisible in F1 and F2 taken individually; the square is only visible from the correlation between F1 and F2. Again, some low-level, preconscious visual process operating on the two successive retinal images before they reach consciousness produces the neural information for the

[20] A Random Dot Stereogram (RDS) is a technique created by Dr. Bela Julesz (1971). A RDS describes a pair of 2-D images showing random dots which when viewed with a stereoscope produced a 3-D image.

perceived motion. Any RDS that can produce depth perception can produce the illusory motion. Anstis (1978) speculated "… that stereo vision, which developed much later than motion perception on an evolutionary time scale, took over and adapted many of the technical tricks that the visual system had already devised for seeing movement. It is interesting to note in this connection that retinal ganglion cells that respond to movement of an object in a particular direction are common in the pigeon, the rabbit, and the ground squirrel, animals whose laterally placed eyes look at different parts of the environment, but they seem to be absent in the cat and higher animals with binocular fields… It is conceivable that stereo vision might have preempted some of the neural circuitry which was originally devoted to motion perception."

Autokinetic effect

A completely different but profoundly compelling type of illusory motion, the *autokinetic effect* (AKE), can occur in the dark or under night time viewing conditions. It is defined as a phenomenon of human visual perception in which stationary, small points of light in an otherwise dark or featureless environment appear to move. It is this phenomenon that tricked the naturalist Alexander von Humboldt (1799) into thinking that some stars made oscillatory motions (Wade and Heller, 2003). Subsequently, this was recognized as a visual illusion, and the observed movement was subjective. By 1887 the phenomenon had been dubbed autokinesis, which indicated that it was a relatively easily appreciated illusion. This effect is another example of the notion that the more sparse or degraded is the visual environment, the more likely visual illusions are to present themselves.

Consider what may be one of the sparsest visual environments possible, a small spot of light in a completely dark field. The light could be solidly fastened to a wall. In fact, a subject could pound a nail in the wall and hang the light from it, so as to be convinced it doesn't move. But, after being placed in the dark and staring for a few seconds at the stationary light, the observer would perceive it as wandering about quite freely. It can easily appear to make excursions of as great as 45°. This apparent motion can have a rapid onset. In a classic study of U. S. military aviators, more than 50% of them reported the AKE within 13 seconds of looking at the light in the dark. Some have reported that the light can seem to move quite fast; "… with an apparent velocity of 15-20 degrees per second, giving the impression of a skyrocket or a rapidly moving shooting star" (Graybiel and Clark, 1945).

The phenomenology seems to demonstrate the dissociation among location, motion, and velocity. "The light sometimes appeared to travel quite rapidly in a particular direction yet never seemed far displaced from its original position. In other words, the rate of movement appeared to be more rapid than the rate calculated from the displacement of the light." (Graybiel and Clark, 1945)

The illusion is very powerful and convincing, and is commonly included in the training syllabus of military aviators. This is because they are routinely required to fly under conditions that can easily provoke the AKE, e.g., during night flight or under other conditions of degraded visibility.

Despite the facts that the AKE is relatively easy to elicit in the laboratory with so dramatic effects, and that it almost certainly is important for the control of vehicles under degraded visual conditions in air, space, ground, or even under water, there still is no single universally accepted explanation for the illusion. It is widely accepted that eye movements are important, but at best they can only explain part of the story. For example, eye movements recorded during the AKE may be only quite small, of the order of seconds or minutes of arc, while the observed illusory motion may be several orders of magnitude greater. It is almost as though the very simplicity of the stimulus makes the explanation of the illusion complicated. Among the reasons that the AKE remains so puzzling is that so many factors affect it. For example, if the spot of light is shaped like an arrow, the AKE is in the direction the arrow points. If the light is shaped like a bicycle, horse, or a walking person, the direction of motion is the direction in which the bicycle, horse or walking person is oriented. If the stimulus looks like a balloon, the motion tends to be upward; if the same stimulus is a parachute, the motion is downward (Toch, 1962). The plasticity of the AKE has even enabled the light to spell out words when subjects were told to look for them.

The most widely accepted explanation of the AKE is some version of that proposed by Gregory and Zangwill (1963). This class of explanations calls on the notion of a cortico-cortical feed-forward signal (efferent copy).[21] As Helmholtz first pointed out, physically moving the eyeball causes the entire image on the retina to appear to move. On the other hand, if the eye of an awake person is immobilized so that it cannot move, the person reports that the world seems to jump when attempting or willing the immobilized eye to move. These observations have been taken as evidence that when the nervous system sends an efferent signal to the extraocular muscles to move the eye, the sensory/perceptual systems receive a copy of that efferent signal so that the interpretation of the retinal image includes the anticipated voluntary motion of the eye. The results of the anticipated ocular motion are calculated into the percept. The perception of movement when the eye is passively moved is attributed to the absence of the efferent copy, whereas the perception of movement when the restrained eye attempts to move is attributed to the present of the efferent copy. Consistent with this idea is the demonstration that if the conjunctival sack around the eye is anesthetized, the eye can be mechanically moved in the dark without any perception of where the eye is looking. Apparently, there is no information coming from the eye about its orientation or position other than what is provided by the images on the retina themselves. It is as though the eye has no proprioception other than the visual image, so the efferent copy serves that purpose. Gregory and Zangwill (1963) proposed that, in a degraded sparse visual environment with few stimuli, the AKE is due to spontaneous fluctuations in this efferent copy system. The AKE is a clear example of the complexity of the illusions that occur in a sparse environment. The environment is sparse and the illusions are compelling, but the explanation is complex. Again, the AKE is not just a laboratory curiosity but something that should be anticipated in the real world, though admittedly under some extreme conditions.

Under normal daytime viewing conditions, the AKE rarely occurs. The relative position of an object with respect to other objects in the visual field provides information sufficient to determine whether the object is moving or not. Some have argued that the perception that an object is moving depends on the perception of stability (Dichgans and Brandt, 1978). That is, the perception of motion includes the perception of non-motion. With fewer visual elements there is decreased information about relative non-motion. With fewer elements in the visual field there is less information available concerning the relative distances among the stimulus elements. These relative distances among the visual elements is the very information missing during the AKE, information that would allow the calibration of the null point of the efferent copy.

Induced motion

Induced motion is the incorrectly perceived velocity/direction of the motion of an object caused by background motion (Duncker, 1929). Levine and Shefner (1991) give the following example: ``...consider a cloudy night sky with the moon ducking in and out of the drifting clouds. The moon is actually stationary relative to the clouds, but because the clouds take up so much more room in the visual field than the moon, they appear to be stationary while the moon seems to move in the opposite direction from them."

Unlike the AKE, there really is motion in this situation; the real motion of the cloud near the stationary moon, but the motion was misinterpreted. Simply because there is some structure in the visual field does not necessarily mean that the percept will be veridical. Movement was attributed to the wrong element. This perceptual misunderstanding emphasized the notion that all motion is relative. Without a reference there is no motion.[22] The sensitivity to the motion of a single dot in an otherwise empty visual field is very poor when compared to the sensitivity to the motion of the same dot when another dot is near by. But with two dots; which one moved?

---

[21] *Efferent copy* theory was developed by von Holst and Mittelsteadt (1950) to account for head adjustments made by flies in response to moving stimuli. The efferent copy mechanism is invoked to explain such phenomena as how a person perceives a motionless world when he/she shifts his/her eyes but perceives a moving world when pushed by someone else (Frijda, 2006).

[22] With respect to the AKE, the stimulus element is seen to be in motion with respect to one's self.

A simple form of this induced motion illusion involves two elements; a stationary element and a moving element that is larger than the stationary one, e.g., the moon and a cloud, respectively. When Duncker first reported this illusion of induced motion, he imaged a 2-cm spot of light on a much larger piece of cardboard. When the cardboard moved back and forth in front of the observer who was about a meter from the cardboard, it seemed that the stationary spot moved. Furthermore, it was irrelevant whether the cardboard moved and the dot remained stationary or if the dot moved on a stationary cardboard; both situations produced the identical perception – the dot seemed to move under both conditions. That is, the illusion went only in one direction; a moving spot did not make the cardboard appear to move. Duncker argued that the smaller object (the spot) seemed embedded in the larger one, the cardboard, and that the larger one provided a frame of reference for the observer to interpret the relative motion. The spot never served as the frame of reference for the cardboard, the spot never induced the percept of background motion. Duncker generalized his argument to say that the room in which the experiment took place, or at least the wall behind the cardboard, provided the cardboard's frame of reference. To induce motion in the cardboard, its frame of reference would have to move.

This suggested a further experiment. In a dark field, a stationary spot is surrounded by a rectangle that moves laterally. The perception of motion, however, is that the spot moves laterally in the direction opposite to the rectangle's actual motion. This merely recapitulates Duncker's original situation. If now a larger stationary ring is introduced, surrounding the moving rectangle, the percept changes. The rectangle is now seen as moving, as well as the dot. The ring provides the frame of reference for the rectangle, which enables the actual lateral motion of the rectangle to be visible. But the rectangle remains the frame of reference for the spot, which though in reality is stationary, is still also seen as moving. If the ring is removed, the motion of the rectangle disappears, but the induced motion of the spot continues (Wallach, 1959).

As another compelling example of induced motion, consider driving at night and a bicycle crosses from one side of the road to the other in front of the vehicle. On the rim of one of the wheels of the bicycle is a reflector. As the bicycle moves forward at a constant speed, the wheel rotates and the reflector makes a circle around the hub of the tire as it moves forward. The trajectory of the reflector incorporates the circular movement around the hub as well as the hub's forward motion, so the path the reflector traces is not a simple circle but, as illustrated in Figure 12-35, a series of arches reminiscent of a suspension bridge (i.e., a cycloid).

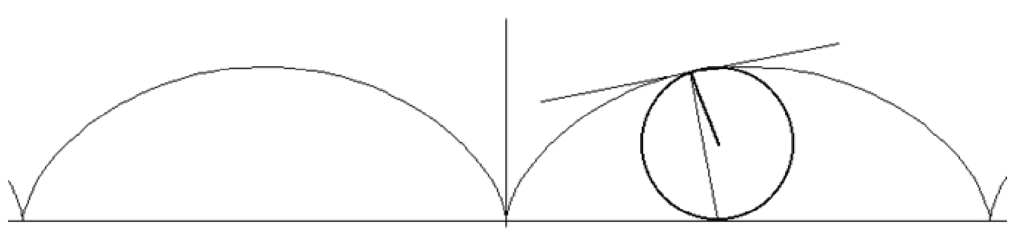


Figure 12-35. The trajectory of the reflector on a bicycle wheel a cycloid).

If the bicycle crossed from left to right in front of the automobile, the wheel is turning clockwise from the perspective of the automobile. Figure 12-35 illustrates the path of the wheel and reflector starting with the reflector at the very bottom. As the wheel turns clockwise, moving forward to the right, the reflector traces an upward path on the backside of the wheel traveling from the six o'clock position to the twelve o'clock position, followed by a downward path as the reflector travels from the twelve back to the six o'clock position.

Now consider the same situation during the day. Although the reflector follows the same path, it is virtually impossible to see cycloid motion described above. Instead, the wheel appears to move forward and the reflector proceeds in a circle around the wheel's hub. The hub and wheel provide the frame of reference for the reflector's circular motion while the rest of the bicycle, terrain, road, etc. provide the frame of reference for the hub and wheel's forward motion.

What happens when a second reflector is added to the hub and the bicycle again makes the crossing in the dark? The reflector on the hub translates to the right and the one at the rim again makes the cycloid motion. The rim reflector can appear to make several different types of motion. It may seem to circle around the hub or it may seem to add a loop component to the bottom part of its path, or it and the hub reflector may seem to be rotating around some third imaginary point midway between them, as though they are on the ends of a tumbling stick. For all of these, the path of the rim reflector seems to include some kind of backward component that is induced by the translation motion of the hub reflector in the rightward direction.

The literature is full of examples of induced motion effects that are complicated or difficult to predict. It should be no surprise however that in visually sparse situations with degraded, underspecified stimuli, the percept may be difficult to predict or ambiguous. The relative motions among the stimulus elements may come together in a fashion different from the sum of their individual components. Consider the two dots in Figure 12-36A. The dots move simultaneously, in the direction of the arrows, then they moved back to their original position. The individual motion of each dot by itself is clear and unambiguous; but when the two dots move at the same time the two paths sum into something else. The two dots seem to move along the diagonal toward and away from each other shown by the arrows pointing to each other in Figure 12-36B. These dots, moving diagonally form a group which itself has a diagonal path of its own, described by the dashed arrow. It is as though the dots become the frame of reference for each other in terms of a relative motion component. The residual motion forms a common motion component in terms of yet another frame of reference. Some have claimed that the person who is viewing the stimuli provides this other frame of reference; that is, the common motion is referenced egocentrically. This is the relatively straight forward notion that the individual perceives this common motion in terms of his or her visual field, as noted by Duncker himself.

It should be pointed out that Duncker (1929) mentioned, almost in passing, another situation that is now widely recognized as extraordinarily important for simulators, virtual reality, as well as HMDs. In his original experiment the large (26 by 39 cm) rectangular cardboard moved to induce motion in the 2-cm stationary dot. He described the situation when the subject was close to the dot, from 30 to about 100 cm. "When the subject now fixated the point he experienced the feeling of being moved with it to and fro. In some cases this was so strong as to cause dizziness. The subject had himself become a part of the induced motion system and was (phenomenally) 'carried along with it.'" Duncker seems to be describing an example of the illusion of self motion induced by visual stimulation, usually of the visual periphery. This illusory self motion, technically referred to as *vection*, was noted in the very beginning of this discussion of dynamic illusions with the example of a stationary car at the red light appearing to be rolling backward.

Figure 12-36. The left panel, A, shows the actual motion of a pair of dots. Individually, the motion of each dot is unambiguous. The right panel, B, shows the apparent motion when the two dots are presented simultaneously.

The Pulfrich effect

Consider a monocular see-through display that provides information to one eye superimposed on the binocular view of the distal world. Or, consider a binocular night vision device (NVD) with each eye viewing the world through two image intensifier tubes. In both cases the two eyes are viewing the world through different optical systems, and in each case the two eyes may be seeing images with different characteristics, including possibly different brightnesses. The monocular display could be imposing a filter that reduces the amount of light to one eye;[23] or the brightness of the two NVG tubes may be mismatched. Such luminance differences between the two eyes are known to affect the apparent depth of moving images. This dynamic illusion, called the Pulfrich effect, is important for optical systems designed to provide separate images to each eye. It also reveals important aspects of how the visual system works, aspects that are essentially unrelated to the kinds of illusions discussed so far.

The Pulfrich effect has been defined as "the apparently ellipsoid or circular excursion of a pendulum actually swinging in a plane perpendicular to the direction of view when a light-absorbing filter is placed in front of one eye" (Cline, Hofstetter and Griffin, 1980).

It is easy to demonstrate the Pulfrich phenomenon; a common classroom demonstration uses a pendulum and a filter of some sort. The filter does not have to be at all precise; the lens of a generic sunglass is sufficient. The pendulum swings back and forth in a left right direction in front of an observer who is watching the swinging motion with both eyes. An observer with normal binocular vision under normal viewing conditions sees the pendulum swinging back and forth in the frontal plane. Now the individual puts the filter in front of one eye, say the right eye, so that the right eye is still seeing the same image as the left eye, only dimmer because of the optical density of the filter. The pendulum is still seen with both eyes, but no longer seems to be swinging back and forth in the frontal plane. Instead, the pendulum seems to be swinging out of the plane in an arc. Specifically, with the filter in front of the right eye, as the pendulum moves toward the left, the path seems to bow away from the observer and as the pendulum swings back toward the right, it seems to bow toward the observer. In other words, its path seems to have a counter clockwise component to it when seen from above.

This effect is more powerfully seen against a rich contoured background, which is one of the differences between this effect and the other dynamic illusions discussed earlier. Since it does occur in a rich environment, good visibility is no guarantee against its occurrence. If anything, good visibility makes the illusion stronger.

The most widely accepted explanation of this phenomenon rests on the basic idea that nerve conduction is not instantaneous. Rather, it takes a finite amount of time for information to travel through the nervous system and that the speed of the conduction depends upon the stimulus luminance. Moreover, somewhere in the visual system the information from the two eyes must come together so that the information for the two eyes can be compared. This is the fundamental basis of stereoscopic vision which provides the ability to see depth. This aspect of vision depends on a comparison of the neural information arriving from the two eyes. Under normal conditions, the information arrives at approximately the same time. But the comparison of the neural information from the two eyes will be disrupted if the timing of the signals from the two eyes is sufficiently altered. That's what the filter does to cause the Pulfrich effect; it disrupts the relative timing of the signals.

As a general rule of vision, the dimmer the visual stimulus, the longer is its latency. This has been demonstrated in a great number of ways and is a consistent finding in many experiments. With the filter in front of the right eye, the neural signals from the right eye are delayed relative to those from the left eye[24]. If the pendulum moves from the right to the left and the signal from the right eye are delayed relative to the left eye, the right eye signal shows the pendulum lagging behind, that is, to the right or temporally in the visual field to where the pendulum would be without the filter.

With the pendulum moving from the right to the left and the filter in front of the right eye, the right eye's image, which lags behind that of the left eye, is shifted more laterally than it would otherwise be without the filter. The relative displacement of the image towards the temple in the right eye relative to that of the left eye informs the visual system that the right eye image is distant; at the position where the lines-of-sight from the two eyes would intersect. The same logical arguments hold for the return trip of the pendulum. With the filter still over the

---

[23] This is certainly the case in the monocular HMD, the Integrated Helmet and Display System (IHADSS), used in the AH-64 Apache helicopter. See Chapter 4, *Visual Helmet-Mounted Displays*.

[24] It may be argued that the Pulfrich effect is not an illusion; but, as will be made clear by its explanation, the effect reflects an exquisite sensitivity to luminance differences between stimuli to the two eyes.

right eye the right eye signal still lags. As the pendulum moves rightward, the image of the pendulum in the right eye visual field is nasal to where it would be without the filter. Since it is nasal, it is interpreted (seen) as closer. [25]

The Pulfrich effect depends on relative neural conduction speed; and those speeds can be manipulated in a number of ways, for example by dark adaptation. One eye can be made more sensitive than the other eye by dark adapting the one independently of the other. In this case, a dim light presented to the more sensitive dark adapted eye would look brighter than a more intense light presented to the less sensitive, non-dark adapted eye. Yet, the latency of the more sensitive dark adapted eye would be slower than that of the less sensitive non-dark adapted eye, regardless of the apparent relative brightness of the two lights.

It has even been reported that the luminance imbalance between the two eyes that produces the Pulfrich effect can produce differences in size, distance, and velocity judgments. This effect reportedly occurred just by looked out the side window of a relatively slowly moving car while putting the filter over either the leading or following eye. "With the filter over the leading eye, the velocity of the vehicle seemed increased, with the lens over the following eye, it seemed reduced. Objects by the roadside seemed further away or nearer according to whether the leading or following eye was looking through the lens." (Robinson, 1998) It is relatively easy to envision how electro-optical systems that provide different displays to each eye could have similar effects. The luminance presented to the two eyes inadvertently could be sufficiently different to cause differential visual latencies for the two eyes. Even if the user notices the luminance differences, the possible impact on perception might not be anticipated or realized. Over time, the eyes could well adapt independently to the luminance of their individual displays; so that the user would not even notice the differences. None the less, the differences in latency, and the distortions resulting from the Pulfrich effect would remain.

During the designing of the AH-64's monocular IHADSS HMD in the late 1970s, vision scientists expressed considerable concern over the potential for such problems as the Pulfrich effect. However, after nearly three decades of fielding, number studies of AH-64 Apache pilot visual problems and complaints have failed to confirm this concern (Rash, 2008).

Motion aftereffects (MAEs)

Motion aftereffects (MAE) were introduced earlier in the development of the notion that the perception of motion is a basic visual sensation rather than simply derived by the displacement of an object over the retina or across the visual field. MAEs are frequently referred to as the waterfall illusion in reference to the initial report by Addams (1864) that after a prolonged period of viewing the downward rush of waterfall, stationary objects seemed to have an upward motion to them. Helmholtz (1909) noted that while watching the landscape pass from the window of a carriage, the interior of the railroad car, when looked at, seemed to move in the opposite direction. Purkinje (1825) noted the MAE after watching a military parade pass in review. The effect is easily experienced today while riding a bicycle, watch the ground pass, then stop; the stationary ground seems to flow in the opposite direction.

MAEs were originally attributed to eye movements, an idea that was soon shown to be inadequate. Consider, for example, a stimulus shaped like the blades of a windmill with a diameter of about 2°. An observer looks at the center of the windmill blades as they rotate clockwise at a speed that permits them to be clearly visible individually rather than fuse into a blurred disk. After a few minutes, the rotation stops so that the blades are stationary. Nonetheless, the blades seem to rotate counterclockwise demonstrating a MAE. In other words, the stimulus generates a strong apparent motion in the absence of real motion. But this MAE cannot be attributed to eye movements because eye movements normally do not have such a circular component to them.

---

[25] It may be noted that when the Pulfrich effect is demonstrated between the two eye, and the eyes are following the pendulum, the depth effect probably includes an adjustment of line of sight between the two eye. The effect however, has also been demonstrated within one eye using stimuli of different intensities, in which case the depth effects are monocular. This only demonstrates further that the Pulfrich effect derives from the impact of luminance on conduction speed.

Furthermore, consider an elaboration of this experiment. The same 2° windmill rotates clockwise, but the observer looks slightly below it, so that in this case the windmill stimulus is in a part of the visual field above where the subject is looking. The subject holds fixation for a few minutes setting up the conditions for a MAE. When the windmill stops rotating the MAE is seen; the stationary windmill again seems to rotate counterclockwise; but only as long as it falls on the part of the retina that had been stimulated by the motion. If the eye changes fixation position so that the stationary windmill falls on a new location, say for example, the eye is now fixated above the windmill, the MAE immediately disappears. But it reappears when the eye returns to the original fixation position. This demonstrates that the MAE is localized; it depends on local stimulus conditions. The MAE is not uniform over the whole visual field, as would be the case if it depended on eye motions.

In fact, two MAEs can be set up in the same eye at the same time, one going in one direction while the other goes in the opposite direction. These and similar demonstrations have been taken as strong evidence that the MAE depends upon localizable neural activity. It should be noted that the MAEs have been set up with rotating spirals, so that the stimulus seems to be expanding or contracting depending on the direction of the spiral and the turn. The MAE in this case is one of depth, so MAEs can occur in the third dimension as well. Furthermore, they can occur with RDS displays that logically require the function of a post-retinal component.

The present discussion included MAEs because they are powerful and easily experienced so they can be expected to affect some aspects of performance in environment that include motion. Furthermore, MAEs have been very influential in the development of the theory underlying many parts of vision. The activity of parallel spatiotemporal channels underlies, or at least enables, much of visual perception. These individual spatiotemporal channels are thought to be essentially independent or parallel. This independence includes a channel's relative involvement with or contribution to MAEs. That is, the different spatiotemporal channels' underlying motion perception adapt independently. The extent of a channel's adaptation is determined by its relative sensitivity (i.e., tuning) to such different stimulus parameters as orientation, temporal and/or spatial frequency, color, as well as disparity and/or depth. Such independence among these channels means that, given the right conditions, at the same time and in the same retinal location, MAEs can occur in some motion channels but need not occur in others. This raises the possibility that multiple MAEs may be active at any one time in any one section of the retina, given the right stimulus conditions. It is reasonable to expect that these different MAEs tend to be mutually self-consistent in the real world. But, such self-consistency may not be the case in a fabricated world of HUDs and see-through displays. It may be that disparate MAEs generated by artificial environments in different parallel channels simultaneously may have consequences that have not yet been appreciated.

Real motion

The above discussion raises the question: "How relevant is illusory motion to the perception of real motion?" This simple question has a simple answer: It depends. There are different types of illusory motion; each of the ones described above almost certainly reflect different motion processes. In fact, they were chosen in part to illustrate the range of phenomena that can produce illusory motion; and this is not a complete list by any means. The same thing may be said about the perception of real motion. The relation between illusory and real motion depends on the specific stimulus conditions, and on how the question is posed.

Consider, for example, stroboscopic motion in which two relatively brief spots of light are flashed alternately on two different locations on the retina. Some inter-stimulus interval (ISI) between the two flashes will produce a unambiguous perception of motion; say with an ISI of about 75 ms. Some researchers question whether the motion seen in this situation has any bearing on the motion seen when a single spot of light moves between the same two retinal locations. On the other hand, some argue that the very similarity of the appearance of real and illusory motion points to the activity of a common set of processes.

This similarity between these two types of motion has been at least in part responsible for what is described as a lock-and-key notion. Specifically, the extent to which motion is seen depends upon the extent to which the stimulus is within the input tolerances of the processes that underlie motion perception. The simple fact that real and illusory motion look so much alike means that they must have something in common; and common sense suggests that the more they have in common, the more alike they look, although this does not prove the case. A second point is that real and illusory motion both are capable of generating a MAE. To the extent that such MAEs reflect the continued activity of some sort of process involved with motion perception, could indicate the extent to which the real and illusory motion stimulate such a motion sensitive process. Currently, there is anatomical and physiological evidence bearing on the issue as well, helping to clarify the similarities and differences. Under normal viewing conditions with real objects that move in a continuous fashion, the retinal image is certainly not continuous because of the eye's constant blinks and twitches. It is as though the visual system has been designed so that it cannot differentiate between the jerkiness of real moving images and ones that are stroboscopic. As far as the eye itself is concerned, there may not be all that much difference between the real and illusory motion.

Real and illusory motion combined

Remember the cowboy movies with the spokes of the wagon wheel going backwards, or the movie in which the aircraft starts it engine and the propeller appears to reverse direction repeatedly? These were images of the real motion of the wheel that combined with intermittent or illusory motion of the film to produce an emergent perception that was not present in isolation in either of the original stimuli. Depending on the technology the same thing could happen with HUDs and HMDs. To understand how the wagon wheel effect emerges, consider a wheel with evenly spaced spokes. If the wheel moves forward from left to right, the spokes rotate around the wheel's hub in a clockwise fashion. Consider the situation in which the image of the spokes are systematically photographed and displayed such that: (a) for the first image, a spoke is at the noon position; (b) for the second image the wheel has changed its clockwise rotational speed so that a spoke is at the eleventh o'clock position; (c) for the third image the rotational speed of the wheel has changed so that a spoke is at the ten o'clock position; (d) for the fourth image a spoke is at the nine o'clock position; and so on. When the whole sequence of groups is projected, it will look as though the wheel has a spoke moving in a counter clockwise direction as the wheel itself moves across the screen from left to right.

It should also be noted that, as far as vision is concerned, all the spokes are identical so it is irrelevant which of the spokes is at the designated position. This perception (is it an illusion?) depends upon the fact that all the spokes are perceptually equivalent; vision does not discriminate among the different spokes. If the wheel has two equally spaced spokes, it is irrelevant whether $5/12^{ths}$ or $11/12^{ths}$ of a rotation is completed; either of these puts one of the perceptually equivalent spokes at the 11 o'clock position. In other words, the phenomenon is not due merely to the sampling rate of the display, which in this case is a film, but also to the equally important fact that visual system treats all the spokes identically.

This suggests that in some sense the visual system calculates the equivalency of the different spokes. The perception of motion and form are not completely separable. For example, consider a bunch of dots that are essentially indistinguishable, like a school of fish, or a distant galaxy. There is the aggregate, the school, and its elements, the fish. Each element has its own motion which is somewhat separate from the summed motion of the aggregate. The shape of the aggregate changes because of the motion of its elements yet it retains its overall identity. If the aggregate passes over a background of other identical random elements, the aggregate still maintains its identity as it moves. In other words, no matter how amorphous is the aggregate shape derived from an average or shared motion component, the aggregate still has an identity that is motion-dependent.

These types of situations can be expected to occur with see-through systems that superimpose imagery on views of the real world or from other sensors. The imagery component, derived from some form of electronic synthesis with its time constants, produces illusory motion that combines with the view of the real world with its moving elements. These moving elements may be real, as in a see-through design of a HMD, or with the output of electro-optical sensor such as a forward-looking infrared (FLIR) or a head-mounted image intensification night vision device. These systems combine multiple apparent motions, either real or synthesized.

Real motion incorporates illusions

When a real object moves away from an observer, the size of the object's image on the observer's retina gets smaller, and the motion of the image on the retina decreases. The ideas underlying size constancy and retinal image size (see Chapter 10, *Visual perception and Cognitive Performance*), apply to velocity. For example, two gulf carts traveling at the same speed, simultaneously, across a football field, one at the 10-yard line and one at the 60-yard line will, to someone standing at one of the goal posts, appear under daytime viewing conditions to be about the same size[26] and to be traveling at about the same speed. The phenomenal or apparent distance sets the scale rather than the distance traversed on the retina.

This has been studied in the laboratory. One classic experiment measured the apparent velocity of downwardly moving black dots as they passed through a rectangular aperture that was vertically oriented, like a window (Brown. 1931). Two such devices were used in the experiment, one placed closer to the observer than the other. The observer's task was to adjust the speed of one set of dots so that they appear to be moving at the same speed as the other set. When the subject set the apparent velocity of the one set to equal the other, the physical velocities were essentially identical; distance made no difference. But there are confounding factors. Since the two devices are at different distances, the retinal size of the dots, the space between the dots, and the aperture are different between the two displays. The retinal images from the closer device are larger than those from the further device. The observer probably was setting the velocity of the dots of the near device to traverse across the aperture in the same time as it took the dots to traverse across the aperture of the far device. The observer was equating the rate of change with respect to the figure. The constancy of apparently velocity was the result of the differences in apparent size.

So the two displays were placed at the same distance right next to each other but one display was half the size of the other. Its aperture was halved as were dot diameter and the spacing between the dots. The observer's task was the same – to set the velocity of the downward streaming motion of the dots so that it appeared equal. In the dark, when the observer saw only the two displays, the velocity of the smaller display was just about half that of the larger one. The velocities were set with respect to the relative size of the displays. The displays provided the frame of reference. But as soon as the lights were turned on so that the surrounding room was visible; this relationship broke down. The surrounding room provided the frame of reference for the apparent velocity.

The experiment matched the velocity of the dots with a frame of reference that was either other dots or the surrounding environment. But the experimental result is not unambiguous. The velocity may just as well be the rate of displacement of the moving dots with respect to the frame of reference. The rate of displacement needs some consideration. It may be that the rate of displacement describes a change in the configuration more than it describes motion *per se*. For example, the matching may have been based on the perception that a dot was closer to bottom of the aperture than the dot in the other display, or that the dot in one display remained centered longer but in the other display it was off center more quickly. In other words, the velocity that was so neatly studied in the experiment, may have measured a velocity that is dependent more on a change in form or structure that a 'pure' velocity. Often the problem becomes one of identifying the actual stimulus components that provide the frame of reference, which in turn, determine the perception.

Because of the importance of binocularity for display technology, a related experiment should be noted. It is very similar in design but instead of using dark dots against a light background in an aperture; luminous dots were presented in a totally dark room (Rock, Hill and Fineman, 1968). One display was located at four times the distance of the other display. The observer's task was the same as previously described, to set the motion of the far display so that it matched that of the near one. The observer saw no frame of reference in this study. The luminous dots were in the dark. The only frame of reference was the one provided by the observer, that is, an egocentric one.

---

[26] Even if the carts are of different make and model so that they differ in size and shape, the individuals driving them would provide a size reference.

When the observer used both eyes to make the adjustment, the speeds were approximately equal; but when the observer viewed the displays with only one eye through an artificial pupil the speed of the far display was four times that of the near one. With both eyes and the natural pupil, even though the dots were completely in the dark, the observer had information about their distance because of the eyes' convergence and accommodation. The artificial pupil eliminated this distance information so that the velocity judgments were based on the velocity of the retinal image in the absence of distance information. This demonstrates, again, the extraordinary power of accommodation and vengeance as cues for distance, albeit cues which are essentially unconscious. This also demonstrates the extraordinary measures that need to be taken to ensure that the retinal image determines perception.

The Ames window illusion

Real motion in the real world also can be surprisingly ambiguous. Boring (1942) tells of Sinsteden who noted that the blades of a windmill seen obliquely in silhouette against the bright evening sky seemed to reverse their direction of rotation periodically. The blades appeared to rotate clockwise, then counterclockwise. According to Robinson (1998) the same effect can be seen with rotating radar antenna in the middle distance. Figure 12-37 pictures a windmill in silhouette from an oblique angle. Notice the middle vane of the three visible vanes. The curved arrow shows the direction in which this vane moves. It is pointing to the left; but which way is that, clockwise or counterclockwise? If we are approaching the windmill from the front, then the vanes are turning in the counterclockwise direction. If we are approaching the windmill from behind; then the vanes are turning in the clockwise direction. If we can't tell the direction, if it is ambiguous, then the rotation can be one or the other. If the frame of reference is ambiguous, so is the motion.[27] In this demonstration it is not only the velocity but the direction of motion that is underdetermined.

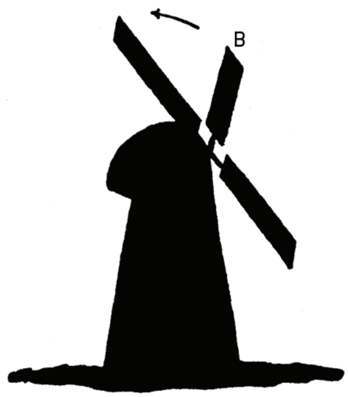


Figure 12-37. Whether the vanes are turning clockwise or counter clockwise depends on whether the vane with the question mark is near or far. Is the left or right side seen as the plane of rotation? (after Boring, 1942)

Ames (1951) developed what has become a relatively well-know demonstration of the ambiguity of perceived motion. This demonstration involved a trapezoidal window as illustrated in Figure 12-38.[28] The left and right sides of the window are parallel, although of unequal lengths, while the top and bottom are of equal length but are not parallel. When the trapezoidal window is seen viewed in a frontal plane, which is equivalent to the window

[27] This effect is well shown by casting the shadow of a slowly rotating vane upon a screen, thus removing all information of which is the back and which the front.

[28] The Ames trapezoid (Ames window) is a style of window which, when observed frontally, appears to be a rectangular window but is, in fact, a trapezoid.

being parallel to the page, the window looks like a regular rectangular window seen at an angle. For the demonstrations, the window's long, nonparallel bottom side was mounted on a vertical shaft, and the window rotated on this shaft at about 3 to 6 revolutions per minute (rpm). From a top down view, the window rotated in a clockwise direction. The observer viewed it with either one or two eyes from a distance of about 10 to 25 feet (x to x meters). The window was illuminated in an otherwise dark room.

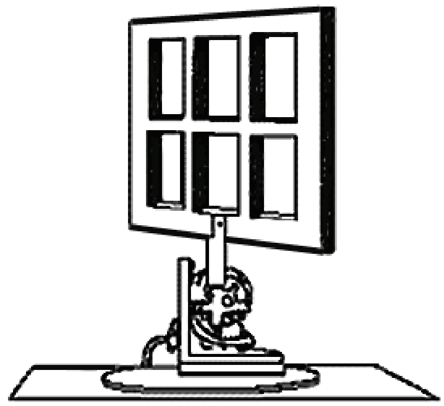


Figure 12-38. Ames trapezoid or window.

Although the rotation of the window was at a consistent speed in a constant direction, it certainly did not look as though that were the case. According to Ames: "As the trapezoidal window slowly rotates about a vertical axis, instead of appearing to rotate completely around, it appears to oscillate back and forth through an angle of about 100°."

To understand what is going on, remember that the parallel vertical sides of the window are of unequal lengths so that the top and bottom pieces of the window, though equally long, are not parallel. When the window is in the frontal plane, the observer simply assumes that the left and right parallel sides are equally long and the two non-parallel sides to be identical and horizontal. Since the visual system perceives the two vertical sides to be essentially identical, even though one is substantially longer than the other, the visual system creates a perception that is consistent with the given stimulus conditions.[29] If the two vertical sides of the window are identical, then the longer side has to be closer than the shorter side. This also means that the non-parallel top and bottom pieces can be seen as horizontal and therefore, parallel, just as they should be in a normal window. In order for the observer to see the trapezoidal window as rectangular means that the visual system fails to recognize that the window is in the frontal plane. Instead, the window appears to be at an angle such that the short side is seen to be at a greater distance than the larger side.

Imagine that the window is in the frontal plane with the long vertical side on the right and the shorter side on the left. In this orientation the long side is at the 3 o'clock position and the short one is at the 9 o'clock position relative to an observer at the 6 o'clock position. As the window rotates clockwise through 360°, the long side approaches the observer positioned at the 6 o'clock position. From the observer's point of view, the window's long side approaches as it sweeps leftward; and, after the long side passes the 6 o'clock position, the long side seems to recede at it continues to sweep leftward to the 9 o'clock position. Of course, as the long side sweeps from the 3 to 9 o'clock position passing through the 6 o'clock position, the short side sweeps from the 9 to the 3 o'clock position, passing through the 12 o'clock position.

The question is what happens as the long side continues its journey from the 9 o'clock to the 12 o'clock position. The key to this is to remember that the long side always seems to be closer to the observer than the

[29] Note that the stimulus conditions include the observer's past experiences.

frontal plane while the short side always seems to be further away from the observer than the frontal plane because of their relative size. According to this way of thinking about the window's appearance there is really no paradox or confusion at all. When the long side passes through the 9 o'clock position, it moves leftward, but since the long side seems to be closer to the observer than does the frontal plane, the long side seems to be approaching the observer at the 6 o'clock position rather than moving toward the 12 o'clock position, which is what it is actually doing. Conversely, as short side moves from the 3 to the 6 o'clock position it always seems to be further from the observer than does the frontal plane, so the short side seems to be heading back to the 12 o'clock position.

These illusions are very powerful, apparently sufficiently powerful to force the visual system to accept impossibilities simply in order to be consistent with the illusion of oscillation. The visual system cannot free itself from the basic illusion that the window is rectangular. All the subsequent perceptions are forced to conform to that basic misperception.

But this is exactly Ames' point. The visual system is more than just easily confounded; the visual system conforms its perceptions to its expectations. According to this idea, the visual system can neither accept, nor anticipate the reality of a trapezoidal window. All its experience is with rectangular ones. Even when the observer knows about the trapezoidal nature of the window, it is not enough to conform the perception to the reality. The visual system makes as much sense of the world as it can, and it uses its past experiences as the basis to accomplish this.

This is more than just an early demonstration that perception is as much a top down as bottom up affair. Therefore, the visual system configures the stimuli to fit its understanding of the world, and this understanding is what previous experience had prepared it to expect. The difficulty of overcoming these expectations may be one of the reasons why spatial disorientation occurs. All of the observers who reported to Ames that they saw the window oscillate were in a sense experiencing unrecognized spatial disorientation. They had no idea that they were mistaken. This is a problem that must be addressed with HMDs and HUDs.

From the discussion so far, it would seem that vision is a rather passive process. The studies presented have emphasized a stationary receptor system responding to a rather simple, artificially-sparse pattern of lights in a correspondingly simple, artificially-sparse environment. In a sense, studies deriving from this tradition, which is commonly referred to as physiological optics,[30] are designed to reveal mechanisms, processes, or functions operative in the visual neurosensory system of the organism. The basic assumption underlying the tradition of physiological optics has been that there is not much difference among physiology, sensation, perception, and psychology.

But even the most sedentary of seeing beings do not remain completely stationary throughout their life cycle. The visual processes that underlie the perception of motion resulting from the movement of an organism through the environment are certainly as important for the survival of the organism as are the visual processes that underlie the perception of the motion resulting from the movement of elements in the environment surrounding the organism when it is stationary. Over the last half of the twentieth century there has been an increasing emphasis on considering the visual environment of the organism as it moves and acts in the world. This emphasis on the moving organism owes much to the influence of J. J. Gibson (1959; 1966; 1979).

This newer approach has been called *ecological optics* to contrast it with physiological optics. Ecological optics emphasizes the visual ecology in which the behaving organism acts and is far less interested than physiological optics in what goes on inside the organism per se, and far more interested in the interaction between the organism and the environment. This approach to the study of vision sees the organism and the visual environment affecting each other in a tight feedback loop. According to ecological vision there is no guarantee that any of the painstaking studies of physiological optics with its carefully controlled pulses of light under rigorously controlled lighting conditions have any bearing on the way the visual system functions in the real world. The field of

[30] This tradition has its roots in Herman von Helmholtz (1866/1963) *A Treatise on Physiological Optics*. 3 Volumes (J.P.C. Southall, Ed. and Translator). New York: Dover.

ecological optics challenges the very validity of the microscopic level of analysis of the laboratory studies. This point of view raises the possibility that many of the results of these laboratory studies may be simply experimental artifacts.

At least two factors contribute to the increased importance of ecological optics. The first is the undeniable artificiality of traditional physiological optics. The second is the rise of computational vision, including virtual realities, computer or artificial vision, image analysis, display technology, and so forth. The extent to which real world visual information analysis, the domain of ecological optics, has increased in importance is proportional to the extent that computational vision has moved into the real world.

The ambient optic array

The idea of the ambient optic array, which is central to ecological optics, may be considered simply as the different patterns in the light that surrounds the organism's eyes. It is the different brightnesses, colors, shades, shadows, and so forth, that the eyes see as they look at the world, real or virtual. The ambient optic array is not composed of the individual points of light devoid of structure, substance, or meaning. That is the purview of physiological optics, the response of the visual to the parameters of the individual light stimuli. Ecological optics is concerned with the information about the environment that is contained in the pattern of stimulus parameters that comprise the ambient optic array. The fundamental idea is that the information in the ambient optic array is sufficient; it contains all the information that the organism needs. The ambiguities of visual illusions, impossible figures, the Ames demonstrations, and the like, are due to the artificiality of the contrivances that intentionally under-specify the stimulus. The stimuli may be amusing and even illustrative, but what they mean in the real world of real perception is limited; they are not much more informative than any of the other studies derived from physiological optics.

Consider the real world that a pilot confronts from the cockpit. The horizon divides the sky from the ground and the kind of information the optic array contains differs from the two areas. The sky is characterized by open expansive areas, gradual changes of shade and brightness, and possibly clouds. All of these contain relatively low spatial frequency information. On the other hand, the high spatial frequency information resides on the ground plane. The topography and terrain provide this information. Much of it can be described as visual texture. Even such large objects as landing fields or football stadiums become indistinct, lose their individual identity, and become visual texture at a great enough distance and in a certain visibility. This is not the only difference between the sky field and the ground field; texture is only one of the more obvious and basic.

Texture differences between sky and ground survive when the pilot lands the aircraft and stands on the ground. The visual array from the sky is still characterized by low spatial frequencies while the ground contains objects that recede into texture at distance. As the pilot stands beside the aircraft, the texture of the local asphalt remains, as do elements of its granularity. But as the ground stretches across the airfield, to the fence that borders it, the various components of the ground merge into an indistinct average.

The difference between sky and ground remains a basic characteristic of terrestrial vision, and one of the clearest dimensions of the difference is texture. The surfaces of all objects have some type of structure that is relatively, but not perfectly, homogenous. Each surface has characteristic texture or granularity, the density of the surface elements. There are ways to quantify texture and texture differences between objects. The merging and changes in optical texture with distance is an important source of information in the ground plan optic array.

A convenient example is the regular black/white checkerboard pattern of a tiled floor. From one wall, the regular check pattern spreads out to the other walls. Its regularity is obvious as is its flatness. In fact, the regularity is visually consistent with its flatness; any deviation from flatness would be immediately evident as an irregularity. But despite the fact that all the tiles look square, not a single image from a single tile is square. Further, the outline of every tile is different. The differences are lawful, described by linear projective geometry.

This texture information is not restricted to the side view. Wherever texture is in the visual field, there is some kind of texture flow during self-motion. With forward velocity, texture elements stream in a radial direction when the eye is oriented in the direction of motion. There is a tendency to think of this flow field as providing heading information; but this is probably an exaggeration since the eye is constantly scanning the visual array. The eyes of a pilot, driver, or other moving individual are sampling moving texture gradients as the eyes change their line of sight and with each line of sight change the eye picks up different information from different segments of the array. In terms of information content, the radial flow field may be the one with the least amount of dynamic information.

One of the challenges of a HUD is the ability to provide texture information when it is important. One problem is that texture information requires high frequencies, which usually involve relatively high resolution and the associated hardware and computational overhead that high resolution requires. So much of the behavior that is determined by texture is non-conscious. It is hard to notice when the texture is missing. Our normal visual systems function perfectly well at night. Observers fill in the missing textures without even noticing. Is it noticeable when these are wrong or inaccurately realized in a virtual display? They certainly do not seem to be a problem in some animations. The visual system may be very forgiving about the inaccuracies in the representation of texture. However, such inaccuracies may be deadly when controlling an automobile or aircraft. The degradation or inadequate representation of texture information in fact may have contributed to some of the AH-64 accidents attributed to dynamic illusions.

A final word on visual illusions

The main premise underlying the preceding sections on static and dynamic illusions is that there are many processes that lead to illusory perceptions, and that these are far more likely to occur under degraded visual conditions. The visual system is very good at filling in missing information while ignoring other, even conflicting, information in order to create a coherent picture of the world. Most of the time, the process is unnoticeable. Visual perceptions need only be sufficiently adequate for function; i.e. survival, and that is the limit of the scale of precision or accuracy required.

The images provided by HMDs and other electro-optical displays provide information to the user. There are some applications in which the user is passive, merely observing the display; for example, watching the animation. The accuracy needed for that task is certainly not the same as needed for the successful control of a system, such as holding a helicopter in a hover. For this task, texture and shear may be vital. The more controversial point is that the successful completion of the task does not necessarily prove that the perceptions were accurate; it means merely that the task was successfully completed. Whatever misperceptions may have occurred did not prevent the successful completion of the task. The next time the task is attempted under similar visual conditions, the misperceptions may be more intrusive.

There is a tendency to consider the rich images from natural world as the gold standard against which displays should be judged. The realism of photo realistic synthetic reality displays has an intuitive appeal. Part of the shortcoming of this intuitive appeal is its naiveté. Realism itself is full of potential illusions, and realism usually is good enough for every day tasks. But when confronted with tasks that go beyond those for which the visual system has evolved, misperception can occur. Assuming survival, blind to the errors in perception, learning may not have occurred for the next time.

Illusions and HMDs

This discussion of visual illusions has argued that they are not just curiosities but an integral part of normal vision. Furthermore, they make possible pseudo-reality, virtual realty, conformal and other advanced displays (e.g., HMDs). These displays depend on the ability of the visual system to see what is *not* there, or equivalently, its very fallibility and failure to see what *is* there. The successful implementation or migration of these emerging display technologies to head-mounted systems presumes the understanding and control of the illusions that fundamentally underlies the ability of the visual system to make sense of the displays.

The illusions (misperceptions of reality) that occur when using HMDs and NVGs are often not unique to these devices (Crowley, 1991; Crowley, Rash and Stephens, 1992), but the effects of these illusions may be exacerbated because of certain characteristics of HMDs and NVGs. This section briefly discusses several reported NVG and HMD illusions.

NVG illusions

The misperception of depth is probably one of the most commonly reported illusions when using NVGs (Crowley, 1991; Miller and Tredici, 1992; U.S. Army Safety Center, 1991). Although not unique to NVGs, the use of NVGs does increase the probability and, perhaps, the severity of the misperception of depth during flight. The characteristics of NVGs that exacerbate the misperception of depth include:

- Reduced visual acuity (thereby reducing stereopsis capability [Wiley, 1989] and reducing texture gradient perception [Miller and Tredici, 1992])
- Lack of color (reduces aerial perspective)
- Potentially unbalanced light levels in the two channels (causing the Pulfrich effect [Crowley, 1991; Pinkus and Task, 2004])
- Limited field of view (reduces geometric perspective)
- Elimination of the physiological link between accommodation and convergence. Accommodation depends on the NVG eyepiece setting, whereas convergence depends on the distance of the object being viewed and the NVG input/output optical axes alignment (Miller and Tredici, 1992)

Crowley (1991) provides aircrew comments extracted from his survey that exemplify many of these distance misperceptions, e.g.:

- "*A break in cloud cover allowed a large amount of moonlight to illuminate ground between aircraft and ridgeline (about 10 miles) giving illusion of hills being much closer (only 5 miles away).*" This problem (brighter objects appearing closer than the really are) could occur without NVGs, but the NVGs can enhance the effect due to their limited dynamic range and the auto gain feature.
- "*We were lead of flight of 2…even though radar altimeter was functioning, both aircraft descended to within 35 feet of ocean surface with no visible change in ocean surface.*" This is an example of the problem of no color and relatively low resolution of the NVGs making it more difficult to see already limited surface texture (such as the ocean surface) to gauge distance (altitude in this case).

Crowley (1991) noted that three respondents in his survey described a "3-D effect" or "disturbed depth perception," which they attributed to brightness differences in the two channels of the NVGs. Crowley noted this could be due to the Pulfrich effect, which can give rise to a false stereopsis-induced depth perception. Pinkus and Task (2004) found in a controlled laboratory study that a brightness ratio of 1 to 1.26 (one channel is 26% brighter than the other) between channels is enough to make the Pulfrich effect statistically detectable, at least some of the time. A sample of fielded NVGs measured in the early 1990's indicated approximately 8.5% had a brightness imbalance of 1.24 or more, indicating the Pulfrich effect could be responsible for some depth misperceptions with NVGs.

It is especially difficult to gauge the distance or identify small light sources or groups of light sources with or without NVGs. However, for small, bright light sources NVGs produce a "halo" effect around the light source that can make it appear. The size of the halo depends on the particular image intensifier tube and the size can even vary across the surface of a single tube. Also, the NVGs make point sources of light much brighter than they would appear to the unaided eye and the lack of color in the NVGs makes it more difficult to differentiate between different light sources. All of these factors can produce misperceptions when using NVGs. A couple of relevant light source related misperceptions from Crowley's survey (1991):

- "…while flying over water, joined on the red running light of an oil tanker vice wingman."
- "…noticed a very bright light at my 10:00 – same altitude, very close. …made a hard right turn to avoid what I thought was another aircraft. Flight engineer reported that it was an automobile on a hill."

Several respondents in the Crowley survey (1991) report misperception of the slope of the ground during landing with a helicopter when using NVGs. Two such reports indicate that the slope could be more or less than what was perceived:

- "Troop insertion…misjudged percent of slope on landing to the ground. Hit left skid high…the slope was much more severe than anticipated."
- "A student I was instructing assessed flat ground as nearly 15° and refused to land even when ordered to."

There are several possible explanations for the misperception of slope such as reduced texture visible through the NVGs, false geometric perspective due to terrain features, or lighting intensity effects. Another possible explanation of this misperception may be the fiber optics image rotators used in NVGs. These devices are intended to re-invert the image from the micro-channel plate of the NVGs so that the image viewed through the eyepiece will be right side up. However, there is a tolerance on the fiber optics rotators of +/- 1°. If the two oculars of an NVG happen to have image rotators that are within tolerance but in opposite directions (one plus 1° and one minus 1°), then it is possible to have an image rotation difference between the two channels of 2°. This rotation difference could induce a false stereopsis that could make flat ground appear to be sloping toward or away from the observer depending on the direction of the rotational difference. This is an area that could use more research.

Misperception of motion is another commonly reported illusion that can occur with or without NVGs. The limited field of view and reduced visual acuity when using NVGs can add to the already existing conditions that can produce motion illusions. One NVG visual effect reported by one of Crowley's responds that may fit in this category was especially interesting:

> "*…while flying over smooth water in a turn, the reflected stars in the lake could be seen…after looking inside the cockpit to outside, the appearance of the stars when looking down in to the turn produced severe vertigo*." Since the NVGs amplify light making stars (and reflection of stars) much more visible than they would be to the unaided eye, this effect would be enhanced with the use of NVGs. The pattern of stars reflecting off of the smooth water would appear stationary because the image of the stars is at optical infinity. This would provide the illusion that the helicopter was stationary when, in fact, the pilot knew that the aircraft was flying at a relatively high rate of speed. The conflict between knowing that the aircraft is moving relatively fast and seeing the star pattern below as stationary (instead of streaming by as cultural lighting would) may have added to the vertigo.

Several respondents in the Crowley (1991) survey reported faulty attitude judgments:

> "*Ridgelines at various angles behind each other produce false and confusing horizons.*" Although these illusions may also occur without NVGs the limited field of view and reduced visual acuity are most likely significant contributing factors.

HMD illusions

Most HMDs that are fielded today have a monocular display such as the Joint Helmet Mounted Cuing System (JHMCS) and the IHADSS but binocular HMDs have also been fielded such as the TopOwl® HMD by Thales. In night vision mode, one can expect to see all of the illusions described above and a few more because the optical inputs for the TopOwl™ in night vision mode are spaced apart significantly more than the interpupillary distance (IPD) of the wearer. This will give rise to hyperstereopsis, explained earlier in this chapter. The JHMCS is currently limited to monocular symbology only and is therefore limited as to what illusions it can create since it is not providing an image of the outside world. The remainder of this section briefly discusses illusions reported for the monocular IHADSS.

Since the IHADSS is a monocular device, none of the binocular-based illusion mechanisms can occur. Also, the majority of illusions reported when using the IHADSS probably also occur without the IHADSS. Rash and Hiatt (2005) noted that a survey of 40 pilots that had recently returned from Operation Iraqi Freedom (OIF) reported significantly fewer instances of static and dynamic illusions for the IHADSS compared to a previous survey conducted in 2000 (n = 216 for this survey). The explanation for the difference, as suggested by the aviators in the 2005 survey, was that in the peace-time year 2000 era flight hours were limited and were primarily for maintaining proficiency with the IHADSS, and therefore pilots would fly relying almost entirely on the monocular helmet display unit (HDU) of the IHADSS to the detriment of other possible visual information available to the non-HDU eye. During OIF flight time was much higher and, the explanation goes, pilots could pay attention to both visual inputs (the un-aided eye and the eye looking at the IHADSS display) and therefore maximize the information available. Whether this is the correct explanation or it is simply a matter of “more flight time makes a pilot more proficient,” one thing is clear: it is possible to reduce the instances of illusions through increased flying experience.

The two most reported static illusions (Rash and Hiatt, 2005) were faulty height judgment and faulty slope estimation. Although both of these illusions can occur without the IHADSS they are most likely enhanced by the IHADSS because of the reduced visual acuity (about 20/60 Snellen), limited field of view, the monocular viewing (as opposed to binocular), and possible slight errors in view angle between the pilot’s line of sight (though the image) and the sensor’s direction of view (where the sensor is pointed). Also, with the sensor mounted so far from where the pilot’s eyes, the pilot must mentally compensate for the shifted view point, which can lead to height and slope misperceptions.

The two most reported dynamic illusions (Rash and Hiatt, 2005) were undetected drift and faulty closure judgment. Again, these can occur when flying without the IHADSS but are probably enhanced because of all of the same limitations listed above.

This section on NVG and HMD illusions was intended to provide a sampling of some of the types of illusions that can occur with these devices and is not a comprehensive treatise on the topic. Hopefully, it provides some insights into the types of illusions, and their root causes, that are actually reported by aircrew members when operating with NVGs and HMDs.

Spatial Disorientation

Spatial orientation has been described as “the most fundamental of all behaviors that humans engage in” (Previc and Ercoline 2004). The process of spatial orientation is one that humans are scarcely conscious of while operating in a normal environment, i.e., standing on the surface of the earth experiencing one gravity (1G) of downward force and with usable visual cues.

Spatial disorientation (SD) represents a failure to maintain spatial orientation; this is relatively uncommon on land outside of neurological clinics but unfortunately all too frequent in the air. There is an exception to this generalization in that the increased use of remote sensors to pilot military ground vehicles has lead to a marked increase in incidences of SD (Johnson 2004). In the aviation arena, collaborative work by Hixson et al. (1977)

showed that SD accounted for 7% of all known-cause accidents with 24% of those involving fatalities. Later studies by Durnford et al. (1995) and Braithwaite, Groh and Alvarez (1997) showed that SD was a major or contributing factor in 32% and 30% of all mishaps, respectively. The dramatic increase over the Hixson study is attributed to a different definition of SD used in the latter studies, in which the criterion used was that no accident would have occurred in the absence of SD. A later study by Curry and McGhee (2007) using the same criteria on rotary-wing accidents in the second Iraq conflict showed an increase to 37%. Since the 1970s, the most commonly used and widely accepted definition of SD is that of Benson (1978), namely the situation occurring "…when the aviator fails to sense correctly the position, motion or attitude of his aircraft or of himself within the fixed coordinate system provided by the surface of the earth and the gravitational vertical." Also often used is the clause of Vrnwy-Jones (1988) that includes as SD a misperception relative to another aircraft or known stationary object. Geographical embarrassment (getting lost) is specifically excluded but would be included in a broader definition of *situation awareness* (SA).

SD can be regarded as a wholly contained subset of the wider area of SA; thus, if one has SD, then one must have lost SA. Whereas, a pilot can lose SA (land at the wrong airport for example) while maintaining spatial orientation throughout. However, the factors that predispose to a loss of SA, such as high task intensity or sleep deprivation, also often predispose to SD. Many SD accidents can be attributed to pilot distraction or channelized attention in flight, during which the aircraft slips into an unusual attitude so gradually as to remain undetected by the pilot's vestibular or somatosensory systems (Albery, 2006). This form of SD is called Type I, where the pilot does not recognize the fact that he/she is disoriented and is therefore extremely dangerous; a typical Type I SD accident would be the graveyard spin.[31] The other form of SD, Type II, occurs when the pilot is aware of the disorientation but is able to combat it by use of the aircraft instruments or handing control to another pilot. This is much less likely to end in an accident and is a regular feature of most pilot careers (Holmes et al., 2003).

Perception of orientation

Human beings orient themselves in space using a combination of different senses mediated by cerebral function. This system is "designed" to operate in the natural environment which until very recently did not include the cockpit of an aircraft. The ability to maintain posture, balance, locomotion and to stabilize the head are dependent on a combination of sensory inputs from the eye, inner ear and somatic sensors with a minor contribution from auditory signals. These are all important in a multiple loop feedback system to allow humans the necessary control over their own bodies. Some of these inputs overlap, and central processing in the brain coordinates all the separate and complimentary components to a coherent set of outputs. This is all very well within the human evolutionary niche, but in the air some and sometimes all of these systems can provide erroneous information making the central processing job more difficult and leading to SD. Therefore, one could regard SD as a normal physiological function (or, perhaps, condition) when the body is subjected to the altered environment of flight.

Figure 12-39 represents the interplay of the various sensory and perceptual components that are involved in the maintenance of spatial orientation in flight. Sensory input to this model is conventionally divided into *subconscious* and *conscious* fractions. The subconscious element consists of the ambient visual system for visual positioning in space, the vestibular system to detect angular and linear accelerations including gravity, and the tactile and proprioceptive systems detecting linear acceleration and inertial force. At the conscious level, focal vision is for detecting the complexity of the central visual field, including flight instruments, while the auditory system provides sound localization cues. These conscious systems require interpretation and intellectual constructs and therefore place a load on central processing before their addition to the whole of the sensory dataset. This full central-processing compares the current situation with learned internal models and generates estimates of the current position, motion and attitude of the aircraft and the airman within.

---

[31] A graveyard spin is a sub-threshold increase in angle of bank and pitch down that is not recognized by the pilot until such time as the very tight spiral is unrecoverable, all aboard ending in the graveyard.

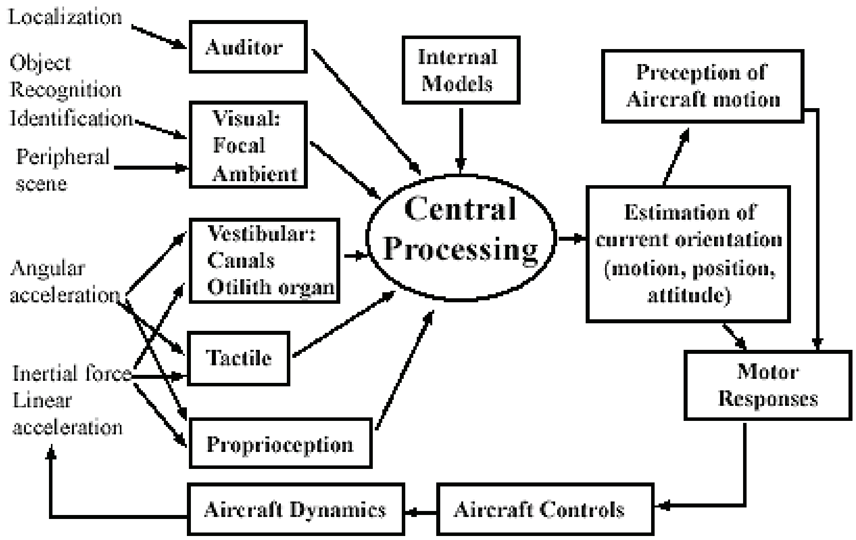


Figure 12-39. Schematic diagram of the spatial orientation mechanisms in flight (adapted from Benson, A. J., Spatial Disorientation - A Perspective (2002).

Vestibular contribution to orientation

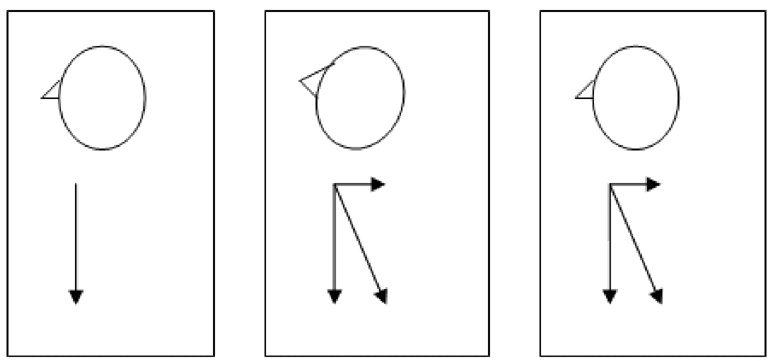
The inner ear, as well as housing the organ of hearing, contains the vestibular apparatus, or organ of balance. The vestibular sense, as previously mentioned, is largely unconscious and only achieves prominence with illness or vigorous stimulation, producing nausea or dizziness. It consists of two major parts: The semicircular canals as one portion, and the utricle and saccule as the second portion. Morphologically, both portions have a similar basis of a fluid filled space into which project inertially sensitive structures that are attached to nerves leading to the brain. A detailed explanation of their function is beyond the scope of this volume, but the subject is given a thorough treatment in Previc and Ercoline (2004). Essentially, the semicircular canals are three accelerometers oriented in the planes of yaw, pitch and roll when the head is normal to the horizon and vertical with respect to gravity. Similarly, the utricle and saccule are accelerometers sensitive to linear accelerations and tilt of the head relative to the gravitational vertical. The functioning of these structures is characterized by some common features. They all have thresholds of detection below which they will not detect accelerations and therefore provide no input to central processing despite an acceleration being present. The semi-circular canals have a detection threshold expressed in Mulder's law,

$$\alpha \, \tau = 2.5 \text{deg/s} \qquad \text{Equation 12-1}$$

where α is the magnitude of the angular acceleration, and τ is the time of application of that acceleration. Simply stated, the weaker the accelerative force, the longer it must be applied to be detected. There is a threshold of angular acceleration below which the canals will not respond; this is likely to be around 0.14, 0.5 and 0.5 deg/sec^2 for accelerations in pitch, roll and yaw, respectively, when sustained for 10 seconds or more (Clark, 1967). They also rapidly habituate (to the order of 20 to 30 seconds) to a maintained acceleration. For example, in a level coordinated turn with no visual reference, the sensation of turning will be lost with a subsequent reversal of sensation when the turn is stopped, producing "the leans" illusion.

The evolutionary purpose of the vestibular apparatus is to enable steady-focused vision during rapid head movements, and the reflexes that enable this are rarely beneficial in the aviation environment. For instance, the counter-rotation illusion described above can lead to nystagmus, an involuntary oculomotor response that destabilizes the retinal image.

The utricle and saccule (collectively known as the otolith organ) function in a similar manner to the semi-circular canals but detect linear rather than angular accelerations. They are also responsible for our sensation of gravity and therefore have a baseline level of activity providing the vertical reference *up*. In zero-G environments, it is not unusual for individuals to feel upside down when they are not – a potent source of motion sickness. In a normal gravitational environment, the otolith organ also senses the attitude of the head relative to the gravitational vertical. This can lead to problems in the flight environment, as this means that the otolith organ cannot differentiate between a head tilt and a sustained linear acceleration (Figure 12-40). Many aircraft have been lost in conditions of limited visibility when a take-off ($+G_X$) acceleration was misperceived as a backward head tilt, hence pitch up and a forward correction applied to the controls resulting in a nose-over into the ground or sea. This is known as the somatogravic illusion. The otolith organ does have a detection threshold, but there is a wide variation in values quoted (Guedry, 1974). The most accepted values for detection of sustained linear acceleration are 0.01 m/sec^2 (0.03 feet/sec^2 or 0.001G) for a supine subject and 0.06 m/sec^2 (0.20 feet/sec^2 or 0.006G) for an erect subject (Meiry, 1965). This is a very fine level of detection, and for all practical aviation purposes the otolith organ will detect all translational accelerations. However, the organ does not detect linear velocity so is of no utility, for instance, in detecting the ongoing drift of a hovering helicopter once the initial acceleration has stopped.



Upright Backward Head Tilt Forward Acceleration

Figure 12-40. With gravity coming from directly above, the otolith organ will detect head tilt and acceleration as producing the same resultant gravitational vector.

One of the primary roles of the vestibular system is stabilization of the retinal image. This allows humans to simultaneously move through their environment and be able to continue to see what they are looking at. In normal circumstances, vision provides approximately 85% of our orientation information; unfortunately, the aviation environment cannot be described as normal. The dynamic milieu of flight can produce problems with image stabilization, and visual acuity starts to fall when the velocity of inappropriate eye movements exceeds 3° to 5°/second. Clearly, if there was no compensatory reflex, then head movements, which tend to be of high frequency, would disturb vision. The relatively slow retinal processing (around 70 ms) cannot compensate for head movement, but the vestibular system via the vestibulo-ocular reflex (VOR) with latency of less than 16 ms can. The VOR has been quantified in the three dimensions of pitch, roll and yaw and has been shown to relate back to the gravitational vertical rather than the head orientation after perturbation, indicating a central processing function after the pure reflex of the VOR (Angaleski and Hess, 1994; Merfield et al., 1993). In other words, stabilization of the VOR functions for Earth-fixed but not head-fixed targets (Cheung, 2004). This presents a problem in flight as the aircraft movements are often of high rate and amplitude resulting in the breakdown of the combination of vestibulo-ocular and optokinetic mechanisms and the destabilization of the retinal image. For instance, prolonged rotational stimulation with a sudden cessation can lead to rotational nystagmus and the disruption of vision. Nystagmus is the flicking of the eye in a plane of movement and is named for the fast phase. In an initial acceleration to the right, the eyes drift to the left (slow phase). As the rotation continues, the eyes flick back to the right (fast phase) and then start the slow drift to the left again. This is a normal part of the stabilization mechanism of gaze, also seen in ice skaters. Unfortunately, when the sustained turn is stopped, there is an initiation of a reverse nystagmus due to the inability of the vestibular system to detect sustained velocity. This reverse nystagmus can be severe enough to make fixation on instruments or the outside scene impossible and lead to significant spatial disorientation.

Tactile and proprioceptive contribution to orientation

The somatosensory system is a widespread and diverse sensory system comprised of the receptors and processing centers that process touch, temperature, proprioception (body position), and nociception (pain). It consists of cutaneous tactile sensors and proprioceptors in muscles, ligaments, tendons and joint capsules. Together these sensors provide information on the body's orientation with respect to gravity and influence numerous force-feedback loops that help to determine conscious and unconscious muscular action to maintain that orientation. They are also fundamental in the everyday activities of movement and fine motor control. Thus, the output from these sensors is both pervasive and powerful and led aviators to the impression that they could "fly by the seat of their pants." Unfortunately, these somatosensors are as vulnerable to confusion in the flight environment as is the vestibular system. Consider an aircraft accelerating in the longitudinal axis at 2G, the tactile sensation on the skin of the pilot's back is exactly the same as in an aircraft accelerating vertically at 1G. In addition to this basic problem, the nature of the sensors themselves can produce erroneous orientation information. One special sensor in skeletal muscle is called the spindle, which is particularly sensitive to stretch. The amount of its activity is crucial to our knowledge of where our limbs are. Experiments have shown that under conditions of vibration (Goodwin, McCloskey and Matthews, 1972) or high G turning (Lackner and Levine, 1979), the spindles produce erroneous information leading to an uncertainty as to the position of our limbs. There are several other sensory inputs from muscle, tendon and joint capsule that all detect stretch, compression and activity and therefore provide information to central processing. All of these can and do provide inaccurate information under one or other circumstance of the flight environment. In the event that a pilot has no dependable visual references, the somatosensory and vestibular systems both provide information that can either confuse or override the other. The unusual force patterns of some flight maneuvers that stimulate vestibular illusions also stimulate somatosensors to provide the brain with an altered perception of orientation and even a changed feel of the aircraft controls (Lackner and Dizio, 1989).

All of the somatosensors are likely to produce misleading information to the brain when exposed to the altered gravitational environment of flight, particularly in rotary-wing flight with six degrees of freedom. A system produced by evolution to provide powerful unconscious information regarding body position and motor control while in the 1G environment on the earth's surface may not function appropriately in an aircraft. The very power of these stimuli is a cause of disorientation in flight without visual cues and should convince any pilot that they cannot fly "by the seat of their pants."

Auditory contribution to orientation

The human auditory system gathers sound information and passes it to the auditory cortex in the brain for interpretation. One of the system's functions is to localize sound and this is achieved by detecting the differences in sound intensity and time of arrival incident on the ears, all facilitated by the shape of the external ear (Batteau, 1974). The predominant portion of sound that contributes to its localization is the low frequency sound and its interaural time difference (Wightman and Kistler, 1992), which is that portion of the audible sound spectrum that is heavily masked in fixed and rotary-wing aircraft. However, individuals with noise induced high tone hearing loss, such as many aircrew, do not have impairment of sound localization as long as the sound remains audible (Lorenzi, Gatehouse and Lever, 1999) (see Chapter 5, *Audio Helmet-Mounted Displays*). In spite of these problems, efforts have been made to create virtual auditory displays. These are systems that produce signals reaching the listeners ears similar to those that would arise from real sources around the listener (Shinn-Cunningham, 1998). Spatial information has usually been presented using visual displays because the spatial acuity of the visual channel is much greater than the auditory channel. However, the visual channel is often overloaded in the flying task, and virtual auditory displays have been shown to marginally reduce workload when used to provide additional orientation information (McKinley and Ericson, 1997). In addition they have been found to be useful in monitoring numerous radio channels at once (Gardner, 1995), the so called "cocktail party effect." An argument can be made that this easing of communications difficulty also could offload the aircrew somewhat allowing more capacity for other tasks. However, the overall auditory contribution to aircrew orientation is likely to remain small because of limited auditory acuity and the potential for multi-sensory interference in a high task workload environment.

Visual contribution to orientation

The visual system in the presence of good visual cues predominates in providing the human with orientation information. Estimates of the extent of this predominance concur at approximately 85%, and this preponderance over the other sensory systems is known as visual dominance (Howard, 1982). This predominance of visual orientation inputs arises from three major factors: Firstly the visual environment is 3-D, and inputs can come from anywhere around us and not just from a point source, such as an auditory signal. Secondly, the precision of the visual system allows high resolution, especially in the central area of vision. Thirdly, the output from the visual system does not habituate like that of the vestibular system when exposed to steady state motion but continues to provide accurate information concerning the spatial layout of the environment over time. This last function is crucial in accurately assessing relative motion (Previc and Ercoline, 2004). The visual system does have limitations with respect to other sensory modalities, largely because it is relatively slow. The processing of a visual signal takes approximately 100 ms, whereas a vestibular signal is processed in a tenth of that time. This is why humans are much more effective at tracking crossing targets with head movement rather than eye movements alone with a stationary head.

There are two fundamental modalities in the visual system (Leibowitz and Dichgans, 1982) familiar to any pilot: focal mode and ambient mode. The focal mode is concerned with fine discrimination, focus and object recognition, essentially the "what." The focal mode is primarily driven by the central area and is represented by the light being focused by the eye on the fovea centralis, the small area of the retina rich in cone cells. The ambient mode is concerned with orientation in space, essentially the "where." This mode is primarily concerned with the light falling on the rest of the retina outside the fovea, an area populated largely by rod cells with some cones to provide color perception. The ambient vision responds to large stimuli such as horizon, sun position and immediate terrain, which permits running without falling over whilst tracking a target with focal vision. There has been postulation of further complication of the perception of 3-D space (Previc, 1998), but this goes beyond the scope of this text. Another important factor to note about these two visual modes is their method of central processing. The focal mode is a largely conscious, with attention required to interpret the scene incident on the fovea. The ambient mode, however, is almost entirely unconscious, allowing for intuitive orientation without active continuous cognition. The difference in processing also accounts for the difficulty that many pilots have in gaining their orientation information from the aircraft instruments, and why instrument flying is a learned skill that must be practiced for proficiency. These issues that occur in instrument flying also apply to many of the symbology sets presented in HMDs; this will be discussed in more detail later.

As previously discussed, the primary role of the ambient visual system is to provide us with our position in 3-D space, and to do this a number of different cues are used: The vividness of a visual scene is important with brightness, contrast and sharpness of texture appearing to diminish in the distance (Figure 12-41).
.

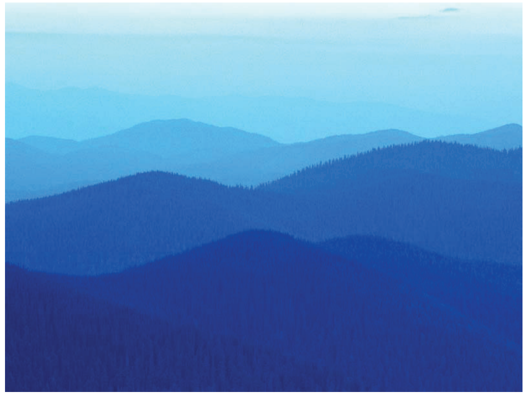


Figure 12-41. Detail and color diminish with distance

The ambient visual system also contributes to the estimation of distance in the areas of perspective and compression. Perspective is familiar in the art world with painters using the construct of a “vanishing point” to align sight lines within a painting to a notional point in the distance (Figure 12-42).

Figure 12-42. Pietro Perugino's usage of perspective in this fresco at the Sistine Chapel (1481– 82) helped bring the Renaissance to Rome.

Compression is the optical tendency for detail in a visual scene to appear closer together in the distance, for instance the ties on a set of railroad tracks (Figure 12-43).

Of all the ambient cues to distance estimation perspective is probably the most important (Sedgwick, 1986), although it does not completely dominate other cues. This is particularly important for pilots approaching a runway where they have a perspective model of what an approach should look like. If the runway is sloped or has a different width to length ratio than previously experienced, then the false perspective can lead to misjudgment of height over the ground (Figure 12-44).

Figure 12-43. Compression and perspective on a set of railroad tracks.

Figure 12-44. The importance of perspective in determining runway slope.

Other key cues are those of motion flow and parallax, which interact with linear perspective to allow the perception of relative motion. All of these cues act within the framework of the surface of the earth and the gravitational vertical supplied visually by the celestial bodies. All of these ambient visual cues are effectively monocular, as anything more than 6 meters (20 feet) away is close to optical infinity, and therefore the incident light rays are effectively parallel. This has important implications for the design of optical displays that attempt to utilize the unconscious ambient system for orientation. A good illustration of the importance of ambient vision in orientation is the autokinetic illusion. In this illusion, a watcher fixates on a small light in an otherwise darkened space. After some minutes, the light will appear to move in a random fashion as the watchers perception of their own orientation breaks down.

The visual world changes as humans move through it, and this change is consistent with the observer's 3-D vector. This apparent motion of the visual surroundings is termed the optical flow field (Gibson, 1966). The optical flow rate of objects in the visual field during motion of the observer can be described as:

$$\text{rate } \beta = (v / r) \sin \beta \qquad \text{Equation 12-2}$$

where β is the angle of the target from the center of the visual field in degrees, v is the forward velocity, and r is the radial distance to the target. It can be seen that this formula expresses the real world phenomenon that the farther an observer is away from something the slower it appears to be moving. Thus, a jet pilot flying at 400 knots and 200 feet (61 maters) perceives the ground passing at the same rate as a helicopter pilot flying at 120 knots and 60 feet (18.3 meters). This can cause problems for pilots who use this visual cue to judge their height above ground, as their perception will vary with their own velocity. A very familiar illusion is caused by the optical flow field; if there is motion in the field which creates the same retinal image as self-motion through space then there is a misperception of movement. The everyday demonstration if this would be when sitting on a stopped train in a station while another train moves away from the next platform. The perception of movement in the opposite sense to a movement in the optical field is called vection and was first described by Ernst Mach in the late 19^{th} century. Vection can be a particular problem in helicopters hovering close to the ground. The rotor downwash produces movement on the surface of water, in dust/snow or even on a cornfield. This apparent movement is rarely completely radial from the machine and can give rise to a very strong sensation of motion that the pilot may attempt to correct, resulting in drift and a possible accident. Many helicopters have ended flights on their sides as a result of blowing snow or dust and the vection illusion combining with the whiteout or brownout to rob the pilot of a stable visual reference. Vection appears to be a function of the periphery of vision and experiments have shown that the stimulus must be outside the central 50° of regard (Previc and Neel, 1995). Indeed if the object moving in the periphery is concentrated on then the vection disappears. The final feature of vection that is of interest to those involved in SD work is that slower stimuli have a tendency to produce a stronger vection response (Berthoz, Pavard and Young, 1975). The vection response drops off at angular velocities of 60°/sec (1 Hertz), if one imagines the train moving away from the station then there is an initial surge of vection followed by a diminution as the train gets faster. As well as increasing speed diminishing it, the vection response tends to habituate between 10 and 20 seconds after starting, thus making the first few seconds the most disorienting and therefore the most dangerous.

Another interesting function of the ambient visual system is that perception of speed is influenced by the nature of what is going past in the optical field (Denton, 1980). A driver on a desert highway will consistently underestimate the vehicle's speed, whereas a driver on a road through a forest will consistently overestimate. This is due to the relative richness of the visual field and can lead to disorientation and accidents in ground vehicles. The same phenomenon tends to not be an issue for fixed wing pilots but very low level helicopter pilots do encounter it.

The predominance of the ambient visual system in framing the human perception of orientation can be problematic in other ways. In some special cases such as those of the Ames rooms, an illusion of a perceptual framework can override the objective evidence provided by the focal vision (Dwyer, Ashton and Broerse, 1990) (Figure 12-45).

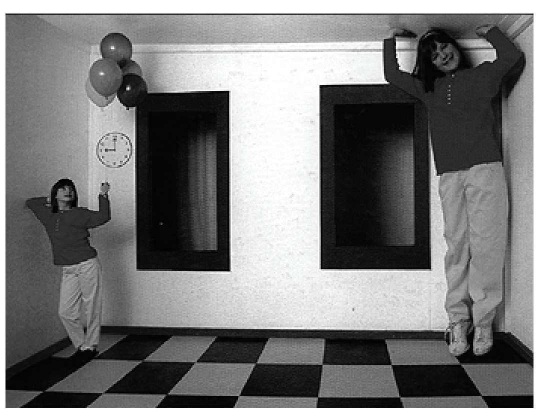


Figure 12-45. The Ames room illusion (Dwyer, Ashton and Broerse, 1990).

This ambient mode dominance often produces problems in flight. For instance, in mountain flying with no true horizontal horizons and the well-known illusion caused by flying above a flat cloudscape which is a few degrees off true horizontal. In ground vehicles where the outside peripheral scene is obscured this phenomenon can lead to underestimation of a slope that the vehicle is about to traverse as the inside of the vehicle is providing the only stable reference.

Cortical processing contribution to orientation

The vestibular, somatosensory and ambient visual systems discussed thus far all produce output that we process very largely subconsciously; only the focal visual system and auditory cues are mainly processed consciously. Thus most of our orientation information is received, processed and acted upon entirely without thought and is a very powerful tool in everyday life. Humans can exist without some of these inputs and the brain has enough plasticity to overcome significant deficits but in a fully functioning body there remains the possibility and even likelihood of disorientation given the right (or wrong) circumstances. Conscious thoughts can override unconscious ones as can be demonstrated by the learned skill of instrument flying but this is not a robust situation and the balance can easily reverse. This breakdown of the conscious primacy has been demonstrated in a flight simulator (Leduc et al., 2000), where performance measures during recovery from simulated disorientation were found to be significantly degraded after sleep deprivation. Physiological responses are known to be altered during fatigue, sleep deprivation, high workload, anxiety, excessive heat or cold and increased altitude. Many or all of these factors are present in a significant number of today's military missions both for aircrew and ground operators, resulting in increased disorientation in these environments (Bushby, Holmes and Bunting, 2005; Curry and McGhee, 2007).

In basic terms the ambient visual system orients the human in 3-D space, all three subconscious systems then combine to deliver information on the movement of the body through that space and the orientation of the body and limbs throughout. This is then overlaid by the conscious mind which in turn refers to the mental models that person has built up through their lifetime. These mental models can be regarded as experience, the high-hour instrument pilot has a well developed model of what should be happening and in most cases this will aid the control of the aircraft. There are circumstances where the mental model can interfere with cortical processing, for instance the pilot who lands with the wheels up after having 'checked' three greens. These are known as cognitive lapses but are really the result of accepting ingrained and usually highly useful mental models and are very rarely performed by inexperienced pilots as their mental models are not as well developed (Swauger 2003).

The fundamental problem faced by pilots and some ground vehicle operators is that some or all of their subconscious sensory systems can be relaying erroneous information some or all of the time. Thus most or all of their orientation information must come from vision, a system not immune to problems, and this sifting of unreliable and reliable information is a significant cognitive burden. All of this is compounded by the very strong evolutionary pressure to believe the senses and the associated mental models accumulated through experience.

Spatial disorientation and helmet-mounted displays (HMDs)

The HMD was defined in Chapter 3 (*Introduction to Helmet-Mounted Displays*) with a useful diagram reproduced below (Figure 12-46). This will form the framework of this section as each facet of the whole is examined. This diagram is labeled as specific to Army Aviation visual HMDs but it could as easily refer to any HMD in the aviation or ground environments. Until recently helmet trackers were exclusive to aircraft but a rudimentary device has been deployed in armored vehicles to control a slewing external camera system. The two major areas of the HMD system that can predispose to SD are the image source and the display, although the interaction of both with the helmet tracker can also be of importance.

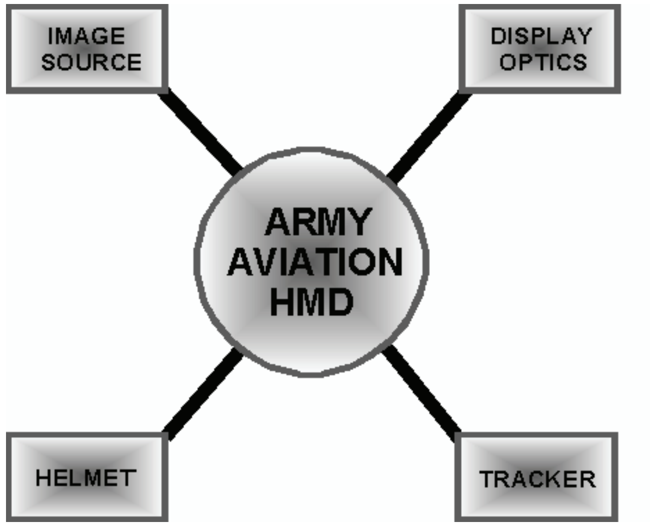


Figure 12-46. Block diagram of a basic U.S. Army rotary-wing aviation HMD (from Chapter 3, *Introduction to Helmet-Mounted Displays*).

SD and HMDs – General principles

HMDs have been around since the 1970s if NVGs are considered as true HMDs. NVGs mostly only present the outside scene and do not require a head tracker to function, but they will be considered in this discussion of SD issues of HMDs as they are generally attached to a helmet and certainly can produce episodes of SD. The more recent branches of the HMD family grew out of a desire to allow targeting information to be displayed to a pilot when looking off-axis of their craft rather than at a vehicle mounted display. Both of these HMD branches are growing more complex; with the presentation of synthetic images, flight instrument symbology, novel flight displays and sophisticated artificial targeting environments.

Particular types of HMD will be discussed later in this chapter but there are some general principles relating to what is presented to the eye that will be explored first. These principles relate to limitations of the technique and technology of HMDs and their likelihood to cause or worsen SD.

The picture of the outside world

All HMDs produce a representation of the world outside the person, vehicular crew station, or cockpit. The design can be as simple as viewing (or looking through) a transparent screen onto which symbology is projected or by viewing a fully synthetic outside view. The method by which the view is attained can have fundamental effects on the likelihood of the viewer to suffer SD. The very simplest see-though HMDs should have very little effect on the view outside the cockpit if they are optically correct. Even in a monocular display, the outside image should be easily accepted, but can be complicated by symbology projected onto the HMD, an important point to be discusser later.

The early HMDs, as previously noted were NVGs, and these have produced a host of SD problems since their acceptance into wide usage (DeLucia and Task, 1995). The image from the modern NVG is a binocular, image intensified picture of the night environment. The picture presented to the eye is formed by an image intensifier which transmits in the 600 to 900 nanometer range, the visible red and near infra-red part of the spectrum. This alone can produce illusions specific to NVGs with red lights appearing closer and bluer lights further away. In the past, bright lights would produce a halo effect and one bright light in the visual scene could washout the rest of the image. These along with acuity issues are largely being defeated by the newer versions and the likelihood of crashing whilst using NVGs is decreasing (Antonio, 2000).

Another SD producing problem with NVGs is their FOV, a circular 40° x 40°. Thus, most of the visual scene is not visible without exaggerated scanning head movements. An attempt has been made to alleviate this problem with the Panoramic NVG (PNVG) (Figure 12-47), a device that extends the field of view to 100° x 40°. Fields-of-view are compared in Figure 12-48. Early operational results (Geiselman and Graig, 2002) have shown fixed-wing pilots have an increased SA and task performance when wearing PNVGs over regular NVGs. The U.S. Special Forces community has tested PNVGs and with a few minor caveats has accepted them as a flight safety enhancement in their particular rotary-wing missions.

An additional and complicated SD issue when viewing the world through NVGs is depth perception. There have been reports of depth perception problems in 11% of US Army helicopter pilots (Crowley, 1991) and up to 30% in a U.S. Air Force survey (Baldin et al., 1999). The optical focus of NVGs is toward infinity, which could cause depth perception problems to about 6 meters (20 feet) away from the pilot, although this is rarely of significance. There is debate about how important the binocular cues gained through NVGs are to depth perception; some laboratory studies have suggested that there is very little stereopsis present and that viewers tend to underestimate distance through NVGs (Wiley, 1989). In any event, once the objects being viewed are over 50 meters (20 feet) away the incident light is effectively parallel, and thus the eyes are essentially viewing the same image and the NVGs become equivalent to a bi-ocular display.

Figure 12-47. Panoramic Night Vision Goggles (PNVG).

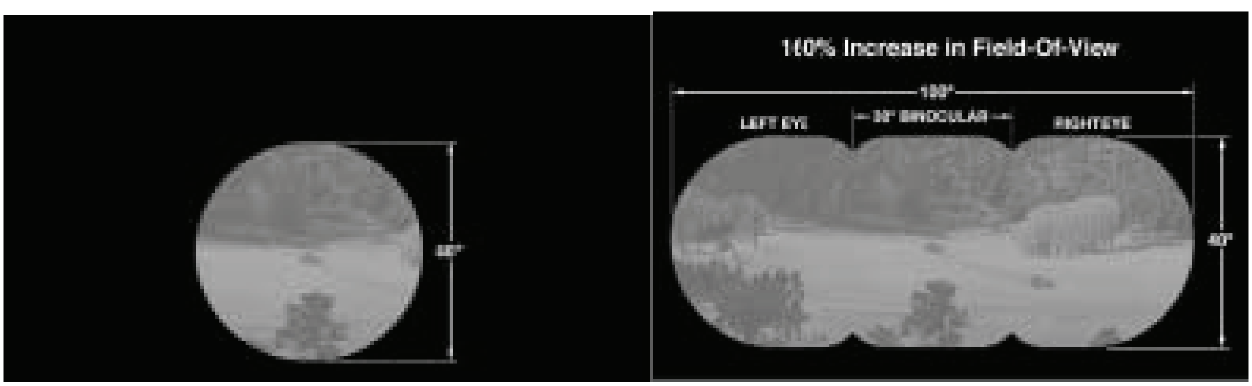


Figure 12-48. Comparison of standard NVG (40° x 40°) vs. PNVG (100° x 40°) fields-of-view.

The next generation of HMDs came about largely as a result of a desire to display targeting information to the user when they were looking off-axis of their aircraft. The early versions projected standard targeting information that had previously only been available when viewing on-axis through a HUD. This technology was found to be a success, and HMDs became a growth area in aircraft systems. The objective of the modern HMD approach is to provide continuous dynamic information that can be monitored while the observer is free to maintain visual contact with the surrounding environment. The see-through HMD has been discussed as has the direct vision (NVG) type. The depiction of the outside scene particularly at night has also been achieved by the use of forward-looking infrared (FLIR), low-light cameras, millimeter radar, etc. to provide an image for the user. The potential for these images to cause SD can come from the position of the sensor relative to the pilot, the image produced by the sensor, the way it is projected in front of the pilot's eye, and more recently how images from different sensors are fused together.

Monocular HMDs

The primary example of a monocular HMD is the IHADSS (Figure 12-49) as used in the AH -64 Apache aircraft (see Chapter 3, *Introduction to Helmet-Mounted Displays*). This display projects a variety of sensor data and flight/targeting symbology in front of the right eye of the pilot and co-pilot/gunner. The sensor images come from a pod mounted on the nose of the aircraft approximately 8 feet (2.4 meters) ahead of the pilot and 3 feet (0.9 meter) below his design eye height. The image can be FLIR, image intensified camera, or a terrain mode from the millimeter-radar mounted on the mast above the rotor system.

Figure 12-49. The monocular Integrated Helmet and Display Sighting System (IHADSS) HMD.

There are several important potential causes of SD in the IHADSS, the most immediately recognized being binocular rivalry (see Chapter 12, *Visual Perceptual Conflicts and Illusions*). With a sensor image plus symbology being presented to the right eye and the outside view or the internal cockpit being presented to the left there is a tendency for one or the other to be attended to. This is known and an attempt is made in training to alleviate the problem in the 'bag phase' by removing the stimuli to the left eye other than internal instruments. Students very often find this phase of training tough and many wash out. Evidence shows that trained pilots still show a degree of binocular rivalry and several hover accidents have been attributed to a lack of attention to the hover vector symbology presented to the right eye (Braithwaite, Groh and Alvarez, 1997). There are other problems specific to the Apache which are illustrative of the issues that face designers of these types of systems and reinforce the requirement to think of all elements of the structure.

The FLIR image used for night pilotage requires a difference in radiant energy from within the environment. This is usual, but in certain temperate conditions, a phenomenon called "thermal crossover" occurs where the background and atmosphere have the same thermal signature. In this circumstance, the pilot effectively becomes blind, losing all scene contrast, and must use instruments to recover the aircraft. Another design issue that has caused SD is the gimballing of the nose sensor, which when traversed from one lateral extreme to the other, produces an image that appears to describe a gentle curve with the apex higher than the periphery. This illusion leads to attempted corrections, particularly when in the hover, by the handling pilot. Both these problems have now been rectified to some extent but were implicated in several SD accidents before they were.

Binocular HMDs

The difference between biocular and binocular displays was discussed in Chapter 3 (*Introduction to helmet-Mounted Displays*), but one interesting development in recent years has lead to concern in the area of hyperstereopsis. In the late 1990s, during the U.S. Army Comanche reconnaissance helicopter program, the helmet design incorporated NVG tubes into the helmet structure itself. These tubes were coupled to an HMD that was also integral to the helmet. One advantage of this design approach is to provide a capability for both image intensification and FLIR imagery; center-of-mass and head-supported weight is also improved. One example of this design approach is the Thales Avionics TopOwl™ system (Figure 12-50). It is used in the European Tiger attack helicopter and has been adopted by multiple countries. However, this design, sometimes referred to as a hyperstereo design, introduces some unique considerations for SD, particularly hypstereopsis (see Chapter 12, *Visual Perceptual Conflicts and Illusions*).

Figure 12-50. The TopOwl™ HMD (Thales Avionics).

The hyperstereopsis issue associated with the NVG tubes being mounted on the sides of the head and therefore creating a greater than normal effective eye separation distance affects distance perception and perspective. Hyperstereo designs and their visual effects have been investigated for their potential use in helicopters (Armburst et al., 1993; Kalich et al., 2007). While some issues require additional study, some data do suggest that pilots can develop strategies to compensate for the hyperstereopsis phenomenon and may even be able to achieve some level of adaptation. It is still open to question as to if performance on certain flight tasks is improved or degraded by the use of hyperstereo HMDs.

A final word on SD and HMDs

All HMDs provide the viewer with a view of the outside world. Almost all also provide the viewer with targeting information and also with flight symbology. The latter is a rich source of potential SD with a particular emphasis on how aircraft attitude information is displayed in an HMD. There are two basic methods of displaying attitude information; the conformal method displays the horizon as it actually is wherever the viewer is looking. Thus if a pilot is looking in any direction in good Visual Meteorological Conditions (VMC) then the projected horizon would overly the real horizon. The non-conformal display provides the pilot with the view as if he/she were looking on-axis out of the front of the aircraft whatever the orientation of the pilot's head and eyes. There is considerable disagreement about which method is superior in terms of providing SA and avoiding SD. In addition to the conformal/non-conformal debate there is also a good deal of dispute about how to present the attitude information to the eye. The two major ones are the inside-out display where the depiction of the aircraft is steady and the horizon moves, and the outside-in display where the horizon stays steady and the aircraft symbol moves (Jenkins, 2003). Further, there are novel displays such as the 'Grapefruit' (Ercoline, Self and Matthews, 2002), the Arc-Segmented Attitude Reference (ASAR) (Wickens et al., 2007) and the Oz (Still and Temme, 2007), among many others.

The standard research methodology for assessing these various display configurations has been the rate of Control Reversal Errors (CREs) during recovery from an unusual attitude, usually in a flight simulator. Unfortunately many of the various studies contradict one another with a degree of partiality towards whichever display type is the product of that organization. There has been a suggestion that pilots trying to determine their orientation from HMD symbology are sometimes confused and this produces a delay in recovery from an unusual position or control reversals in doing so (Ligget and Gallimore, 2002). This may provide new ideas for the design of HMD symbology that could reduce SD by referencing the theoretical underpinnings of normal pilot orientation, in particular by using the ambient visual system. This approach could also utilize the potential for HMDs to have a large field-of-regard by being placed close to the eye.

A question asked by a Swedish group (Eriksson and von Hofsten, 2002) was whether visual displays can be constructed in such a way as to convey the crucial information that supports spatial orientation. HMD technology allows a large field-of-regard and the potential to provide peripheral cues. Work by Kappé (1997; Kappe, van Erp and Korteling, 1999) used an HMD with a detailed image in the frontal direction surrounded by a sparse peripheral image in a driving simulator. He found that the peripheral displays had a clearly beneficial effect on driver orientation and performance even though the peripheral displays contained relatively small amounts of information. This idea of providing peripheral cues in an HMD format have been pursued by several groups with varying methodologies varying from simple horizon lines in the periphery to full novel displays such as OZ – a computerized system that provides pilots with a symbolic picture of flight status without requiring slow instrument reading in a conventional manner (Still and Temme, 2007). This display utilizes a relatively sparse peripheral field with both horizon lines and "star-field" vection cues to produce the visual sensation of motion (Figure 12-51).

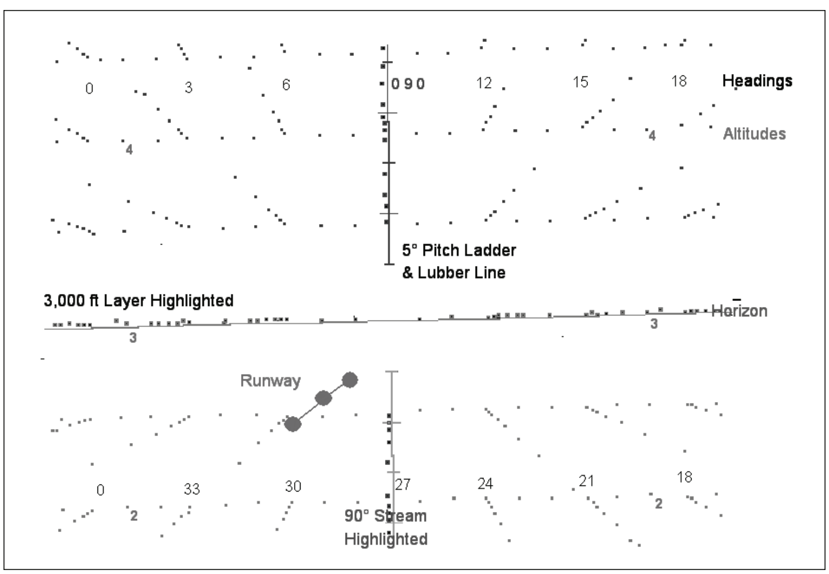


Figure 12-51. The Oz display (Still and Temme, 2007).

In summary, it would seem that HMDs have the potential to make the problem of SD worse or possibly to make it dramatically better. The designers of these devices would do well to look at the basic neurophysiology of orientation in humans. Early pilots found that they could not "fly by the seat of their pants," and flight became much safer with the advent of the standard instrument panel in the twenties. All pilots know however, that instrument flight is a learned skill that must be practiced and one to which our visuo-motor function is ill-suited. The possibility exists that well designed HMDs can make flight in poor visual conditions much safer by using the intuitive and hard-wired human orientation mechanisms evolved for life on the ground.

Luning

The FOV of HMDs is limited by the size of the optics. This historically has limited the FOV of a single ocular to approximately 40°. For a binocular system, when both eyes see the identical full image, the HMD is known as having a fully-overlapped FOV. If for design reasons, the size of the monocular fields are at a maximum and cannot be increased without incurring unacceptable costs such as reduced spatial resolution, or increased size and weight of the optics, then the size of the fully-overlapped FOV may not be sufficient.

In order to increase the extent of the visual world available via an HMD to Warfighters (especially for aviators), an optical approach known as partial binocular overlap has been explored. In a partially-overlapped design, the wider FOV consists of three regions---a central binocular overlap region seen by both eyes and two flanking monocular regions, each seen by only one eye (Figure 12-52). There are perceptual consequences for displaying the FOV to the human visual system in this unusual way. These perceptual effects have been a concern to the aviation community because of the potential loss of visual information and the visual discomfort (Alam et al., 1992; Edgar et al., 1991; Kruk and Longridge, 1984; Landau, 1990; Melzer and Moffitt, 1989).

Partially-overlapped binocular displays contain binocular overlap borders, which in terms of the FOV separate the binocular overlap region and the monocular regions. In terms of the monocular fields, these borders separate the portion exclusively seen by one eye from the portion seen in common with the other eye. In normal unencumbered vision, the binocular overlap borders, dividing the natural FOV, are not experienced explicitly (see Gibson, 1979, for a good discussion) and are only cognitively identified and located with attentional effort. However, in artificial viewing situations such as HMDs, where the monocular fields are smaller than in natural viewing, these borders are accompanied by a perceptual effect that in the display literature has come to be known as *luning* (CAE Electronics, 1984; Moffitt, 1989).

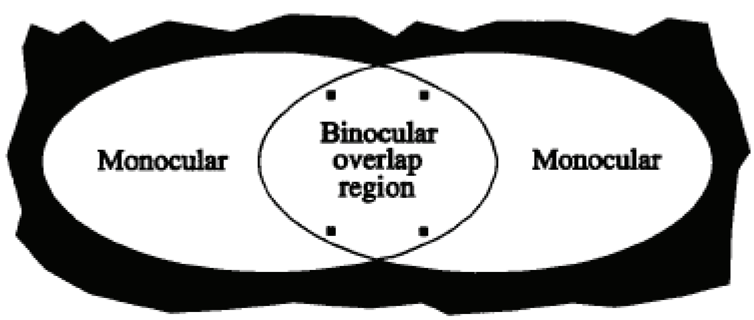


Figure 12-52. The partially-overlapped FOV mode with a central binocular overlap region seen by both eyes and two flanking monocular regions (Klymenko et al., 1994),

Luning is a visual perception characterized by a subjective darkening of the visual field in the monocular regions of partial binocular overlap displays. It was so named (Moffitt, 1989) because of the crescent shapes of the darkened monocular regions adjacent to the circular binocular overlap region. It is most pronounced near the binocular overlap border separating the monocular and binocular regions, gradually fading with increasing distance from the border. The prominence of luning fluctuates over time and appears not to be strongly under attentional control (see Figure 12-53).

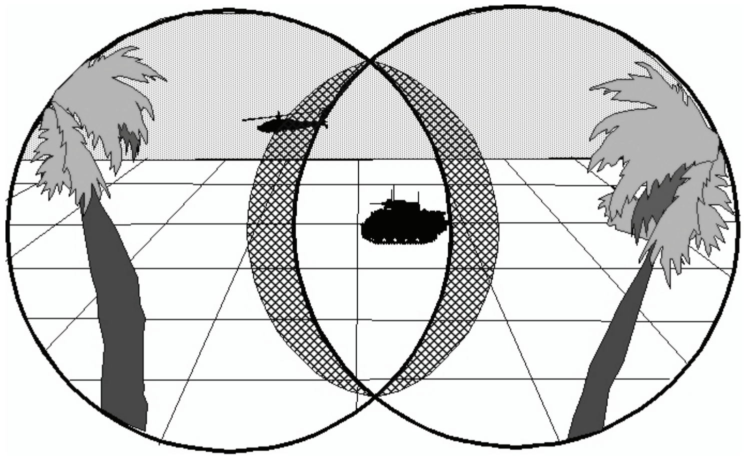


Figure 12-53. Luning in partial-overlap HMDs (Rash, 2001).

Luning may be related to binocular rivalry and suppression. Binocular rivalry refers to the alterations in the appearance of a binocular stimulus which is dichoptic, i.e., where each eye's image alternately dominates the phenomenal binocular FOV by suppressing the other eye's input. Over time, one and then the other eye may successfully compete and dominate awareness. Suppression refers to the phenomenal disappearance of one eye's input due to monocular dominance by the other eye. Partial suppression refers to the partial disappearance of one eye's input. In the partial binocular overlap display mode, each eye's monocular region is the result of dichoptic competition between a portion of its monocular field and the other eye's monocular field border and dark background. If the background is completely suppressed, the total FOV looks natural, where the binocular and monocular regions are both seen as one continuous visual world. If an eye's monocular region is partially suppressed by the dark background of the other eye, then this dark background will appear in monocular regions of the first eye with the greatest darkening – luning – occurring near the binocular overlap border.

In the monocular regions of partial binocular overlap displays, both the dichoptic differences in luminance and the presence of the monocular edge—the luminance drop--at the binocular overlap border likely affect luning. This luminance transition between the monocular field and the background occurs in what we shall refer to as the *noninformational* eye. During fusion it is matched to a region within the monocular field of the *informational* eye. There are a number of interocular inhibitory processes in addition to binocular rivalry of dichoptic stimuli (Fox, 1991), which may also contribute to luning (e.g., see Gur, 1991, on Ganzfeld fade-out and blackout, and Bolanowski and Doty, 1987, on blankout). Binocular rivalry and the interocular inhibitory process of suppression due to rivalry between dichoptic stimuli is can be one working hypothesis of luning. There are different types of binocular rivalry including piecemeal dominance, binocular superimposition, and binocular transparency (Yang, Rose and Blake, 1992). Binocular transparency describes the percept when both dichoptic stimuli are seen simultaneously, but appear "scissioned," or segregated in depth; superimposition describes the situation in which both dichoptic stimuli appear to occupy the same space; and piecemeal dominance refers to small isolated parts of each eye's image dominating the binocular percept. Since luning is a change in apparent brightness (a darkening of a region), which can spread or recede overtime, this particular occurrence of binocular rivalry (see Kaufman, 1963) theoretically appears also to be related to the ubiquitous contrast, and color, spreading phenomena (see Grossberg (1987) for a catalogue and neural net theory of such phenomena), such as neon color spreading (see Nakayama, Shimojo and Ramachandran, 1990). Luning appears to emanate from the binocular overlap border and is attenuated by placing physical contours in the location of this border, that is, in the location within the homogeneous monocular field of the informational eye that binocularly corresponds to the edge of the monocular field of the noninformational eye (Melzer and Moffitt, 1991).

A potential ecological overview of the luning phenomena incorporates what recently has recently come to be known as DaVinci stereopsis (Nakayama and Shimojo, 1990). First extensively studied in modem times by Barrand (1979), DaVinci stereopsis refers to binocular occlusion, which refers to the situation in which an object in the FOV, such as one's nose, may occlude only one eye's view of more distant objects (see Gillam and Borsting, 1988). Explaining luning based on DaVinci stereopsis requires us to first analyze the optical geometric constraints imposed by the real world on an observer (see Melzer and Moffitt, 1991). That is, what real world situation, such as viewing through an aperture or viewing past an object in front of the face, corresponds to the artificial display mode of the HMD that causes luning? The visual system may have natural responses to these situations. For example, the tendency to suppress the foreground region of an aperture may be one such response. Also, there may be no one real world situation which perfectly corresponds to an HMD display, thus leading to conflicting visual responses. There are a number of potential ecologically salient visual geometric configurations one could evoke for each type of artificial display situation; however, only recently have researchers begun to examine the visual system's natural tendencies to interpret a viewing situation in terms of these real world configurations(e.g., see Nakayama, Shimojo and Silverman, 1989; Shimojo and Nakayama, 1990).

Klymenko at al., (1994) investigated factors that affect the perception of luning in the monocular regions of partially-overlapped HMDs. These factors included: (1) the convergent versus the divergent display modes for presenting a partial binocular overlapping FOV (Figure 3-3, Chapter 3, *Introduction to Helmet-Mounted Displays*), (2) the display luminance level, (3) the placement of either black or white contours versus no (null) contours on the binocular overlap border (Figure 12-54), and (4) the increasing or decreasing of the luminance of the monocular side regions relative to the binocular overlap region. Eighteen Army student aviators served as subjects in a repeated measures design. The percentage of time luning was seen was the measure of the degree of luning. The results indicated that the divergent display mode systematically induced more luning than the convergent display mode under the null contour condition. Adding black contours reduced luning in both the convergent and divergent display modes, where the convergent mode retained its relatively lower magnitude of luning. The display luminance level had no effect on luning for the null or black contour conditions. Adding white contours reduced luning by an amount that depended on display luminance where there was less luning for lower display luminance levels, but no systematic effect of display mode. Changing the luminance of the monocular regions (relative to the binocular overlap region) reduced the amount of luning, where a decrease in luminance produced more of a reduction in luning than an increase. When a partial binocular overlap display is needed to present a larger FOV to aviators in HMDs, the convergent display mode with black contours on the binocular overlap borders appears to be the most reliable of the conditions tested to systematically reduce luning.

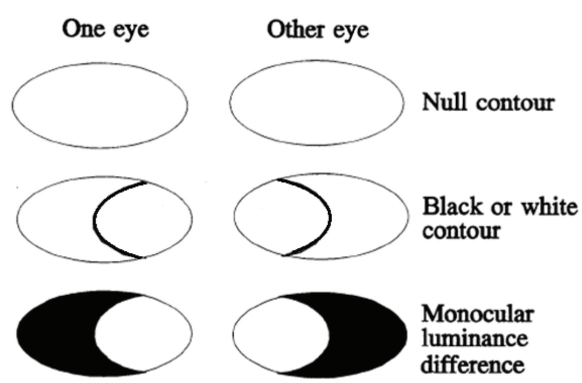


Figure 12-54. Use of border contours by Klymenko et al. (1994) to investigate luning.

References

Addams, R. (1834) An account of a peculiar optical phenomenon seen after having looked at a moving body. *London and Edinburgh Philosophical. Magazine and Journal of Science*, 5, 373–374.

Alam, M.S., Zheng, S.H., Iftekharuddin, K.M., and Karim, M.A. (1992). Study of field-of-view overlap for night vision applications. Proceeding of the 19.92 IEEE National Aerospace and Electronics Conference, NEACON, 3, 1249-1255. Dayton, OH.

Alais, D., and Blake, R. (1999). Grouping visual features during binocular rivalry. *Vision Research*, 39(26), 4341-4353.

Alais, D., and Blake, R. (Eds.). (2005). *Binocular Rivalry*. Cambridge, MA: The MIT Press.

Albery, W. (2006). Development of Multisensory Orientation Technology for Reducing Spatial Disorientation Mishaps. AFRL-HE-WP-TR-2006-0024.

Alpern, M (1953). Metacontrast. *Journal of the Optical Society of America*, 43, 648-657.

Ames, A. (1951). Visual perception and the rotating trapezoidal window. *Psychological Monographs: General and Applied*, 65, No. 324.

Andrews, T.J., Sengpiel, F., and Blakemore, C. (2005). From contour to object-face rivalry: multiple neural mechanisms resolve perceptual ambiguity. Chapter 11. In: *Binocular Rivalry*, Alais, D., and Blake, R., (Eds.), Cambridge, MA: MIT Press.

Angaleski, D.E., and Hess, B.J.M. (1994). Inertial Representation of Angular Motion in the Vestibular System of Rhesus Monkeys. I. Vestibuloocular Reflex. *Journal of Neurophysiology*, 71, 1222-1249.

Anstis, S.M. (1974). Letter: A chart demonstrating variations in acuity with retinal position. *Vision Research*, 14(7), 589-592.

Antonio, J.C. (2000). Reducing NVG-Related Spatial Disorientation Mishaps. *Recent Trends in Spatial Disorientation Research*. November 15 – 17, San Antonio, Texas.

Armbrust, J., Ros, N., Hale, S., and Rabin, J. (1993). Final report, developmental test (DT) of the Night Vision Pilotage System. Fort Rucker, AL: U.S. Army Aviation Technical Test Center. TECOM Project No. 4-AI-100-RAH-008.

Aulhorn, E., and Harms, H. (1972). Visual perimetry. In: *Handbook of Sensory Physiology: Visual Psychophysics VII/4*. Jameson, D., and Hurvich , L.M. (Eds.), New York, NY: Springer-Verlag. p 102-145.

Bachmann, T. (1994). *Psychophysiology of Visual Masking: The Fine Structure of Conscious Experience*. Commack, NY: Nova Science Publishers, Inc.

Baldin, J.B., Dennis, R.J., Ivan, D.J., Miller, R.E., III, Belihar, R.P., Jackson, W.G., Jr, Tredici, T.J., Datko, L.M., and Hiers, P.L. (1999). The 1995 Aircrew Operational Vision Survey: Results, Analysis, and Recommendations. USAF School of Aerospace Medicine, SAM-AF-BR-TR-1999-0003, Brooks Air Force Base, Texas.

Barrand, A.G. (1979). An ecological approach to binocular perception: the neglected facts of occlusion. Doctoral dissertation, Cornell University.

Batteau, D.W. (1964). *The Role of the Pinna in Human Sound Localization*. Cambridge, MA: United Research Inc.

Benson, A.J. (1978). Spatial Disorientation: General aspects. In: *Aviation Medicine. Dhenin*, G. (Ed.). London: Tri-Med Books.

Benson, A.J. (2002). Spatial Disorientation – A Perspective. Part of: *Spatial Disorientation in Military Vehicles: Causes, Consequences and Cures*. RTO-MP-086.

Berthoz, A., Pavard, B., Young, L.R. (1975). Perception of linear horizontal self-motion induced by peripheral vision (Linear vection). *Experimental Brain Research*, 23, 471-489.

Blake, R., (2008). Retrieved from Vanderbilt University binocular rivalry demonstration website: http://www.psy.vanderbilt.edu/faculty/blake/rivalry/BR.html

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind*, 2, 5-38.

Blake , R. (1988). Dichoptic reading: the role of meaning in binocular rivalry. *Perception and Psychophysics*, 44, 133-141.

Blake, R., and Boothroyd, K. (1985). The precedence of binocular fusion over binocular rivalry. *Perception and Psychophysics*, 37(2), 114-124.

Blake, R., and Levinson, E. (1977). Spatial properties of binocular neurons in the human visual system. *Experimental Brain Research*, 27, 221-232.

Blake, R., and Logothetis, N.E. (2002). Visual competition. *Nature Reviews Neuroscience*, 3, 13-21.

Blake, R. (2005). Landmarks in the history of binocular rivalry. Chapter 1. In: *Binocular Rivalry*, Alais, D., and Blake, R. (Eds.), (p. 1- 27). Cambridge, MA: MIT Press.

Blake, R., Sloane, M., and Fox, R. (1981). Further developments in binocular summation. *Perception and Psychophysics*, 30, 266-276.

Blake, R., Westendorf, D.H., and Overton, R. (1980). What is suppressed during binocular rivalry? *Perception*, 9, 223-231.

Bolanowski, S.J., and Doty, R.W. (1987). Perceptual "blankout" of monocular Homogeneous fields (Ganzfelder) is prevented with binocular viewing. *Vision Research*, 27, 967-982.

Boring E.G. (1942). *Sensation and Perception in the History of Experimental Psychology*. New York, NY: Appleton-Century-Crofts.

Braithwaite, M.G., Groh, S., and Alvarez, E. (1997). Spatial disorientation in U.S. Army helicopter accidents: An update of the 1987-92 survey to include 1993-95. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Technical Report 97-13.

Breitmeyer, B.G. (1984). *Visual Masking: An Integrative Approach*. New York: Oxford University Press

Breitmeyer, B.G., and Ogmen H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision*. New York: Oxford University Press.

Brown, J.F. (1931). The visual perception of velocity. *Psychologische Forschung*, 14, 199-232.

Bushby, A., Holmes, S., and Bunting, A. (2005). An assessment of the influence of spatial disorientation upon British military aircraft accidents. *Aviation, Space, and Environmental Medicine*, 76(3), Sect 1, 217-219.

CAE Electronics. (1984). Wide-field-of-view, helmet-mounted infinity display system development. Brooks Air Force Base, TX: Air Force Systems Command. AFHRLTR- 84-27.

Cagenollo, R., Arditi, A. and Halpern, D.L. (1993). Binocular enhancement of visual acuity. *Journal of the Optical Society of America A*, 10, 1841-1848.

Campbell, F.W., and Green, D.G. (1965). Monocular versus binocular visual acuity. *Nature*, 208, 191-192.

Cheung, B. (2004). Nonvisual Spatial Orientation Mechanisms. In: *Spatial Disorientation in Aviation*. Vol. 203. *Progress in Astronautics and Aeronautics*.

Chong, S. C., and Blake, R. (2006). Exogenous attention and endogenous attention influence initial dominance in binocular rivalry. *Vision Research*, 46(11), 1794-1803.

Clark, B. (1967). Threshold for the perception of angular acceleration in man. *Aerospace Medicine*, 38, 443-450.

Crawford, B.H. (1947). Visual adaptation in relation to brief conditioning stimuli. *Proceedings of the Royal Society of London*, 134B, 283-302.

Crewther, D.P., Jones, R., Munro, J., Price, T., Pulis, S., and Crewther, S. (2005). A neural network model of top-down rivalry. Chapter 18. In: *Binocular Rivalry*, Alais, D. and Blake, R., (Eds), (p. 337-353). Cambridge, MA: MIT Press.

Cline, D., Hofstetter, H .W., and Griffin, J. R. (Eds.) (1980). *Dictionary of Visual Science*. Radnor, PA. Chilton Company.

Coren, S., and Girgus, J.S. (1978). Visual llusions. In: *Handbook of Sensory Physiology: Perception VIII*. Held, R., Leibowitz, H.W., and Teuber, H.L. (Eds.). New York, NY: Springer-Verlag.

Crowey, A., and Rolls, E.T. (1974). Human cortical magnification factor and its relation to visual acuity. Experimental Brain Research, 21(5), 447-454.

Crowley, J.S. (1991). Human Factors of Night Vision Devices: Anecdotes from the Field Concerning Visual Illusions and Other Effects. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No: 91-15.

Crowley, J.S., Rash, C.E., and Stephens, R.L. (1992). Visual illusions and other effects with night vision goggles. *Proceedings of SPIE*, *Helmet-Mounted Displays III*, 1695, 166-180.

Crozier, W.J., and Holway, A.H. (1939). Theory and measurement of visual mechanisms. I. A visual discriminometer. II. Threshold stimulus intensity and retinal position. *Journal of General Physiology*, 22, 341-364.

Curry, I.P, and McGhee, J.S. (2007). Spatial Disorientation as a cause of Mishaps in Combat Helicopter Operations. Presented to the *Aerospace Medical Association Annual Meeting*. New Orleans, LA.

Cutting, J.E., and Vishton, P.M. (1995). Perceiving layout and knowing distances: the integration, relative potency, and contextual use of different information about depth. In: *Perceptual Space*. Epstein, W., and Rogers, S. (Eds.). New York: Academic Press. p. 69-117.

de Weert, C.M.M., Snoeren, P.R., and Koning, A. (2005) Interactions between binocular rivalry and Gestalt formation. *Vision Research*, 45, 2571-2579.

DeLucia, P.R., and Task, H.L. (1995). Depth and collision judgment using Night Vision Goggles. *International Journal of Aviation Psychology*, 5(4), 371-386.

Denton, G. G. (1980). The Influence of Visual Pattern on Perceived Speed. *Perception*, 9, 393-402.

Dichgans, J., and Brandt, T. (1978). Visual-vestibular interaction: Effects on self-motion perception and postural control. In: *Handbook of Sensory Physiology: Perception VIII*. Held, R., Leibowitz, H.W., and Teuber, H.L. (Eds.). New York, NY: Springer-Verlag. p. 755-804.

Duncker K (1929) Über induzierte Bewegung (Ein Beitrag zur Theorie optisch wahrgenommener Bewegung). *Psychol Forsch*, 12, 180-259.

Durnford, S.J., Crowley, J.S., Rosado, N.R., Harper, J., and DeRoche, S. **(**1995**).** Spatial Disorientation: A survey of U.S. Army helicopter accidents 1987-92. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Technical Report 95-25.

Dwyer, J., Ashton, R., and Broerse, J., (1990). Emmert's law in the Ames room. *Perception*, 19(1), 35-41.

Edgar, G.K., Carr, K.T., Williams, M., and Clark A.L. (1991). The effect upon visual performance of varying binocular overlap. AGARD Symposium on Helmet-Mounted Displays and Night Vision Goggles. *AGARD Conference Proceedings*, 517, 8-1 to 8-15.

Epstein, W. (1995). The metatheoretical context. In: *Perception of Space and Motion*, Epstein, W., and Rogers, S.J. (Eds.). San Diego: Academic Press Inc.

Ercoline, W.R., Self, B.P., and Matthews, R.S.J. (2002). Effects of three helmet-mounted display symbologies on unusual attitude recognition and recovery. *Aviation, Space and Environmental Medicine*, 73(11), 1953-1958.

Eriksson, L., and von Hofsten, C. (2002). On the Possibility of Couteracting or Reducing G-Induced Spatial Disorientation With Visual Displays. Presented to the *RTO HFM symposium, Spatial Disorientation in Military Vehicles: Causes, Consequences and Cures*, La Coruna, Spain. 15-17 April 2002, and published in RTO-MP-086.

Exner, S. (1875). *Pflügers Arch. ges. Physiol.* 11, 403–432.

Fox, R. (1991). Binocular rivalry. In: *Vision and Visual Dysfunction*, 9, 93- 110. Boca Raton, FL: CRC Press, Inc.

Freeman, A.W., Nguyen, V.A., and Alais, D. (2005). The nature and depth of binocular rivalry suppression. Chapter 3. In: *Binocular Rivalry*, Alais, D., and Blake, R. (Eds.). Cambridge, MA: MIT Press.

Frijda, N.H. (2006). *The Laws of Emotion*. London: Routledge. p. 47-62.

Gardner, W.G. (1995). Transaural 3-D audio. MIT Media Laboratories Perceptual Computing Section, Technical Report No. 342.

Geiselman, E.E., and Craig, J.L. (2002). Panoramic Night Vision Goggle Update. Wright-Patterson Air Force Base, OH: U.S. Air Force Research Laboratory. Technical report, AFRL-HE-TR-2002-0083.

Geudry, F.E. (1974). Psychophysics of vestibular sensation. In: *Handbook of Sensory Physiology*, 6, 3-154, Kornhuber, H.H. (Ed.). Berlin: Springer-Verlag.

Gibson, J.J. (1959). *The Perception of the Visual World*. Boston: Houghton Mifflin.

Gibson, J.J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA : Houghton-Mifflin.

Gillam, B. (1981). False percpectives. *Perception*, 10, 3113-318.

Gillam, B., and Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, 17, 603-608.

Goodwin, G.M., McCloskey, D.I., and Matthews, P.B.C. (1972). The contributions of muscle afferents to kinasthesia shown by vibration induced illusions of movement and by the effects of paralyzing joint afferents. *Brain*, 95, 705-708.

Graybiel, A., and Clark, B. (1945). The autokinetic illusion and its significance in night flying. *Journal of Aviation Medicine*, 16, 111-151.

Gregory, R. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society B*, 352, 1121-1128.

Gregory, R.L. (1996). What are illusions. *Perception*, 25(5), 503.

Gregory, R.L., and Zangwill, O.L.(1963). The origin of the autokinetic effect. *Quarterly Journal of Experimental Psychology*, 15, 255-261.

Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness Perception: II. Binocular theory. *Perception and Psychophysics*, 4, 117-158.

Gur, M. (1991). Perceptual fade-out occurs in the binocularly viewed Ganzfeld. *Perception*, 20, 645-654.

Harding, T.H., and Rash C.E. (2004). Potential see-through performance deficits in U.S. Army developmental helmet-mounted displays. *Proceedings of the SPIE, Helmet- and Head-Mounted Displays IX: Technologies and Applications*, 5442, 273-279.

He, S., Carlson, T, and Chen, X. (2005). Parallel pathways and temporal dynamics in binocular rivalry. Chapter 5. In: *Binocular Rivalry*, Alais, D., and Blake, R. (Eds.), (p. 82-100), Cambridge, MA: MIT Press.

Heinecke, J.K., Rash, C.E., Ranaudo, R., and Hiatt, K.L. (2008). Apache aviator evaluation of dual-technology night vision systems in Operation Iraqi Freedom (OIF) urban combat (Master’s Thesis). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2008-05.

Helmholtz, H. (1909). *Handbuch der Physiologischen Optik*. 3rd Edition, Vol. 3 (translated by Southall [1962]).

Hershberger, M. L., and Guerin, D. F. (1975). Binocular rivalry in helmet-mounted display applications. Dayton, OH: Armstrong Aerospace Medical Research Laboratory. AMRL-TR-75-48.

Hiatt, K.I., Rash, C.E., Harris, E.S., and McGillberry, W.H. (2004). Apache aviator visual experiences with the IHADSS helmet-mounted display in operation Iraqi Freedom. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2004-20.

Hixson, W.C., and Spezia, E. (1977). Incidence and cost of orientation-error accidents in regular Army aircraft over a five-year study period: Summary report. Joint Army/Navy report NAMRL-1238/USAARL Technical Report 77-19.

Holmes, S.R., Bunting, A., Brown, D.L., Hiatt, K.L., Braithwaite, M.G., and Harrigan, M.J. (2003). Survey of spatial disorientation in military pilots and navigators. *Aviation, Space, Environmental Medicine*, 74(9), 957-65.

Holway, A.H., and Boring, E.G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology*, 54, 21-37.

Howard, I.P. (1982). *Human Visual Orientation*. New York: Wiley.

Howard, I.P. (2002). Binocular fusion and rivalry. Chapter 7. In: *Seeing in Depth*, Howard, I.P., and Rogers, B.J. (Eds.). Ontario, Canada: I. Porteous. p. 272-315.

Howard, I.P. (2005). Binocular rivalry and the perception of depth. Chapter 10. In: *Binocular Rivalry*, Alais, D., and Blake, R. (Eds.). Cambridge, MA: MIT Press. p. 169-186.

Jenkins, J.C. (2003). Development of helmet-mounted display symbology for use as a primary flight reference. *Proceedings of SPIE, Helmet-Mounted Displays*, 5079, 333-345.

Johnson, C.W. (2004). The role of night vision equipment in military incidents and accidents. *Human Error, Safety and Systems Development*. Johnson, C.W, and Palanque, P. (Eds.). Kluwer Academic Publishers.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: The University of Chicago Press.

Kahneman, D. (1968). Methods, findings, and theory in studies of visual masking. *Psychological Bulletin*, 70, 404-425.

Kalich, M.E., Rash, C.E., McLean, W.E., Ramiccio, J.G. (2007). A limited flight study for investigating hyperstereo vision. *Proceedings of the SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications* 6557, 65570I.

Kanizsa, G. (1979). *Organization in Vision: Essays on Gestalt Perception*. New York: Praeger.

Kappé, B. (1997). Visual Information in Virtual Environments. PhD thesis, University of Utecht, and TNO, Netherlands. ISBN 9039318921.

Kappé, B, van Erp, J, and Korteling, J.E. (1999). Effects of head-slaved and peripheral displays on lane-keeping performance and spatial orientation. *Human Factors*, 41(3), 453-466.

Kaufman, L. (1963). On the spread of suppression and binocular rivalry. *Vision Research*, 3, 401-415.

Kenkel, F. (1913). Untersuchungen uber den Zusammenhang zwischen Erscheinungsgrobe und Erscheinungsbewegung bei einigen sogenannten optischen Tauschungen. *Zeitschrift für Psychologie*, 67, 358–449.

Klymenko, V., Verona, R.W., Martin, J.S., Beasley, H.H., and McLean, W.E. (1994a). Factors Affecting the Perception of Luning in Monocular Regions of Partial Binocular Overlap Displays. For Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 94-47.

Klymenko, V., Verona, R.W., Martin, J.S., Beasley, H.H., and McLean, W.E. (1994b). Factors affecting the perception of luning in monocular regions of partial binocular overlap displays. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 94-47.

Korte, A. (1915). Kinematoskopische Untersuchangen. *Zeitschrift fiir Psychologie*, 72, 193-206.

Kovács, I., Papathomas, T.V., Yang, M. and Fehér, A. 1997: When the brain changes its mind, Interocular grouping during binocular rivalry, *Proceedings of the National. Academy of Sciences, USA*. 93, 15508-15511.

Kruk, R., and Longridge, T.M. (1984). Binocular overlap in a fiber optic helmet mounted display. *The Image 3 Conference Proceedings*, 363, 363-377. Brooks Air Force Base, TX: Air Force Human Resources Laboratory. Air Force Systems Command. AFHRL-TR-84-36.

Lackner, J.R., and Dizio, P. (1989). Gravitational Effects on Nystagmus and Perception of Orientation. Representation of Three Dimensional Space in the Vestibular, Oculomotor and Visual Systems, Cohen, B., and V. Henn, V., (Eds.). *Annals of the New York Academy of Science*, 545, 93-104.

Lackner, J.R., and Levine, M.S. (1979). Changes in apparent body orientation and sensory localization induced by vibration of postural muscles: Vibratory myasthetic illusions. *Aviation, Space and Environmental Medicine*, 50, 346-354.

Landau, F. (1990). The effect on visual recognition performance of rregistration and overlap for a biocular helmet-mounted display. *Proceedings of SPIE, Helmet-Mounted Displays II*, 1290, 173-184.

Laramee, R.S., and Ware, C. (2002). Rivalry and interference with a head mounted display. *ACM Transactions on Computer-Human Interaction*, 9(3), 238-251.

Leduc, P.A., Riley, D., Hoffman, S.M., Brock, M.E., Norman, D, Johnson, P.A., Williamson, R., and Estrada, A. (2000). The Effects of Sleep Deprivation on Spatial Disorientation. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Technical Report 2000-09.

Lee, S-H., and Blake, R. (1999). Rival ideas about binocular rivalry. *Vision Research*. 39(8), 1447-1454.

Leopold, D.A., Maier, A, Wilke, M., and Logothetis, N.K. (2005). Binocular rivalry and the illusion of monocular vision. Chapter 13. In: *Binocular Rivalry*, Alais, D. and Blake, R. (Eds.), (p. 232-258). Cambridge, MA: MIT Press.

Liebowitz, H.W., and Dichgans, G. (1982). The Ambient VISUAL System and Spatial Orientation. Spatial Disorientation in Flight: Current Problems, AGARD, Neuilly sur Seine, France, p. B4-1 to B4-4.

Levine, M., and Shefner, J. (1991). *Fundamentals of Sensation and Perception*, Pacific Grove, CA: Brooks/Cole Publishing Company.

Ligget, K.L., and Gallimore, J.J. (2002). An analysis of control reversal errors during unusual attitude recoveries using helmet-mounted display symbology. *Aviation, Space, and Environmental Medicine,* 73(2), 102-111.

Long, G.M., and Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: Alternative views of reversible figures. *Psychological Bulletin*, 130(4), 748-768.

Lorenzi, C., Gatehouse, S., and Lever, C. (1999). Sound localization in noise in hearing-impaired listeners. *Journal of the Acoustical Society of America*, 105(6), 3454-3463.

Lythgoe, R. J. and Phillips, L. R. (1938). Binocular summation during dark adaptation. *Journal of Physiology*, 91, 427-436.

MacLeod, D.I.A. and Willen, J.D. (1995). Is there a visual space? In: Luce, R.D., Hoffman, D.D., D'Zmura, M.D., Iverson, G.J., and RomneyA.K., (Eds., *Geometric representations of perceptual phenomena: Papers in honor of Tarow Indow on this 70th Birthday*. Mahway, N.J. Erlbaum.

McKinley, R.L., and Ericson, M.A. (1997). Flight Demonstration of a 3-D auditory display. In: *Binaural and Spatial Hearing in Real and Virtual Environments*, Gilkey, R., and Anderson, T., (Eds.). New York: Erlbaum. p. 665-682.

Meiry, J.L. (1965). The vestibular System and Human Dynamic Space Orientation. Masters Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Melzer, J. E., and Moffitt, K. (1989). Partial binocular-overlap in helmet-mounted displays. *Proceedings of SPIE, Display System Optics II*, 1117, 56-62.

Melzer, J.E., and Moffitt, K. (1991). An ecological approach to partial binocular-overlap. *Proceedings of SPIE, Large Screen Projection, Avionic, and Helmet-Mounted Displays*, 1456, 175-191.

Merfield, D.M., Young, L.R., Paige, G.D., and Tomko, D.L. (1993). Three-dimensional eye movements of squirrel monkeys following postrotatory tilt. *Journal of Vestibular Research*, 3, 123-139.

Miller, R.E., and Tredici, T. J. (1992). Night vision manual for the flight surgeon (AL-SR-1992-0002). Brooks AFB, TX: Armstrong lab.

Moffitt, K. (1989). Luning and target detection. (Company working document.) San Jose, CA: Kaiser Electronics.

Nakayama, K., and Shimojo, S. (1990). Da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Research*, 30, 1811-1825.

Nakayama K., Shimojo, S., and Ramachandran, V.S. (1990). Transparency: relation to depth, subjective contours, luminance, and neon color spreading. *Perception*, 19, 497-513.

O'Shea, R. (2009). Binocular rivalry. Retrieved on 20 Jan 2009: http://psy.otago.ac.nz/r_oshea/BR_bibliography/br_bibliography.html.

Papathomas, T.V., Kovacs, I., and Conway, T. (2005). Interocular groupings in binocular rivalry: basic attributes and combinations. Chapter 9. In: *Binocular Rivalry*, Alais, D., and Blake, R. (Eds.), (p. 155-168).Cambridge, MA: MIT Press.

Patterson, R., Winterbottom, M., and Pierce, B. (2007). Binocular rivalry and head-worn displays. *Human Factors*, 49(6), 1083-1096.

Pinkus, A.R., and Task, H.L. (2004). Night Vision Goggle Luminance Disparity and the Pulfrich Phenomenon. *Proceedings of SPIE, Helmet- and Head-Mounted Displays IX*, 5442, 54-63.

Previc, F.H., and Neel, R.L. (1995). The effects of visual surround eccentricity and size in manual and postural control. *Journal of Vestibular Research*, 5, 399-404.

Previc, F.H. (1998). The Neurophysiology of 3-D space. *Psychological Bulletin*, 124, 123-164.

Previc, F.H., and Ercoline, W.R. (2004). Spatial disorientation in aviation. In: *Progress in Astronautics and Aeronautics*. Reston, VA: American Institute of Aeronautics and Astronautics.

Purkinje, J.E. (1825). Beobachtungen und Versuche zur Physiologie der Sinne. *Neue Beitrage zur Kenntniss des Sehens in Subjektiver Hinsicht* (Berlin: Reimer).

Rabin, J. (1995). Two eyes are better than one: binocular enhancement in the contrast domain. *Ophthalmic and Physiological Optics*, 15(1), 45-48.

Rash, C.E. (2008). A 25-year retrospective review of visual complaints and illusions associated with a monocular helmet-mounted display. *Health and Safety Aspects of Visual Displays, Displays*, 29(2), 70-80.

Rash, C.E., and Hiatt, K.L. (2005). Apache flight experience in operation Iraqi Freedom (OIF) shows reduced incidents of visual illusions and problems with the monocular helmet-mounted display. *American Helicopter Society 61st Annual Forum*, Grapevine, TX, June 1-3.

Rash, C.E., and Martin, J.S. (1988). The impact of the U.S. Army's AH-64 helmet-mounted display on future aviation helmet design. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 88-13.

Rash, C.E., Reynolds, B.S., Stelle, J.A., Peterson, R.D., and LeDuc, P.A. (2004). The role of helmet-mounted displays in A"H-64 Apache accidents. *U.S. Army Medical Department Journal*, April-June, 7-16.

Rash, C.E., Suggs, C.L., Mora, J.C., van de Pol, C., Reynolds, B.S., and Crowley, J.S. (2002). Visual Issues Survey of AH-64 Apache aviators (year 2000). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2002-02.

Rash, C.E., Verona, R., and Crowley, J. (1990). Human factors and safety considerations of night vision systems flight using thermal imaging systems. *Proceedings of SPIE, Helmet-Mounted Displays II*, 1290, 142-164.

Robinson, J.O. (1998). *The Psychology of Visual Illusions*. Mineola, NY: Dover Publications.

Rock I., Hill, A., and Fineman, M. (1968). Speed constancy as a function of size constancy. *Perception and Psychophysics*, 4, 37-40.

Rogers, S. (1995). Perceiving pictorial space. In: *Perception of Space and Motion,* Epstein, W. and Rogers, S. (Eds.). San Diego: Academic Press.

Schumann, F (1987). Contributions to the analysis of visual perception – First paper: Some observations on the combination of visual impressions into units (Translated into English by Anne Hogg). In: *The Perception of Illusory Contours*, Petry, S., and Meyer. G. (Eds.). New York: Springer-Verlag.

Sedgwick, H.A. (1986). Space perception. In: *Handbook of Perception and Human Performance*, Boff, K.R., Kaufman, L., and Thomas, J.P (Eds.), 21.1-21.57.

Sekuler, R., and Blake, R. (2005). *Perception*. New York: McGraw-Hill.

Shimojo, S., and Nakayama, K. (1990). Real world occlusion constraints and binocular rivalry. *Vision Research*, 30, 69-80.

Shinn-Cunningham, B. (1998). Applications of virtual auditory displays. *Proceedings of the 20th International Conference of the IEEE Engineering in Biology and Medicine Society*. 20(3), 1105-1108.

Smallman, H.S., Manes, D.I., and Cowen, M.B. (2003). Measuring and modeling the misinterpretation of 3-D perspective views. *Proceedings of the Human Factors and Eergonomics Society 47th Annual Meeting*, 1615-1619.

Smith, E.E., and Kosslyn, S.M. (2007). *Cognitive Psychology: Mind and Brain*. Upper Saddle River, NJ: Pearson Education, Inc.

St. John, M., Cowen, M.B., Smallman, H.S., and Oonk, H.M. (2001). The use of 2D and 3D displays for shape understanding versus relative-position tasks. *Human Factors*, 43, 79-98.

Still, D.L., and Temme, L.A. (2007). Hover training display: Rationale and implementation. *Proceedings of SPIE-The International Society for Optical Engineering*, 6956, 695607-695607-15.

Swauger, S. (2003). How Good Pilots Make Bad Decisions: A Model for Understanding and Teaching Failure Management to Pilots. *12th Symposium on Aviation Psychology*. April 14-17, Dayton, OH.

Ternus, J. (1926). Experimentalle Untersuchung uber phanomenale Identitat" *Psychologische Forsch*, 7, 81-136. Translated as "The problem of phenomenal identity." In: *A Sourcebook of Gestalt Psychology,* Ellis, W.D. (Ed.), London: Routledge (1936/1999).

Toch, H.H. (1962). The effect of ‘meaning’ on the autokinetic illusion. *American Journal of Psychology*, 75, 605-611.

Tong, F., Meng, M., and Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, 10(11), 502-511.

Tong, F., Nakayama, K., Vaughan, J.T., and Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*. 21, 753-759.

U.S. Army Safety Center. (1991). NVG desert operations. Flightfax April, Ft. Rucker, AL, p. 3-7.

von Holst, E., and Mittlestaedt, H. (1950) Das Reafferenzprinzip. *Naturwissenschaften*, 37, 464–476.

Vrnwy-Jones, P. (1988). Disorientation accidents and incidents in U.S. Army helicopters 1 Jan 1980 to 30 Apr 1987. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Technical Report 88-3.

Wade, N.J. (2005). Ambiguities and rivalries in the history of binocular vision. Chapter 2. In: *Binocular Rivalry*, Alais, D., and Blake, R. (Eds.). Cambridge, MA: MIT Press.

Wade, N.J., and Heller, D. (2003). Visual motion illusions, eye movements, and the search for objectivity. *Journal of the History of the Neurosciences*, 12(4), 376-395.

Wagner, M. (2006). *The Geometries of Visual Space*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Wallach, H. (1959). The perception of motion. *Scientific American*, 201, 56-60.

Wertheimer, M. (1912). Experimentelle Studien fiber das Sehen von Bewegung. *Zeitschriftfiir Psychologie*, 61, 161-265.

Wheatstone, C. (1838). Contributions to the physiology of vision – Part one. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, 128, 371-394.

Wickens, C.D, Self, B.P, Andre, T.S, Reynolds, T.J, Ronald, L. (2007). Unusual Attitude Recoveries with a Spatial Disorientation Icon. *International Journal of Aviation Psychology*, 17(2), 153-165.

Wightman, F.L., and Kistler, D.J. (1992). The dominant role of low-frequency interaural time difference in sound localization. *Journal of the Acoustical Society of America*, 91(3), 1648-61.

Wiley, R.W. (1989). Visual Acuity and Stereopsis with Night Vision Goggles. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 89-9.

Winterbottom, M.D., Patterson, R., Pierce, B.J., and Rogers, J. (2007). Binocular rivalry and attention in helmet-mounted display applciations. Mesa, AZ: Air Force Research Laboratory. AFRL-HE-AZ-TP-2007-06.

Wolf, J.M., Kluender, K.R., Levi, D.M., and Bartoshuk, L.M. (Eds.) (2005). *Sensation and Perception*. Sunderland, MA: Sinauer Associates, Inc.

Yang, Y., Rose, D., and Blake, R. (1992). On the variety of percepts associated with dichoptic viewing of dissimilar monocular stimuli. *Perception*, 21, 47-62.

Yoshitaka, S., Yoshinori, T., and Hideo, T. (2003). Examination of the Pulfrich phenomenon by using a monocular HMD (Head Mounted Display). Japan Air Self-Defence Force, Aeromedical Lab. Reports of Aeromedical Laboratory, 43(1), 1-11.

Zalevski, A., Meehan, J.W., and Hughes, P.K. (2001). Size estimation with night vision goggles. Australian Government, Department of Defence Science and Technology Organization, Report No. DSTO-RR-0201.

13 AUDITORY CONFLICTS AND ILLUSIONS

Angélique A. Scharine
Tomasz R. Letowski

Perceptual Conflicts and Illusions

Perceptual conflicts appear when the brain receives ambiguous information and needs to choose which of the conflicting pieces of information represents the actual stimulation. In some cases, one piece of information dominates the other ones and we are largely unaware of the conflicting information. In other cases, however, all pieces of information are perceptually equivalent and the brain switches spontaneously between alternate interpretations of the received stimulation. If the brain receives several conflicting but perceptually equivalent pieces of information such a phenomenon is also called perceptual multistability and the acting stimulation is called multistable stimulation (Leopold and Logithesis, 1999). In case of two conflicting pieces of information, the phenomenon is called perceptual bistability and corresponding stimulation is called bistable stimulation (Hupe, Joffo and Pressnitzer, 2008).

Multistable stimulation has been extensively studied in the visual domain (e.g., moving plaids, binocular rivalry) but can also occur in the auditory modality. The most common form of multistable auditory perception is auditory grouping of incoming acoustic information that may change depending on the listener's focus on specific temporal events (Bregman, 1990; Hupe, Joffo and Pressnitzer, 2008; Pressnitzer and Hupe, 2006; Van Noorden, 1975).

The grouping of arriving sounds into perceptual events is referred to as *auditory streaming*. Normally, the grouping forms a uniform unequivocal stream. However, bistable streaming may occur under certain conditions. For example, an alternating sequence of high- and low-frequency tones may be perceived as one or two streams. When the tones are similar in frequency and the rate of presentation is slow, the listener hears a single coherent series of tones alternating in time. However, when the difference in frequency is large and the repetition rate is fast the alternating sequence splits perceptually in two unrelated streams of high- and low- frequency tones. Between these two extreme conditions, the listeners may hear either phenomenon by paying attention to different properties of the sequence. For example, Van Noorden (1975) presented the listeners with two tones A and B forming a sequence ABA-ABA-…ABA. Depending on the frequencies of the tones and the duration of the pause between individual ABA chunks, the listeners perceived either a single stream of information in a form of "a gallop" or two parallel streams of high- and low-frequency tones. However, within a certain range of manipulated parameters the listeners could hear either of these two phenomena just by refocusing their attention. Similar effects can be observed when the same tone with alternating intensities is presented to the listener (Van Noorden, 1975).

Perceptual conflicts may also have a multisensory form when information received through two or more senses lacks congruency. For example, several authors reported perceptual conflicts in simultaneous perception of conflicting visual and tactile cues (Adams and Duda, 1986; Heller et al., 1999; Hershberger and Misceo, 1996), and visual and auditory cues (Hupe, Joffo and Pressnitzer, 2008).

Another group of perceptual effects that are not directly dependent on the presence of external stimulation are perceptual illusions. Perceptual illusions are the instances where the cues that the brain relies on to provide specific information about sensory stimulation are poorly correlated with actual physical stimulation. They are not the instances when two incongruent pieces of information compete for our attention but rather the instances where our brain reports a stable and repeatable awareness of stimulation that cannot be directly explained by the physical properties of the acting stimuli. They are distortions of reality that are typically shared by many people (Solso, 2001). They can be explained by various physiological processes but they do not refer directly to the stimulation received and therefore are called illusions.

Perceptual illusions should not be confused with the masking phenomena described in Chapter 11, *Auditory Perception and Cognitive Performance,* where the part of the stimulus is not perceived due to the shadowing effect (masking effects) of the other parts of the stimulus. Masking does not result in a qualitatively new percept but only causes a partial awareness of the stimulus. Illusions are also different from auditory images or hallucinations, which are the sensations created in the absence of stimulus. For example, composers report "hearing a tune" in their head before writing a new piece. Hallucinations usually have a pathological basis but they may also occur occasionally in the real world when a highly expected event does not happen.

Auditory illusions are quite common in perception of music and speech due to our brain's tendency to fill the unexpected gaps in incoming streams of events by a reasonable prediction of what should be there. There are also some between-channel associations that may create illusions or hallucinations of the presence of specific acoustic stimulation that does not take place. For example, seeing lip movement in a noisy environment where no speech is present may result in the illusion of hearing speech. Another example of an auditory illusion is the McGurk effect, described in Chapter 14, *Auditory-Visual Interactions,* where seeing the lips pronouncing sound "ga" and hearing sound "ba" results in illusion of hearing the sound "da" (McGurk and McDonald, 1976). In this case a sound is present but it is heard as a different one.

The initial part of this chapter discusses of the processing of information by the auditory channel and the potential conflict in information reception. The second part describes common auditory conflicts and illusions and their physiological basis. Auditory-visual interactions and related conflicts, are discussed in Chapter 14, *Auditory-Visual Interactions,* and auditory-tactile interactions observed during tactile stimulation of the head are described in Chapter 18, *Exploring the Tactile Modality for HMDs*. Chapter 14 also contains a discussion of practical strategies intended to reduce auditory and auditory-visual conflicts and cognitive overload by proper design of auditory signals (earcons and warning signals) so that they are easily understood and complied with during times of stress and fatigue.

Auditory Scene Analysis

Auditory scene analysis (ASA) is the term coined by a Canadian psychologist Albert Bregman (1990) to describe a variety of processes by which the brain parses the sound arriving at the ear into its various components and groups them together into meaningful events. Each of our ears receives only a single sound pressure wave and this wave consists of the combination of all sounds occurring in the environment. The fact that we have two ears is critical for spatial perception and auditory orientation in space but each of the ears receives a relatively complete collection of auditory events emanating from an infinite variety of sound sources. The brain has the task of analyzing the complex waveform that arrives at the ears into its components and then assigning those components to the auditory events and sound sources creating those components. Some authors divide the ASA tasks into the simultaneous grouping (frequency grouping) and sequential grouping (stream grouping) tasks but in most practical situations both tasks are performed concurrently and aid each other (Plack, 2005).

ASA resembles the task of visual scene analysis performed by the sense of vision. Rather than view the world as a hodgepodge of colors and lights, our visual system follows a set of rules to determine which visual components belong to which visual object. It does this by utilizing a number of Gestalt cues. "Gestalt," a German word for "form," is used to refer to self-organizing principles that form a whole from a collection of features. Some examples of the Gestalt Laws that govern the ways by which auditory and visual objects are formed from their specific features are listed in Table 13-1. An example demonstrating how the Laws of Good Continuation, Simplicity, and Closure affect our visual perception is shown in Figure 13-1. The Laws of Good Continuation and Simplicity suggest that the picture is composed of blocks with simple straight edges rather than a strange object with jagged edges and that the jagged lines are due to a juxtaposition of multiple blocks. Further, the *Law of*

Closure suggests that the edges obscured by other blocks placed in front of them are continuous and form whole objects.

Table 13-1.
A description of auditory and visual Gestalt Laws.

Name	Audition	Vision
Proximity (Belongingness)	Sounds arriving from places *close in space* tend to be grouped	Elements *close together* in space tend to be grouped
Similarity	Sounds with *similar timbre and pitch* tend to be grouped	Elements *shaped alike* tend to be grouped
Good Continuation	Sounds that follow a *regular pitch contour* tend to be grouped	Elements that follow a *regular spatial contour* tend to be grouped
Closure	Interrupted auditory stimuli tend to be perceived as *continuous* when plausible	*Borders are interpreted/completed* to specify shapes
Simplicity (Pragnanz)	Frequencies with *simple harmonic ratios* tend to be grouped	*Prototypical* shapes tend to be regular, simple, symmetric
Common Fate	Sounds with *synchronous rhythm patterns* tend to be grouped	Elements that *move together* tend to be grouped

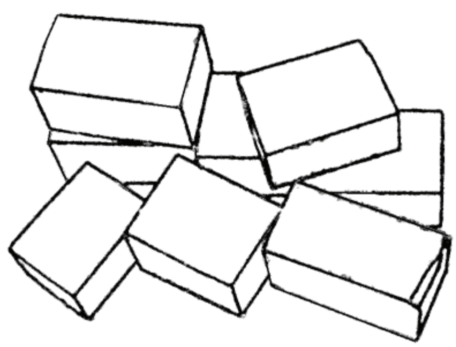


Figure 13-1. Illustration of the effect of three Gestalt Laws -*Law of Good Continuation, Law of Closure* and *Law of Simplicity* - on our perception. We interpret the drawing as a collection of simple blocks rather than a complex collection of random lines.

As shown in Table 13-1 the Gestalt Laws apply not only to vision but also to audition. Their function is to allow the auditory system to group the different features of the complex waveforms arriving at the ears into a plausible set of auditory events. Consider a common auditory environment, that of a kitchen with a radio on and a person cooking dinner while another person sets the table. There will be the sound of the radio, the sound of people talking, the sounds of dishes and pans, and the ambient sounds of the heating or cooling system. These will all arrive at the ears of the listener as two complex waveforms. In order to parse the waveform into the auditory events creating it, the listener will use many of the Gestalt Laws. For example, the *Law of Proximity* (Belongingness) suggests that anything coming from spatial location occupied by the radio is caused by the radio. The *Law of Similarity* suggests that sounds with a similar timbre (Chapter 11, *Auditory Perception and Cognitive Performance*) are caused by the same auditory event, for example, a particular person speaking. The *Law of Closure* allows a person to listen to the talker and understand him, despite occasional masking caused by the radio. The *Law of Simplicity*, also called the *Law of Pragnanz*, which is a German term meaning “good figure,” suggests that we will choose simple plausible interpretations, consistent with our knowledge of the environment. Thus, ringing sounds will be attributed to pans in the kitchen. Different instruments with differently pitched tones playing in a synchronous rhythm will be perceived as a single acoustic event, that is, the band that is playing on the radio because of the *Law of Common Fate*.

The Gestalt rules can be applied in numerous ways, but their primary function is always to help the listener parse the continuous sound wave into the individual sound events that created it. For the most part, this occurs automatically without much effort on the part of the listener and with very few errors. However, in some cases, the sound wave is parsed incorrectly, and sound components fuse together, emerging perceptually as sound objects that are not actually present. When this occurs, they result in auditory illusions. In almost all cases, auditory illusions exemplify a situation in which a Gestalt cue biases the perceptions of the listener. The selected illusions described in this chapter are dependent on the precise coincidence of certain spectral and temporal features; however, they illustrate how the Gestalt Laws work. Auditory events in the real world are somewhat more random, making such illusory events less predictable. However, a good grasp of Gestalt cues will aid in the design of auditory cues and warnings that are easily detected and understood despite masking from noise in the ambient environment. More extensive discussion of the auditory signal design is presented in Chapter 14, *Auditory-Visual Interactions.*

Auditory Conflicts and Illusions

Auditory conflicts and illusions are not clearly differentiated in the literature and sometimes one of these terms is used to describe both classes of phenomena. In some cases it is even difficult to differentiate if a specific perceptual effect should be classified as a conflict or an illusion. Similarly, none of the proposed classification of these effects, such as transmission-based and construction-based effects, is well established and intuitive enough to be included in this chapter. However, all these phenomena are various distorted perceptions of sound pitch, temporal properties, and location of a sound source. Therefore, for the purpose of clarity, they are grouped together in this chapter by the perceptual characteristic that is being distorted; that is, pitch, temporal pattern, and spatial phenomena.

Pitch conflicts and illusions

In general, humans equate pitch with the lowest frequency of a periodic sound (see Chapter 11, *Auditory Perception and Cognitive Performance*). This frequency is called the fundamental frequency of the sound. The physical nature of the object determines the dominant frequency and all the other accompanying frequencies. Most commonly, the dominant frequency is the fundamental frequency and the other frequencies are harmonics that follow a fairly regular pattern of multiples of the fundamental frequency. However, there are also cases where the perceived pitch does not correspond to the lowest frequency of a sound or where higher frequencies do not follow a strict harmonic pattern. In addition, there are some cases where pitch sensation does not follow any of the accepted forms of pitch creation or changes in an unpredictable manner. Such cases are usually referred to as pitch conflicts or pitch illusions.

Recent advances in electronics and sound synthesis make it possible to precisely control the amplitudes and phases of all frequency components of a sound thus allowing us to explore the way pitch is perceived. The first three illusions described in this section, p*eriodicity pitch*, c*ircular pitch* and the t*ritone paradox*, demonstrate the interacting effects of the fundamental and harmonic frequencies on pitch sensation. The illusions that follow, the *octave illusion* and other pitch streaming illusions, illustrate the preference of the auditory system for small intervals and a regular pitch contour. In general, the perceptual system attempts to group sound components into streams that can be easily interpreted and encoded. Therefore, in some cases, rather than parsing the sound wave in a manner consistent with the spatial information arriving to the left and right ear and creating two sound streams having complicated temporal patterns, the sound wave is parsed into two simple patterns neglecting spatial disparity. The conflicting spatial cues are then misperceived to agree with the dominant simplicity of the resulting streams. The next illusions, the s*plit-off illusion* and its derivatives, illustrate the Laws of Good Continuation and Simplicity. The perceptual system attempts to form a simple interpretation of two auditory streams fusing them into a single stream. Sound elements that are inconsistent with this interpretation are either ignored or “split-off” into an extraneous stream. The *Huggins pitch* emerges from a white noise signal because the phase information separates each narrow frequency band from the rest of the signal, essentially causing the listener to perceive a single stream as two separate ones.

Periodicity pitch

Sounds that are composed of several frequency components having a simple harmonic relationship are called tones. Tones that consist of a single frequency are called pure tones and tones that consist of several frequencies are called complex tones (Emanuel and Letowski, 2009). The lowest frequency (F0) of a complex tone is called the fundamental frequency, and the higher frequencies (F1, F2, F3,…) that are integer multiples of the fundamental frequency are called harmonics. The specific harmonic structure of a complex tone gives the tone its characteristic timbre (tone color).

If one presents a complex tone, with a fundamental frequency of F0 and a series of harmonics, the pitch of the tonal complex is normally associated with the pitch of the fundamental frequency F0. Adding or removing harmonics from the complex affects the timbre of the complex, but it does not change its pitch. This will remain true even if several of the first few harmonics are removed and – more remarkably – even if the fundamental frequency is removed from the complex. . Consider, for example, a case shown in Figure 13-2. A complex tone shown in panel (a) consists of a 400 Hertz (Hz) fundamental frequency and its five subsequent harmonics and produces pitch sensation corresponding to 400 Hz frequency. A complex tone in panel (b) does not have a 400 Hz component but maintains the pitch corresponding to 400 Hz frequency. This phenomenon is called *periodicity pitch, residual pitch, or the missing fundamental phenomenon* and is an indication that our auditory system responds to the overall periodicity of the incoming sound wave. The explanation of the periodicity pitch phenomenon lies in the fact that the missing 400 Hz frequency in panel (b) is the highest common denominator of all the frequency components shown in this panel. Thus, both complex waves represented by the spectra shown in panels (a) and (b) have the same basic period of their complex waveforms. The auditory system groups frequency components that share a common mathematical denominator and matches them to a prototypical sound with those components and assigns a pitch value based on that prototype according to the *Law of Simplicity*. This physiological mechanism is supported by an observation that the tonotopic organization of the auditory cortex is based on pitch rather than frequency and, thus, that signal periodicity is transmitted to the brain (Pantev et al., 1989).

The periodicity pitch phenomenon may seem like an artificial type of phenomenon in respect to head-mounted display (HMD) considerations. However, telephone communication is a common example of the occurrence of this phenomenon. Telephones typically transmit frequencies between 300 and 3600 Hz whereas the average fundamental frequencies of male and female voices are 125 and 200 Hz, respectively. This means that they lie below the transmission range of the telephone system. However, one is usually able to correctly identify the gender of the talker on the telephone because the male voice is still perceived as being lower in pitch.

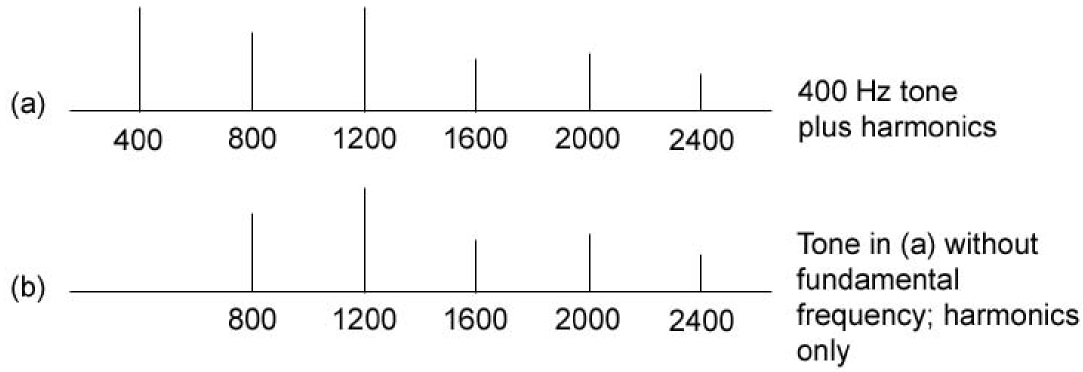


Figure 13-2. Fourier spectra of (a) 400 Hz tone with its lowest five harmonics and (b) the same complex tone without its 400 Hz fundamental. Both complexes produce different timbre sensations but result in the same pitch sensation.

Circular pitch

Pitch is often described as the "highness" or "lowness" of a tone. However, in Western music, pitch relationships are organized into 12-tone sequences defined by the ratio of the tone's fundamental frequency to the scale's tonic scale tone. Each successive octave has a fundamental frequency that is double that of the previous octave. Thus, although the frequency and the associated pitch rise linearly, the musical pitch scale is circular. This concept is shown in Figure 13-3. Such dualism led to two components of pitch: pitch height associated with frequency and pitch chroma associated with music intervals. Pitch height constitutes a basis for arranging the sounds into streams associated with specific sound sources while pitch chroma constitutes a basis for arranging the sounds into specific acoustic patterns (melodies) regardless of the sound source (Rakowski, 1993; Warren et al., 2003).

As noted in the discussion of the periodicity pitch illusion, pitch is not entirely determined by the fundamental frequency, but rather by the relationship of the harmonic frequency components and their fit with a prototypical sound with a certain pitch. Based on this concept, American psychologist Roger Shepard and composer James Tenney developed at Bell Labs a circular set of 12 complex tones, called *Shepard tones* or *Shepard staircase*, which, when played in a continuous loop, make the impression of a indefinitely rising or descending music scale (Shepard, 1964; Tenney, 1969). This phenomenon is frequently called the *circular pitch illusion* or *circular pitch paradox*. The name circular pitch refers to the fact that although the perceived scale progresses continuously in one direction, in fact it is played by the circular repetition of the same 12 tones. It is an auditory analog of the moving barber's pole or *Penrose stairs* illusions in vision (Mussap and Crassini, 1993; Seckel, 2004).

Fundamental frequencies of Shepard tones cover the span of 1 octave and differ from each other by a semitone (~6%). Each complex tone consists of several harmonic components that form a base-2 geometric relationship (1, 2, 4, 8, 16, etc.). Such tones are constructed to have clearly different pitch chroma but to be very similar in pitch height and timbre.

The pitch height ambiguity of Shepard tones is due to the spectral shape of the individual tone complexes. The spectral envelope of all of the complex tones can be described by a single Gaussian function as shown in Figure 13-4. When the Shepard tones are presented serially, the intensities of individual harmonics change slightly so that as the scale ascends, the higher components become less intense while the lower ones become more intense. Thus, at the end of the 12-tone sequence the shifting of intensity weights makes the 13^{th} tone identical to the first one. Since according to the *Law of Proximity*, the perceptual system prefers small intervals, the ear follows the frequency components of successive tones and perceives it as a continually rising pitch sequence. The effect is reversed when the tones are cycled in the opposite direction.

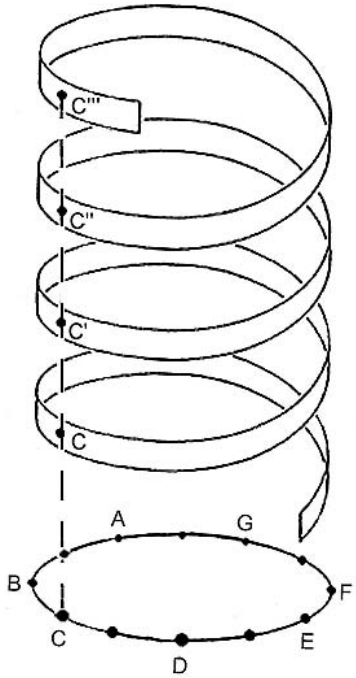


Figure 13-3. Schematic representation of pitch height (vertical axis) and pitch chroma (horizontal axis) (adapted from Shepard, 1982).

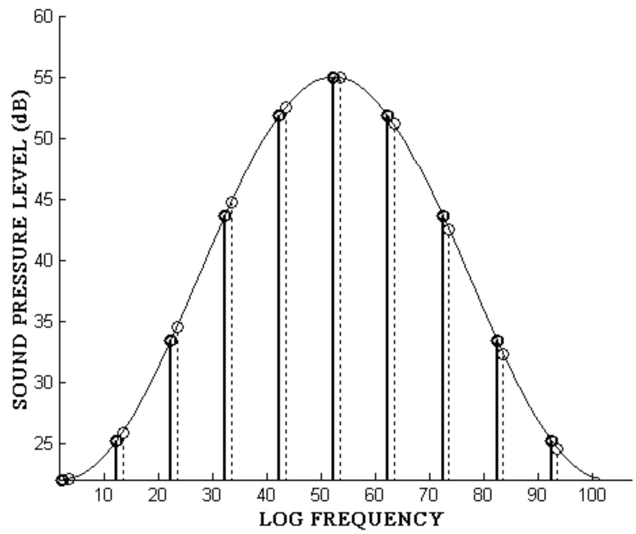


Figure 13-4. Schematic of the Gaussian curve that describes the sound pressure levels of the component frequencies that make up the Shepard tones. The solid lines and dark circles represent the components of 1 tone, the dotted lines and light circles represent the components of the subsequent tone in the scale (adapted from Shepard, 1982).

It is noteworthy that the circular pitch illusion described above had to be intuitively known by music composers prior to the development of Shepard tones. For example, tonal sequences creating this illusion can be found in pieces by Bach and Chopin (Wikipedia, 2008). The illusion is also present in some more modern compositions as well as in the video game Super Mario 64 in association with an infinite staircase.

The circular pitch illusion is best heard when the subsequent tones are presented with short silent intervals between the successive tones. However, it is not limited to discrete steps in frequency or even to octave-based complex tones (Burns, 1981). In 1969, a French composer Jean-Claude Risset working at Bell Labs developed a continuous version of the Circular Pitch illusion known as the Shepard-Risset Glissando or the Continuous Risset Scale (Risset, 1969). The effect is created by 10 harmonically related pure tones that cover the span of nine octaves. All pure tones simultaneously decrease their frequencies in a logarithmic fashion across the span of 10 octaves. The intensity of the overall sound is controlled by a Gaussian function similar to that of the Shepard tones illusion. In addition, the tones differ in the initial phase to maintain the continuity effect.

Jean-Claude Risset is also the author of a circular rhythmic illusion, called the *Risset pattern*, in which the perceived tempo seems to increase or decrease endlessly (Risset 1972; 1986). This illusion is based on the simultaneous presence of several drum beats having simple geometric relations. As the time progresses, the slower beats are made less intense while the faster beats increase in intensity so that the listener gradually changes focus from the slower to the faster beats as they become louder.

The sound effects, based on circular pitch and the Risset pattern, are useful for the simulation of ascent or descent in toys and games. One of the authors remembers having a toy spaceship that played a continuously ascending sound whenever pointed upwards and a continuously descending sound when pointed downwards. A similar sound could easily be used to provide feedback on the directional use of hand-operated controls. It may also have a practical application for warning signal design. In some situations a continuously ascending or descending pitch signal may be needed to force the operator's action. In such cases Shepard tones may be an efficient engineering solution.

The tritone paradox

Despite the circular nature of the Shepard tones, one perceives a directional change because the components are always one semitone apart. So, essentially one is choosing between movement of one semitone in one direction, or 11 semitones in the other direction (Figure 13-5). The simplest interpretation is the shorter distance of one semitone. However, if one hears the first semitone followed by the 6th semitone of the Shepard tones, the movement could be interpreted as going either six semitones up or down. In music, this half-octave interval is called a "tritone" hence the name of this auditory conflict.

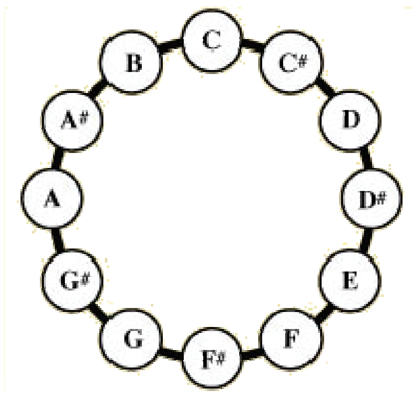


Figure 13-5. Pitch circle. Stepwise changes in the clockwise direction are perceived as ascending. Counterclockwise steps are perceived as descending. The tritone effect occurs when the interval is exactly half of an octave, or halfway around the circle (adapted from Deutsch, 1999).

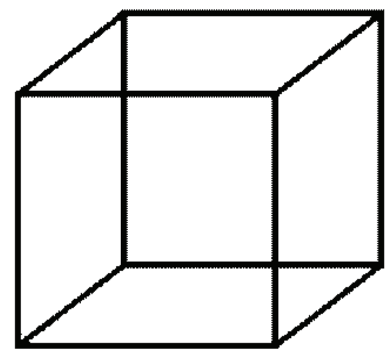


Figure 13-6. Necker cube. Ambiguous line drawing that has two interpretations when perceived as a three-dimensional cube.

When listeners are presented with Shepard tones that are exactly ½ octaves apart, some will perceive the pitch as rising and others as falling. There is usually a listener's bias to hear a particular tritone pair as rising or falling and this remains constant over time. This effect is an auditory analog of the visual conflict represented by the *Necker cube* (Figure 13-6) where a line drawing of a cube can be interpreted as facing either left or right, depending on whether the two intersecting lines on the bottom left are perceived as forming the front or the rear corner of the figure (Marr, 1982). In both cases one can often force oneself to reverse the initial perception.

The octave illusion

In the *octave illusion* (Deutsch, 1974) two pure tones with frequencies an octave apart are presented through the earphones as two tones of equal amplitude alternating between the ears. As a result the listener is presented with a two-tone continuous complex sound with both tonal components switched repeatedly between ears. This sequence of events is shown in Figure 13-7. Deutsch confirmed through several experimental studies that the presented pattern of events is never heard as such. The resulting perceptual effect varies greatly with the listener. Most people hear a complex tone that changes in pitch from high to low as it switches back and forth between the ears so that the high pitched tone is heard in one ear and the low pitched tone in the other ear. Other people hear one tone (low or high) pulsing in both ears accompanied by alternating pulses of the other tone in one ear. Some people reported changes in pitch or in the speed of ear alternations during signal presentation. Still others hear more complex percepts. Reversing of the earphones has no effect on the laterality of perception. What was heard in the left ear remained the same after earphone reversal.

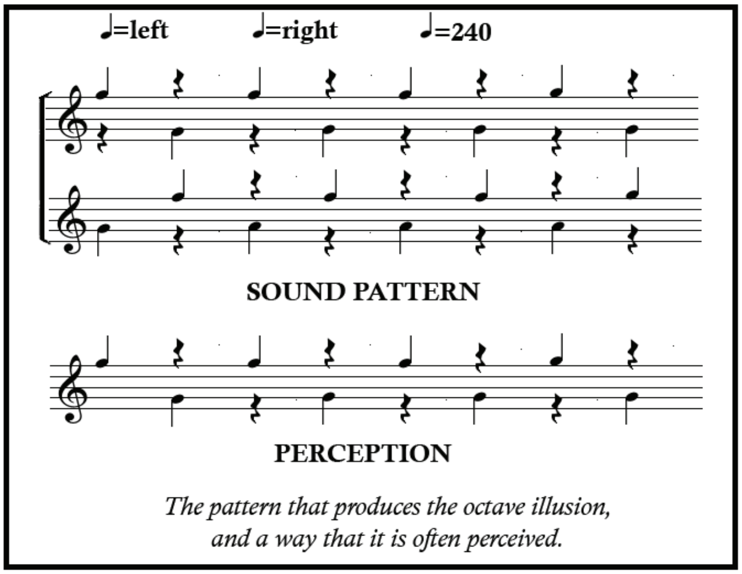


Figure 13-7. Octave illusion. The sequence of tones presented to the left and right ear (upper panel) and the typically perceived sequence of events (lower panel) (Deutsch, 1974).

How is this explained? There are two Gestalt Laws in effect here. The *Law of Similarity* suggests that the same pitched tones should be grouped together as coming from the same sound source. Conversely, the *Law of Proximity* suggests that the sounds occurring in each ear should be grouped together as coming from the same location. It is implausible that a single sound source would be moving from left to right and back again at a high rate of speed. Thus, the perceptual interpretations based on the above two laws conflict with each other and individual perceptions vary, even over time. Therefore, despite the fact that the described phenomenon is called the octave illusion it could also be called the octave conflict. Regardless of its name, there are two notable consistencies in the perceptions reported. First, right handed listeners tend to hear the higher pitch in their right ear and the lower one in their left ear (Deutsch, 1983). No such bias was found to exist for left handed listeners. Second, and most notably, nobody can hear the pitches as they actually occur.

Pitch streaming illusions

The group of *pitch streaming illusions* is based on the same stimulation paradigm as the octave illusion. They all exploit the *Law of Good Continuation,* which states that if several sound components occur that are close to each other in pitch and form a regular pitch contour, they will be perceived as coming from the same sound source as part of the same sound event. This is sometimes true even if the other segregating cues suggest other interpretations. In the *scale illusion* (Deutsch, 1975), a major scale is presented with successive tones alternating from ear to ear. Two versions of the scale are presented simultaneously, ascending and descending, so that when a tone is played from the ascending scale in one ear, the corresponding tone from the descending scale is played in the other ear.

The *Law of Good Continuation* suggests that the tones that form a regular pitch contour will be grouped together. This is what occurs. Listeners usually hear the high tones in one ear and the low ones in the other. As with the octave illusion, right-handers tend to hear the higher notes in the right ear, while no regular bias occurs for left-handers.

Deutsch (1987; 1995; 2003) demonstrates several other illusions such as the *chromatic illusion*, the *Glissando illusion*, and the *Cambiata illusion* that all function the same way as the scale illusion. In each case, tones forming a regular pitch contour or that are close to one another in pitch, are grouped together and appear to come from the same ear. Handedness often plays a role in the assignment of a pitch register to an ear. For right-handers, lower pitches are often assigned to the left ear and higher ones are assigned to the right ear.

A further extension of pitch streaming illusions can be found in the *phantom word illusion* created when words or syllables are used in the place of pure tones. Several of these illusions were developed and described by Deutsch (2003). They all involve two syllables or two words (for example "high" and "low" in a *high-low illusion*) that are played simultaneously one word to each ear switching the ears after each presentation. The listeners always have an illusion that a certain word or a short phrase is played repeatedly but they never hear and report the words as they are actually presented.

The split-off illusion

The *split-off illusion* (Figure 13-8) appears when an ascending tone glide and a descending tone glide are played so that the descending glide begins 200 milliseconds (ms) before the ascending glide ends. The beginning pitch of the descending glide starts out lower than the pitch of the ascending glide at that point in time and the pitch trajectories never cross. However, the percept is that of a continuously rising and falling glide. The final 200 ms of the ascending tone "splits off" and is heard as a separate tone in the middle of the frequency range of the glide.

Several practical realizations of this illusion have been described (Remijn, Nakajima and ten Hoopen; Sasaki and Nakajima, 1996). In all cases, two longer glides are fused together, and components of these glides that are inconsistent with the perception of a simple smooth trajectory are "split-off" and are heard as a separate tone. The auditory system seems to connect the two glides in the simplest way possible according to the Gestalt *Law of Good Continuation*. Any components that are inconsistent with this interpretation are either ignored or parsed away from it as being independent from the fused glides.

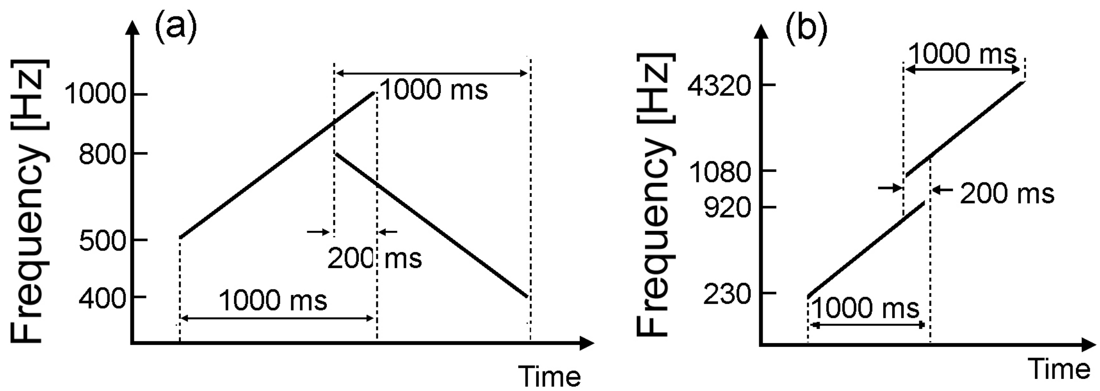


Figure 13-8. Two examples of the split-off illusion. In both cases, the final 200 ms of the first glide "splits off" and is perceived as an independent tone. The rest of the first glide and the second glide are joined perceptually into a single glide that either is (a) rising and falling or (b) rises continuously (adapted from Nakajima, Sasaki and ten Hoopen, 2006).

Huggins pitch (Dichotic pitch)

One Gestalt laws is the *Law of Proximity,* which states that sounds occurring near to each other in space are judged to be from the same sound source. The *Huggins pitch illusion* (Cramer and Huggins, 1958) is a faint pitch sensation that results from specific interpretation by the brain of the cue of interaural phase differences when two similar noise signals are delivered to the left and right ear. In the Huggins pitch demonstration the sounds delivered to each ear are white noise signals with the exception of three narrow bands; 400 to 440 Hz, 500 to 550 Hz, and 600 to 660 Hz. Both signals begin in phase, and then the phase of the signal delivered to one of the ears is advanced by 180° in each narrow band successively. The perceptual impression is that the noise is accompanied by a tone with gradually increasing faint pitch. No pitch is heard when the left or right ear signal is heard alone. Therefore, the Huggins pitch is a result of binaural processing of two slightly different signals.

The physiological explanation of Huggins pitch is still unclear. However, it seems that the brain separates the narrow bands of noise with 180° phase shift from the rest of noise and treats these sounds as coming from a different sound source separated in space from the noise source. Therefore, the Huggins pitch is essentially a spatial effect that that makes it easier to hear the presence of two simultaneous but different sounds. It is classified here as a pitch illusion but it may be equally well classified as a spatial illusion.

The Huggins pitch illusion illustrates an important consideration for HMDs and auditory displays in general. A binaural, spatial (3-dimensional) auditory display is the best way to improve detection of auditory information in operational conditions where noise cannot be reduced. This is true because to the degrees that the different signals are spatially separate, binaural presentation will create masking level differences, (discussed earlier in this chapter), and increase the effective signal to noise ratio.

Temporal conflicts and illusions

Temporal conflicts and illusions are generally related to perception of temporal patterns and the effects of the interstimulus interval on perceived sounds. The latter effects are related to the presence of temporal masking (short intervals) and memory traces (long intervals) and were already discussed in Chapter 11, *Auditory Perception and Cognitive Performance*. Therefore, the focus of the present chapter is on perception of temporal patterns and more precisely on the powerful illusion of pattern continuity whenever such continuity may be assumed.

The continuity effect

The *continuity effect* is an illusion observed when a soft sound is interrupted by a louder sound. Despite interruption, the original sound is heard as maintaining its continuity and the interrupting sound is heard as a separate event. The appearance of this illusion has certain limitations regarding the types of both sounds and their relative intensities but this is a powerful and easy to replicate perceptual effect. This continuity phenomenon was originally described by Miller and Licklider (1950) and called the *picket-fence effect*, but there is little doubt that it was heard and known before. The authors observed that a tone interrupted by a more intense burst of noise was heard as being continuous despite the interruption. The same effect was observed when the tone was replaced by speech. This idea of continuity illusion is shown schematically in Figure 13-9. When a tone interrupted by a temporal gap is presented in quiet (Figure 13-9a), the interruption is heard clearly. However, if a wideband noise burst is inserted into the gap, the tone is perceived to be continuous (Figure 13-9b). It is as though the auditory system assumes that the tone must be continuous and "fills in" the missing information. This effect is reduced if the wideband stimulus has a notch in the same frequency range as the tone (Figure 13-9c), which suggests that the auditory system may be extracting the tonal information from the wideband signal.

The continuity illusion also can be observed for tone glides, music, and continuous environmental sounds, such as rain sound, stream of typewriter sounds, etc. In other words, the continuity illusion works if the sounds before and after interruption are assumed to come from the same sound source (Bregman, 1990; Plack, 2005). Similarly to tones and other continuous sounds, continuous speech signal interrupted by short pauses looses its continuity and intelligibility; and the silent gaps are clearly heard. However, when the silent gaps are filled with bursts of wideband noise, coughs, or other wideband sounds, the listener has the impression that the speech signal is continuous and "hears" the missing parts of the speech sounds. As a result, speech recognition improves in comparison to that of the speech interrupted by the silent gaps. The mental restoration of the original speech masked by short interfering louder wide band sounds has been referred to as *phonemic restoration* (Warren, 1970; Warren and Obusek, 1971).

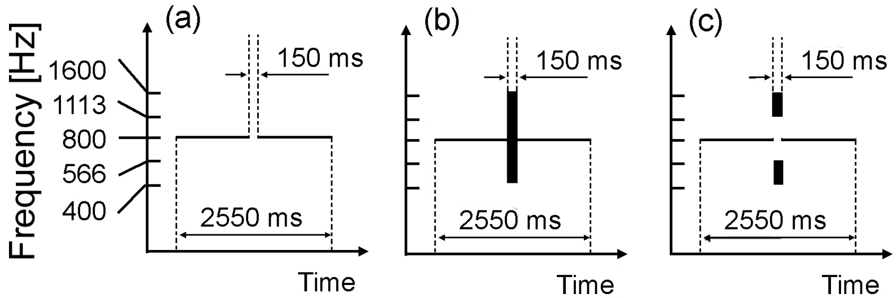


Figure 13-9. Schematic drawing of three time paradigms used to demonstrate the continuity effect; (a) continuous tone with silent gap, (b) continuous with a silent gap filled with broadband noise, and (c) continuous tone with a silent gap filled with two bands of noise surrounding the tone frequency (adapted from Nakajima, Sasaki and ten Hoopen, 2006).

The name *continuity effect* for the continuity illusion was used originally by Thurlow and Elfner to describe the perceived continuity of sound when two sounds alternate in time. The authors presented alternating pulses of a soft pure tone and a loud other sound (noise, another pure tone, etc.) that were alternating in time and observed that the soft tone was heard as a continuous tone (Elfner and Caskey, 1965; Thurlow, 1957; Thurlow and Elfner, 1959). Thurlow (1957), who apparently rediscovered this effect not knowing about the study by Miller and Licklider (1950), has originally called this effect the *auditory figure-ground effect* by an analogy to a similar figure-ground effect in vision (Rubin, 2001) where hidden background seems to be continuous behind foreground objects. An example of visual ground-figure effect is shown in Figure 13-10.

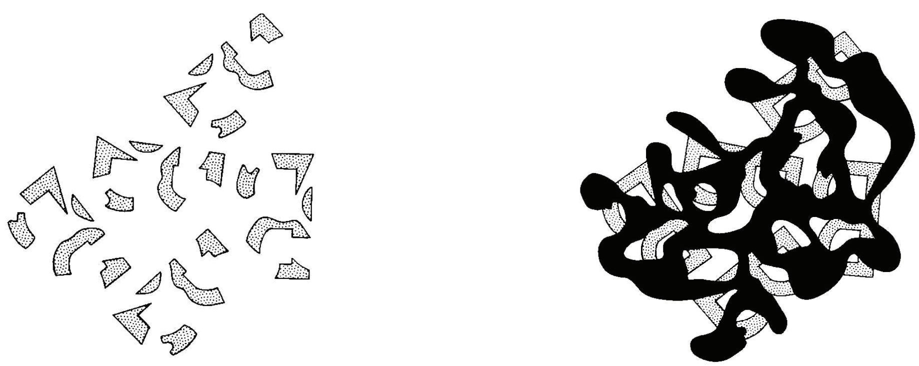


Figure 13-10. Visual example of perceptual restoration. The occluding information provides information about which elements are likely to be continuous and which elements are likely to be discrete (Bregman, 1981).

Since an alternate pulsing is an expanded form of single or multiple interferences, the term continuity effect can be used as a single label for all of the phenomena described here. However, various other names are also used in the literature to describe continuity effects including the *acoustic tunnel effect* (Vicario, 1960, cited by Warren, 1999), *auditory induction* (Warren, Obusek and Ackroff, 1972), and *temporal induction* (Warren, 1999). In addition, Warren and colleagues (Warren, Obusek and Ackroff, 1972) reported that the continuity illusion can be heard when the interfering sound is just a higher level of the original sound. The authors presented the same stimulus (octave band noise centered at 2000 Hz) at two alternating levels (70 and 80 dB sound pressure level [SPL]) and observed that the lower level was heard as always being present while the higher level was heard as an additional sound. This observation led Warren (1999) to differentiate between *homophonic continuity* (the same signal, different levels) and *heterophonic continuity* (different signals, different levels) effects.

A parallel effect to the continuity effect (temporal restoration) is the spectral restoration effect where presence of noise in spectral gaps of filtered speech improves speech intelligibility. For example, Bashford and Warren, (1987) and Shriberg (1992) reported improved speech intelligibility where low-pass or high-pass filtered speech was heard together with a complementary filtered noise. Both of these effects together have been referred to as *perceptual restoration* by Warren (1999)

Pulsation threshold

The *pulsation threshold* is not an auditory illusion but a research method used to assess spectro-temporal analysis performed by the auditory system (Fastl, 1975; Houtgast, 1972; Letowski and Smurzynski, 1983). Since this method is based on some properties of the continuity effect and has important implications for understanding auditory physiology it, deserves a short description.

The pulsation threshold methodology was developed by Houtgast (1972) as an alternative to temporal masking in studying lateral suppression. The procedure uses relatively long maskers and signals so it is much easier to use than temporal masking techniques. This methodology is based on an observation that in order for the continuity effect to occur, the intensity of the louder sound must be such that the softer sound would be totally masked if the louder sound was continuous (Warren, Obusek and Ackroff, 1972; Houtgast, 1972). This means that the barely audible pulsing of the softer sound can be used as a measure of its masked threshold, that is, the amount of the excitation overlap of the two sounds in the cochlea (Platt, 2005). This methodology is especially convenient for measuring masked threshold in the vicinity of the masked tone or its harmonics in tone-on-tone experiments where the potential beats[1] make a simultaneous masking technique quite unusable. It is also important that despite a subjective criterion that is used by the listener in determining whether the soft signal is continuous or pulsing, the pulsation threshold data have relatively low variability (Plack and Oxenham, 2000).

A schematic diagram of the temporal pattern used to measure the pulsation threshold is shown in Figure 13-11. A masking (M) and a masked (m) tone alternate in time and the experimenter adjusts either the level of the masking tone or the level of the masked tone until the continuity effect is heard. The masking tone is usually a wideband noise or a pure tone signal and the masked tone is a pure tone signal. The former adjustment procedures is used to determine the level of the masking tone needed to mask the other tone and the latter procedure is used to measure the masked threshold of the signal in the presence of a given masking tone. The optimum alternation cycle for measuring pulsation effect is about 4 Hz but interruption times can be as long as 300 ms beyond which tonal continuity cannot be maintained for longer periods (Houtgast, 1974; Warren, 1999).

The gap transfer illusion

When a long ascending glide tone is crossed in the middle by a short descending glide tone (Figure 13-12a), the pitch percept of the longer tone is often sigmoidal (Halpern, 1977; McPherson, Ciocca and Bregman, 1994; Tougas and Bregman, 1985). If a short, 100 ms gap is added to the short glide, the pitch is perceived veridically (Figure 13-12b). Evidently, when there is no gap, the *Law of Similarity* encourages us to group the pitch components in the shorter glide with the dominant longer glide. The gap separates those frequencies and allows the longer ascending pitch to be perceived veridically. Nakajima and his colleagues (Nakajima et al., 2000) show

---

[1] *Beats* are the effect produced by interference of waves of slightly different frequency, producing a pattern of alternating intensity.

that the same percept is achieved when the gap is placed in the longer glide (Figure 13-12c). Essentially, the gap is “transferred” to the shorter glide and the long glide is perceived as continuous. Here, the *Law of Good Continuation* dictates that the well established long glide is more likely to be continuous and assigns the nearby frequencies to it.

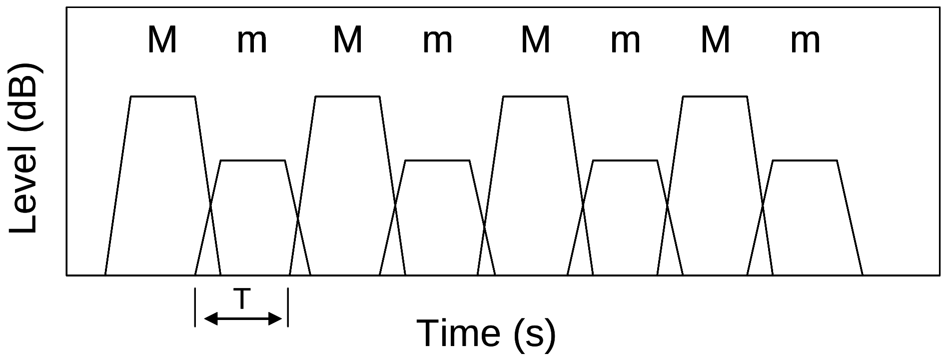


Figure 13-11. An example of the temporal paradigm used to elicit a pulsation threshold. Two signals, masker (M) and maskee (m) alternate in time with onset and offset ramps of about 20 ms to avoid generation of audible clicks during transitions. The duration (T) of each pulse is the same and usually about 100-150 ms.

It is somewhat striking that the auditory system is so susceptible to bias. It the previous example, the gap is in an ascending glide, and the short glide is descending – yet the long glide is perceived as continuously ascending. It is as though the auditory system is unable to adequately process the pitch information, so it “fills in” the missing information with a plausible interpretation. The *Law of Simplicity* essentially posits that the perceptual system will interpret sensory information with the simplest interpretation possible. Take, for example, Figure 13-13. Are the drawings 2-dimensional or 3-dimensional? Either interpretation is valid; however, the probability of a three-dimensional interpretation changes as you progress through the figures. The simplest interpretation of Figure 13-13a is that of a cube. Figure 13-13b is symmetric, but still has irregular shapes, and is still probably interpreted as a cube. However, Figure 13-13c consists of 6 congruent triangles in a symmetric arrangement and either a 2- or 3-dimensional interpretation is equally possible.

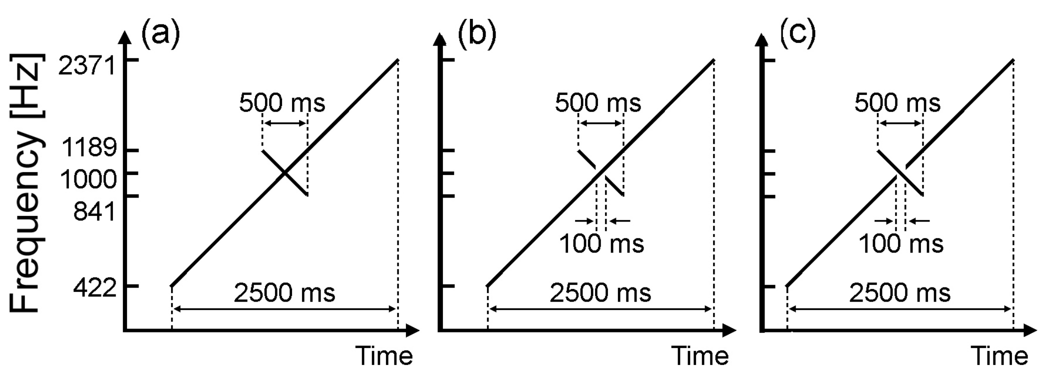


Figure 13-12. The gap transfer illusion. The percept of the event depicted in (a) is usually sigmoidal; the frequency components in the short glide are incorporated into the dominant longer glide. (b) Adding a gap to the short glide allows the event to be heard veridically. Event (c) is perceived to be the same as event (b); the gap is transferred perceptually to the shorter glide and the longer glide is perceived as continuous (adapted from Nakajima, Sasaki and ten Hoopen, 2006).

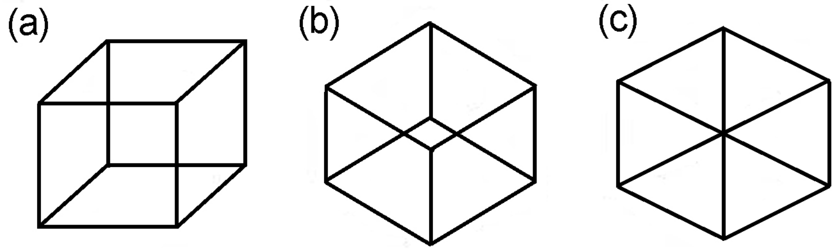


Figure 13-13. Figures demonstrating the *Law of Simplicity*. Observer perceptions are guided by the simplest interpretations of these lines and there is a bias to perceive them as simple, symmetric shapes.

Spatial conflicts and illusions

The Franssen effect

The *Franssen effect* (Franssen 1960; Yost, Mapes-Riordan and Guzman, 1997) is a spatial illusion caused by the temporal difference in the onset of sound in the left and right ear of the listener. It appears if a narrowband stimulus is presented in a reverberant sound field from two loudspeakers located at about ±45º and placed at a certain distance (1 meter or more) from the listener. In the demonstration of this effect the one loudspeaker that delivers the sound is abruptly turned on and then slowly turned off with short and linear offset ramp (20 to 100 ms). The other loudspeaker is gradually turned on and the sound is delivered with an onset envelope that is a mirror image of the offset envelope of the first sound and kept on for some time (e.g., 2 to 10 s) after reaching its steady state level. If the sound is a wideband noise, the noise is suddenly heard in one ear and then "jumps" to the other ear for the rest of the sound duration, as expected. However, when the sound is a narrowband noise or a tone (e.g., 1000 Hz tone) the sound is heard as coming from the first loudspeaker the entire time, that is, from the loudspeaker that delivered a short abrupt sound during the first short period of the sound presentation.

The Franssen effect is a special case of the *precedence effect* (see Chapter 11, *Auditory Perception and Cognitive Performance*). The precedence effect results from the fact that the sound arriving first at the ears determines the perceived location of a sound source in the space. The Franssen effect occurs because people have difficulty localizing sounds with a gradual onset envelope and room reflections obscure the residual localization cues. The effect is the easiest to demonstrate for stimuli in 1.0 to 2.5 kHz frequency range, that is, in the range when neither temporal nor intensity localization cues work very well and it is difficult for people to localize sound sources. The presence of reflections is very important for eliciting the Franssen effect and the effect is difficult to demonstrate in anechoic conditions or through earphones (Hartmann and Rakerd, 1989; Rakerd and Hartmann, 1985). However, under proper listening conditions the illusion may exist for very long sounds; even infinitely long (Bradley 1983). The Franssen effect is an auditory illusion that requires unusual circumstances to occur, but it demonstrates our perceptual reliance on sound onset cues for sound source localization.

The Clifton effect

The *Clifton effect* (Clifton, 1987; Clifton and Freyman, 1989) is a spatial illusion that represents the breakdown of the precedence effect (see Chapter 11, *Auditory Perception and Cognitive Performance*). To demonstrate the Clifton effect a train of click pairs separated by several milliseconds (e.g., 12 ms) is emitted from two loudspeakers located at about ±45º and at a certain distance from the listener. Each pair of clicks consists of a click from one loudspeaker (the lead loudspeaker) and a click from the other loudspeaker (the lag loudspeaker). After initial presentation of several pairs of the clicks, the loudspeakers delivering the lead and the lagging clicks are reversed, that is, the loudspeaker delivering the lead click now delivers the lagging click and vice versa. As predicted by the precedence effect, most listeners perceive a single click arriving from the location of the lead loudspeaker during the presentation of the initial segment of the click train. Immediately after the switch, the clicks are perceived as coming from both loudspeakers. This perceptual effect last for the duration of 3 to 5 click-pairs. After this period of time a single click, coming again from the leading (now the opposite) loudspeaker is heard, as predicted by the precedence effect. The Clifton effect demonstrates human inability to immediately adjust to the change in the position of the leading click and the temporary failure of the precedence effect (Yost and Guzman, 1996). This illusion suggests that the suppression mechanism of the lagging stimulus may be active for some time after termination of the lead sound. The duration of this activity may be dependent on the duration of the exposure to the lead-lag pairs of stimuli coming from the same locations (Litovsky et al., 1999).

Attention and Illusions

The perceptual system is able to acquire a large amount of sensory information, but it is only able to process and interpret a small proportion of it. The maximum rate of information that can be processed by a single sensory channel is called *channel capacity*. The concept of channel capacity is identical with the concept of working memory capacity, which refers to the maximum amount of sensory information that can be held temporarily in a storage buffer (Baddeley, 1992). This storage buffer is needed by the sensory system in order to continuously process incoming information. The processed information is then either used for current decision-making processes or stored in long-term (permanent) memory. The other terms that convey the same meaning as working memory are short-term memory and operational memory.

Miller (1956) posited that one could hold seven plus or minus two (7 ± 2) items of information in short term storage, roughly the equivalent of a phone number. Broadbent (1975) argued that the capacity of short-term memory is even smaller and the memory can only hold approximately three items. Either way, the capacity of working memory is very small and the concept of information "item" is somewhat nebulous. It seems that during information processing small items (chunks) of information that are well known or meaningful are grouped together into bigger and bigger chunks and each chunk can constitute an item. In the case of a phone number, the information item could be just a digit or it could be the area code or it could be the entire number.

Regardless of the precise concept of information item the absolute capacity of short term memory is very limited. Therefore, the perceptual system must be selective in which information is processed or attended to. There are numerous theories proposing how attention is allotted to a scene and which events generally elicit deeper processing (for a complete discussion, see Jones and Yee, 1993). However, since the overall attentional resources are capped at a certain level, they only can be increased in one channel by reducing those of another.

The existence of perceptual illusions demonstrates the fact that in order to facilitate efficient processing of information, the perceptual system relies on a number of heuristics to process sensory information and often "fills in" missing or contradictory information with plausible interpretations when information is incomplete (Shinn-Cunningham, 2008). Therefore, from the neuroscience point of view, perceptual illusions create an important key to understanding how the brain processes incoming auditory information and which parts of the information are not being processed. Also, from the display design point of view, the existence of illusions highlights the need to consider attentional limits and minimize complexity of incoming information, as well as considering their ecological validity to the perceptual system itself. Failure to do so increases the probability of errors due to incorrectly interpreted events. Knowledge of how the auditory system parses information can help to determine ways to increase the effective signal to noise ratio, both perceptually and cognitively. In addition, knowledge of common perceptual biases can also highlight potential sources of information loss and situations where the use of redundancy across modalities is warranted.

References

Adams, J., and Duda, P. (1986). Laterality of cross-modal spatial processing. *Cortex*, 22, 539-552.

Baddeley, A.D. (1992). Working memory. *Science*, 255, 556-559

Bradley, D.A. (1983). Room acoustics and listening. *Journal of the Acoustical Society of America,* 73, 17.

Bregman, A. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Bregman, A.S. (1981). Asking the "what for" question in auditory perception. In: *Perceptual Organization,* Kubovy, M., and Pomerantz, J.R., (Eds.). Hillsdale, NJ: Erlbaum.

Broadbent, D.E. (1975). The magic number seven after fifteen years. In: *Studies in Long-Term Memory,* Kennedy, A., and Wilkes, A. (Eds.), (pp. 3-18). New York: John Wiley and Sons.

Burns, E.M. (1981). Circularity in relative pitch judgments for inharmonic complex tones; The Shepard demonstration revisited, again. *Perception and Psychophysics*, 30, 467-472.

Clifton, R.K. (1987). Breakdown of echo suppression in the precedence effect. *Journal of the Acoustical Society of America,* 82, 1834-1835.

Clifton, R.K., and Freyman, R. (1989). Effect of click rate and delay on breakdown of the precedence effect. *Perception and Psychophysics,* 46, 139-145.

Cramer, E.M., and Huggins, W.H. (1958). Creation of pitch through binaural interaction. *Journal of the Acoustical Society of America,* 30, 412-417.

Deutsch, D. (1974). An auditory illusion. *Nature,* 251, 307-309.

Deutsch, D. (1975). Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, 57, 1156-1160.

Deutsch, D. (1983). The octave illusion in relation to handedness and familiar handedness background. *Neuropsychologia,* 21, 289-293.

Deutsch, D. (1987). Illusions for stereo headphones. *Audio Magazine,* 71, 36-48.

Deutsch, D. (1995). *Musical Illusions and Paradoxes*. La Jolla, CA: Philomel Records.

Deutsch, D. (1999). The processing of pitch combinations. In: *The Psychology of Music,* Deutsch, D. (Ed.), (pp. 349-411). San Diego: Academic Press.

Deutsch, D. (2003). *Phantom Words and Other Curiosities*. La Jolla, CA: Philomel Records.

Elfner, L.F., and Caskey, W.E. (1965). Continuity effects with alternating sounded noise and tone signals as a function of manner of presentation. *Journal of the Acoustical Society of America,* 38, 543-547.

Emanuel, D., and Letowski T. (2009). *Hearing Science*. Baltimore: Lippincott Williams and Wilkins.

Fastl, H. (1975). Pulsation patterns of sinusoids vs. critical band noise. *Perception and Psychophysics*, 18, 95-97.

Franssen, N.V. (1960). Some considerations on the mechanism of directional hearing. Ph.D. Dissertation. Delft University of Technology (Technische Hogeschool). The Netherlands.

Halpern, D.L. (1977). The effect of harmonic ratio relationships on auditory stream segregation. Unpublished research report, Psychology Department, McGill University.

Hartmann, W.M., and Rakerd, B. (1989). Localization of sound in rooms IV: The Franssen effect. *Journal of the Acoustical Society of America,* 86, 1366-1373.

Heller, M., Calcaterra, J., Green, S., and Brown L. (1999). Intersensory conflict between vision and touch: The response modality dominates when precise, attention-riveting judgments are required. *Perception and Psychophysics,* 61, 1384-1398.

Hupe, J.M., Joffo, L.M., and Pressnitzer, D. (2008). Bistability of audiovisual stimuli: perceptual decision is modality specific. *Journal of Vision,* 8, 1-15.

Hershberger, W., and Misceo, G. (1996). Touch dominates haptic estimates of discordant visual-haptic size. *Perception and Psychophysics,* 58, 1124-1132.

Houtgast, T. (1972). Psychophysical evidence of lateral inhibition in hearing. *Journal of the Acoustical Society of America,* 51, 1885-1894.

Jones, M.R., and Yee, W. (1993). Attending to auditory events: the role of temporal organization. In: *Thinking in Sound: The Cognitive Psychology of Human Audition,* McAdams, S., and Bigand, E. (Eds.), (pp. 69-112). Oxford, UK: Oxford Science Publications.

Leopold, D., and Logithesis, N. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Science,* 3, 254-264.

Letowski T., and Smurzynski, J. (1983). Evaluation of electroacoustic devices by equivalent scale methods. *Archives of Acoustics*, 8, 133-154, 1983

Litovsky, R.Y., Colburn, H.S., Yost, W.A., and Guzman, S.J. (1999). The precedence effect. *Journal of the Acoustical Society of America,* 106, 1633-1652.

Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices, *Nature,* 264**,** 746-748.

McPherson, L., Ciocca, V., and Bregman, A.S. (1994). Organization in audition by similarity in rate of change: Evidence from tracking individual frequency glides in mixtures. *Perception and Psychophysics,* 55, 269-278.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 63, 81-97.

Miller, G.A., and Licklider, J.C.R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America,* 22, 167-173.

Mussap, A.J., and Crassini, B. (1993). Barber-pole illusions and plaids **-** The influence of aperture shape on motion perception. *Perception,* 22, 1155-1174.

Nakajima, Y., Sasaki, T., Kanafuka, K., Miyamoto, A., Remijn, G., and ten Hoopen, G. (2000). Illusory recouplings of onsets and terminations of glide tone components. *Perception and Psychophysics*, 62, 1413-1425.

Nakajima, Y., Sasaki, T., and ten Hoopen, G. (2006). *Demonstrations of auditory illusions*. Retrieved 26 August, 2008 from: http://www.kyushu-id.ac.jp/~ynhome/ENG/Demo/illusions.html#title

Pantev, C., Hoke, M., Lütkenhöner, B., and Lehnertz, K. (1989). Tonotopic organization of the auditory cortex: Pitch versus frequency representation. *Science,* 246, 468-488.

Plack, C.J., (2005). *The Sense of Hearing,* (p. 212). Mahwah, NJ: Lawrence Erlbaum Associates.

Plack, C.J., and Oxenham, A.J. (2000). Basilar-membrane nonlinearity estimated by pulsation threshold. *Journal of the Acoustical Society of America,* 107, 500-507.

Pressnitzer, D., and Hupe, J.M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology,* 16, 1351-1357.

Rakerd, B., and Hartmann, W.M. (1985). Localization of sounds in rooms II: The effect of a single reflecting surface. *Journal of the Acoustical Society of America,* 78, 524-533.

Rakowski, A. (1993). Categorical perception in absolute pitch. *Archives of Acoustics*, 18, 515-523.

Remijn, G., Nakajima, Y., and ten Hoopen, G. The split-off effect and the perception of auditory continuity (as cited in Nakajima, Sasaki and ten Hoopen, 2006).

Risset, J.C. (1969). Pitch control and pitch paradoxes demonstrated with computer-synthesized sounds. *Journal of the Acoustical Society of America,* 46, 88.

Risset, J.C. (1972). *Musical Acoustics*. Paris: IRCAM.

Risset, J.C. (1986). Pitch and rhythm paradoxes: Comments on "Auditory paradox based on a fractal waveform". *Journal of the Acoustical Society of America,* 80, 961-962.

Rubin, E. (2001). Figure and ground. In *Visual Perception*, Yantis, E. (Ed.), (pp. 225-229). Philadelphia: Psychology Press.

Sasaki, T., and Nakajima, Y. (1996). An illusory reconstruction of auditory elements. *Journal of the Acoustical Society of America,* 100, 2751.

Seckel, A. (2004). *Masters of Deception: Escher, Dalí and the Artists of Optical Illusion*. New York: Sterling Publishers.

Shepard, R.N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 36, 2346-2353.
Shepard, R.N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, 89, 305-333.
Shinn-Cunningham, B.G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, 12, 182-186.
Shriberg, E.E. (1992). Perceptual restoration of filtered vowels with added noise. *Language and Speech,* 35, 127-136.
Solso, R.L. (2001). *Cognitive psychology* (6th Ed.). Boston: Allyn and Bacon
Tenney, J. (1969). Computer music experiences: 1961-1964. *Electronic Music Report #1*. Utrecht, Netherlands: Institute of Sonology.
Thurlow, W.R. (1957). An auditory figure-ground effect. *American Journal of Psychology,* 70, 653-654.
Thurlow, W.R., and Elfner, L.F. (1959). Continuity effects with alternately sounding tones. *Journal of the Acoustical Society of America,* 31, 1337-1339.
Tougas, Y., and Bregman, A.S. (1985). The crossing of auditory streams. *Journal of Experimental Psychology: Human Perception and Performance,* 11, 788-798.
Van Noorden, L. (1975). Temporal coherence in the perception of tonal sequences. Ph.D. dissertation. Eindhoven University of Technology, Leiden (The Netherlands).
Vicario, G. (1960). L'effetto tunnel acustico. *Rivista di Psicologia,* 54, 41-52.
Warren, J.D., Uppenkamp, S., Patterson, R.D., and Griffiths, T.D. (2003). Separating pitch chroma and pitch height in the human brain. *Proceedings of the National Academy of Sciences USA,* 100, 10038-10042.
Warren, R.W. (1970). Perceptual restoration of missing speech sounds. *Science,* 167, 392-293.
Warren, R.M. (1999). *Auditory perception: A new analysis and synthesis*. Cambridge, UK: Cambridge University Press.
Warren, R.M., and Obusek, C.J. (1971). Speech perception and phonemic restorations. *Perception and Psychophysics,* 9, 358-362.
Warren, R.M., Obusek, C.J., and Ackroff, J.M. (1972). Auditory induction: Perceptual synthesis of absent sounds. *Science,* 176, 1149-1151.
Warren, R.M., Hainsworth, K.R., Brubaker, B.S., Bashford, J.A. Jr., and Healy, E.W. (1997). Spectral rtestoration of speech: Intelligibility is increased by inserting noise in spectral gaps. *Perception and Psychophysics,* 57, 275-283.
Wikipedia. (2008). Shepard tone. Retrieved 1 August 2008 from: http://en.wikipedia.org/wiki/ Shepard_tone
Yost, W.A., and Guzman, S.J. (1996). Auditory processing of sound sources: Is there an echo in here? *Current Directions in Psychological Science,* 5, 125-131.
Yost, W.A., Mapes-Riordan, D., and Guzman, S.J. (1997). The relationship between localization and the Franssen effect. *Journal of the Acoustical Society of America,* 101, 2994-2997.

14 AUDITORY-VISUAL INTERACTIONS

Thomas G. Ghirardelli
Angélique A. Scharine

Multisensory Perception

Most displays or other devices designed to communicate information focus on a single sensory modality. This is perhaps a consequence of the fact that most research examining human perceptual capabilities has focused on a single sensory system at a time. However, most events in the natural environment generate physical information affecting multiple sensory modalities. This information is typically co-located in both space and time, and our perceptual systems have evolved to create a single coherent representation of our environment. Recent research has increasingly acknowledged the importance of these regularities, and the topic of multisensory integration became critical for understanding global awareness of the environment. This chapter presents an overview of the most recent research into the integration of auditory and visual information and presents considerations for the design of multisensory (i.e., auditory-visual) displays. The initial section of the chapter presents comparison of the basic features of the auditory and visual modalities and examines the effects of interacting visual and auditory stimuli. Included in this section is a discussion of the relative strengths and weaknesses of each modality and the situations in which one modality dominates. This section is followed by a discussion of perceptual effects when information from both modalities is complementary or conflicting and the role of multisensory attention on the perception of auditory-visual information. The final section of the chapter presents a discussion of the applications of auditory-visual integration principles to audio-visual display (e.g., helmet-mounted displays [HMDs]) and signal designs. Research on the interaction of a third modality, tactile or haptic displays, is just beginning and will not be discussed extensively here, although many of the same considerations can be expected to apply (see Chapter 18, *Exploring the Tactile Modality for HMDs*).

Dominant Characteristics of Audition and Vision

Many of the interactive effects that will be discussed are driven by the unique characteristics of each modality. These affect which modality is preferable for a specific type of information and when a sensory mode will dominate. First, auditory information is characterized by temporal changes in the sound pressure wave arriving at the ear, and a complete sphere of receptivity around the head, albeit with differential sensitivity. In many cases, sounds are transient, meaning that they have terminated prior to the observer's response. The observer must either remember the features of the sound or supplement the sound information with visual information in order to make the response. On the other hand, visual information is characterized largely by changes in the intensity and/or spatial frequency of light waves across a limited spatial region the field-of-view, or field-of-regard.[1] Although some objects can move or change with time, the majority of the scene elements will remain constant over time.

Perhaps, because auditory information is primarily temporal, the temporal resolution of the auditory system is more precise. We can discriminate between single and pairs of clicks when the gap is only a few tens of microseconds (Krumbholz et al., 2003; Leshowitz, 1971). Perception of temporal changes in visual modality is much poorer, and the fastest visible flicker rate in normal conditions is about 40-50 Hertz (Hz) (Bruce, Green and Georgeson, 1996).

[1] Field-of-regard includes head movements, but not torso movements; field-of-view refers to the eye only as limited by whatever stops are in the field, e.g. glasses, NVG, etc.

In contrast, the maximum spatial resolution (contrast sensitivity) of the human eye is approximately 1/30°, a much finer resolution than that of the ear, which is approximately 1°. Furthermore, the relative temporal stability of visual information means that the observer has time to visually locate a visual object in his or her environment before resolving its details. An auditory object must be localized while it is still sounding, or remembered after the auditory event. Consequently, the visual modality tends to dominate spatial perception. This has consequences when visual and auditory information conflict; and will be discussed in the section on the capture effect.

Conversely, as noted previously, humans are sensitive to sounds arriving from anywhere within the environment; whereas, the visual field is limited to the frontal hemisphere, and "good" resolution is limited to the foveal region. Therefore, while the spatial resolution of the auditory modality is cruder, it can serve as a cue to events occurring outside the visual field-of-view.

Information presented by a display system must be remembered for at least as long as it takes the user to respond to it. This "short-term memory" (Klatzky, 1975) or "working memory" (Baddeley, 1982) refers to the limited storage capacity where we first process the stimuli originating from the environment. Its capacity is very limited and varies with modality. One common technique used to test the capacity of short-term memory is to present a list of words and then test for recall. This usually results in a pattern of results called the *serial position effect*, where the items at the beginning and end of the list are more likely to be recalled than those in between. When the mode of presentation is varied so that one can compare the effect of visual or auditory presentation, there is no difference in recall of items at the beginning of the list, but there is a slight improvement in memory for auditory items at the end of the list. However, since sound is transient and vision can be static, an auditory message is best accompanied by a visual message that can remain on the display until dismissed.

The modality effect appears to be eliminated for long term memory. Visual and auditory events are equally likely to be recalled. There does seem to be an effect of level of processing; so redundancy is advantageous. As the number of modes that information is presented in increases, the amount of processing of that information and the probability that it will be attended to also increases.

The perceived intensity of sound is referred to as *loudness* and the perceived intensity of light is referred to as *brightness* (Stevens and Marks, 1965), and each of these depends on the characteristics of the specific stimulus (sound or light) and the context in which the stimulus appears. Since intensity can be used to convey the importance or urgency of a signal, it is important to consider how the two modalities compare perceptually. When Stevens compared perceived intensity of a 75-4800 Hz band of noise and of a white light, he found close functional similarity between both sensory functions. The levels for which the two stimuli were perceived as equal depended somewhat on the test constraints (experimenter-paced or self-paced) and are shown in Figures 14-1 and 14-2.

Interaction of Audition and Vision

An interesting question is: What will happen to the perception of a visual stimulus when presented simultaneously with an auditory stimulus? Does the neural stimulation combine, improving detection? Or, does sensory input from one modality inhibit that of the other modality? The answers depend on several factors. For example, both Kravkov (1934) and Hartmann (1933) found facilitative effects of auditory tonal stimulation on visual thresholds and visual acuity. Others have found similar effects for broadband signals (Watkins, 1964; Watkins and Feehrer, 1964). Maruyama (1957, 1959) found that this effect is dependent on the frequency and intensity of the auditory stimulus. Kravkov and subsequent researchers have found that sensitivity to green light increases as sound intensity increased, but that this effect is reversed for orange-red light (Allen and Schwartz, 1940; Kravkov, 1936, 1939; Letourneau, 1972; Letourneau and Zeidel, 1971). Other studies have found inhibitory effects (Davis, 1966; Maloney and Welch, 1972) and that the effect is dependent on the temporal relationship between the stimuli in each modality (Broussard, Walker, and Roberts, 1952; Coleman and Krauskopf, 1956; Ince, 1968) or with no effect whatsoever (Symons, 1963).

Even if a stimulus in one modality is known to be irrelevant to the observer's response, any signal in the irrelevant modality may serve to enhance processing in the relevant modality. For example, Stein et al. (1996) found that observers' judgments of the intensity of a light-emitting diode (LED) were increased by the co-occurrence of an irrelevant noise burst, regardless of whether the noise originated from the same location as the LED or not.

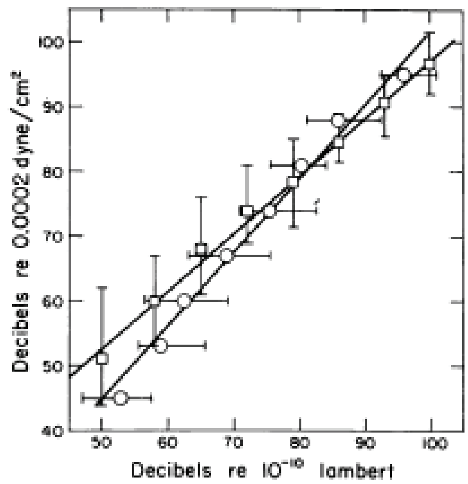


Figure 14-1. Equal-sensation functions for loudness and brightness, showing the levels of luminance and sound pressure that appeared equal in subjective intensity (Expt. 1). Squares: sound adjusted to match light; (Expt. 2). circles: light adjusted to match sound. The vertical and horizontal line segments show the interquartile ranges of the adjustments. Duplicate sentence deleted These ranges become much smaller when the intercept variability is removed.[2]

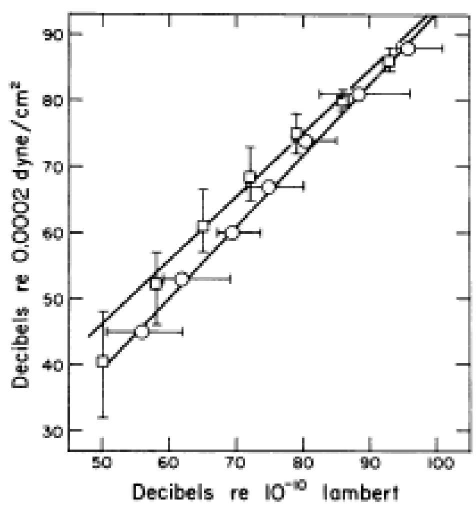


Figure 14-2. Equal-sensation functions for loudness and brightness, showing the levels of luminance and sound pressure that appeared equal in subjective intensity. Squares: sound adjusted to match light; circles: light adjusted to match sound.

[2] Note: Figures taken from "Cross-modality matching of brightness and loudness" by J. C. Stevens, and L. E. Marks 1965, *Proceedings of the National Academy of Sciences of the United States of America*, 54, 407-411. (Reprinted with permission.)

Because the findings are so inconsistent, it is tempting to dismiss them altogether. Often the reported findings were obtained under greatly restricted experimental conditions. For example, dark adapted observers seated in a dark room with no extraneous distractions showed small improvements in their ability to detect simple Gabor patches,[3] colors, pure tones and narrowband noise bursts. When one considers the contrast provided by the normally rich environment, these minute differences in threshold may not be significant. However, there do seem to be two consistent effects observed in auditory-visual studies. The first is that the combined neural activation of the two sensory modalities increases the probability that at least one sensory event will be detected. The second is that there does seem to be a limit to the amount of sensory input that can be attended to at one time. Therefore, if a light is flashed at the same time that a tone is presented, the observer may or may not detect the tone. If they are offset in time, so that the flash precedes the tone, this is less likely to happen.

Colavita (1974) seems to be the first one who described the above inhibitory effect and the effect when observers failed to respond to an auditory stimulus occurring simultaneously with a visual stimulus, known as the Colavita effect. The Colavita effect is more likely to occur if the auditory and visual events have a lower probability of co-occurrence, and if the events are spatially co-located (Koppen and Spence, 2007; Sinnett, Spence, and Soto-Faraco, 2007). In a natural environment, when an auditory and a visual stimulus occur simultaneously and in the same location, there is a good probability that they were caused by the same event and less obvious characteristic of the event is more easily missed.

The Colavita effect is reduced if the sound occurs prior to the visual event (Koppen and Spence, 2007). In a natural environment, sound often serves as a cue to visual events occurring outside our visual focus. Although a sound may sometimes distract an observer from the task of detecting a visual target (Turatto, Benso, Galfano, and Umilta, 2002), in general it signals a possible visual event – drawing the observer's attention to the visual target.

Auditory-Visual Synergy and Redundancy

Given that both the auditory and visual information from a single event will inform the observer about that event, it is not surprising that research has focused on the effects of information redundancy. The primary finding from such studies is that observers are faster when responding to redundant bimodal stimuli (e.g., a light and a sound) than they are to either of the component unimodal stimuli alone. Such signals are redundant because observers are instructed to make the same response to the presence of either the light or the tone. This *redundant signals effect* (RSE) (Miller, 1982, 1986) is greater for spatially congruent than for spatially incongruent stimuli (Gondan et al., 2005) and might be the result of the separate processing of the two different signals, with the response triggered by whichever processing finishes first. Based solely on the theory of probability, the resulting *race model* predicts the improvement because the two signals would produce a faster response than either signal alone.

If the response time is faster than that predicted by the race model, then the evidence supports the *coactivation model* which proposes that the two separate signals are integrated and processed together. (For a detailed discussion of the race and coactivation models, see Miller [1982]). Recent behavioral and electrophysiological studies have provided strong evidence for the coactivation by redundant bimodal stimuli of separate brain areas responsible for processing unimodal sensory information, and thus support the coactivation model (Giard and Peronnet, 1999; Molholm et al., 2002). For example, Giard and Peronnet devised two objects. Each object could be defined by a visual feature alone, an auditory feature alone, or by a combination. Object A consisted of a circle that morphed into a horizontal ellipse and/or a 540 Hz tone. Object B consisted of a circle that morphed into a vertical ellipse and/or a 560 Hz tone. Each of the six possible stimuli was presented equally often and observers made a speeded discrimination. Observers identified the objects more rapidly and more accurately when both features were presented than when presented with either visual or auditory features alone. Neurophysiologically, they found that event-related potential (ERPs) to multimodal objects were temporally, spatially, and functionally

---

[3] A *Gabor patch* is a luminance profile where the intensity at the center is the maximum grayscale value and the intensity at the edge of the diameter is one grayscale step above the background.

distinct from those to unimodal objects and these differences appeared very early in the processing of the objects (e.g., within 200 ms poststimulus).

Auditory-visual search

In addition to altering perceptual judgments and facilitating processing of redundant targets, information from a different modality may facilitate processing in other ways. Bolia, D'Angelo, and McKinley (1999) found that auditory cues speeded responses to targets in a visual search task. The targets were configurations of 2 or 4 LEDs amongst distractors consisting of 1 or 3 LEDs. The total number of targets and distractors (i.e., the set size) was 1, 5, 10, 25, or 50 items. Auditory cues were pink noise that was presented either from a loudspeaker at the same location as the target or at the same virtual location via spatialized headphone presentation. Auditory cues that were co-located with targets resulted in search times that did not increase significantly with increasing set size. Virtual auditory cues produced response-times (RTs) that increased with set size but only by 40 ms per item compared to increases in search time of more than 240 ms per item for trials in which no auditory cues were presented. This study showed the benefit of adding redundancy via auditory information by speeding localization of a visual target, even in the presence of non-target distractors. Although Bolia, et al., do not explicitly identify attention as the source of this facilitation, this is consistent with findings from studies that specifically address the role of multisensory attention as we shall see later in this chapter.

Auditory-visual synchrony

When designing or purchasing a HMD device that has both auditory and visual displays, it is important to consider the degree to which the auditory output is synchronized with the visual output. Obviously, perfect synchrony, though optimal, may not be possible due to technical constraints. The human brain is accustomed to a certain amount of asynchrony between auditory and visual information due to the fact that sound travels more slowly than light and as such, our tolerance for asynchrony is asymmetric (Stone et al. 2001). A number of studies have been conducted to determine the limits of our ability to detect auditory-visual asynchrony. To some extent, these limits depend on the type of auditory-visual information being transmitted. Vatakis and Spence (2006a) found that the stimulus onset asynchrony (SOA) required for detecting asynchrony between video and audio clips was lowest for simple non-speech sounds, higher for speech, and highest for piano and guitar music. They suggest that tolerance increases as the source familiarity decreases and the complexity increases (Vatakis and Spence, 2006b). The visual portion of speech can lead audition by more than 240 ms before asynchrony becomes noticeable (Dixon and Spitz, 1980; Grant and Greenberg, 2001; Grant, van Wassenhove, and Poeppel, 2003; Munhall et al., 1996). This limit is supported by neurophysiological research that shows that the temporal interval during which multisensory enhancement can occur in animals is about 200 ms (King and Palmer, 1985; Meredith, 2002; Meredith, Nemitz, and Stein, 1987; Stein and Meredith, 1993). The window is smaller for nonspeech items; auditory lags of 112 to 188 ms can be detected (Dixon and Spitz, 1980; Lewkowicz, 1996). These same studies show that there is less tolerance for lagging vision; the limits found ranged between 40 to 80 ms, with the exception of Dixon and Spitz, who found tolerance for lags up to 131 ms for speech passages. Therefore, a conservative guideline might be that visual output should lead sound by no more than 100 ms and lag by no more than 40 ms. Any asynchrony larger than this may be noticeable, depending on the source.

Capture effect

Another important consideration when designing or choosing an audio system for an HMD is to realize that information received in one sensory channel can be affected by information received through another channel. This phenomenon is called the *capture effect*. One of the most familiar examples of this phenomenon is the *ventriloquism effect* (VE) (Howard and Templeton, 1966). The VE refers to our tendency to perceive sounds as coming from the same location as a visual event, as would be the case of perceiving the sound as coming from the ventriloquist's dummy. In this case, the location of a visual object, the dummy, captures the perceived location of the sound source, the ventriloquist. Thomas (1941) describes the tendency for listener judgments of sound source location to be biased in the direction of a flickering light, especially if the light is in sync with the sound. The perceived location can either be fused with the visual source, or shifted towards the source.

The effect is strong and compelling for smaller angles of 20° to 30°. Thurlow and Jack (1973) report that it is greatly decreased at 60° (but still occurred at least some of the time for 6 out of 10 participants). Although Thurlow and Rosenthal (1976) observed some capture at 170°, it is probable that this is due to the human tendency to confuse the auditory location of sounds near 0° and 180° (see Chapter 12, *Visual Perceptual Conflicts and Illusions*).

In the case of the ventriloquist, the percept of the sound source location is *fused* with the apparent visual source of the sound (Figure 14-3). Cognitive factors affect the strength of the VE by increasing the likelihood that the visual and auditory sources will be fused (Radeau and Bertelson, 1977). For example, researchers have varied the apparent probability that the visual object is the source of the sound, using video monitors, puppets and stationary objects as the visual targets (Thurlow and Jack, 1973; Warren, Welch, and McCarthy, 1981) (Figure 14-4). As might be expected, the sound is more likely to be fused with the visual object if the visual object appears to be a probable source of the sound.

At times a visual object will capture the location of the auditory object and bias it towards the visual object even without them actually being perceived as a fused object. For example, Bertelson and Radeau (1981) reported that the attraction of auditory localization towards visual objects may occur even when fusion is not present, that is, the stimuli are not correlated. Thus the visual capture may depend strongly on the synchrony of auditory and visual stimulations and not necessarily on the realism of the auditory-visual pair (Radeau and Bertelson, 1977). The extent of the visual capture depends on the distance between the locations of the visual and auditory stimuli and is the strongest around the midline (Hairston, Wallace, Vaughan, Stein, Norris, and Schirillo, 2003).

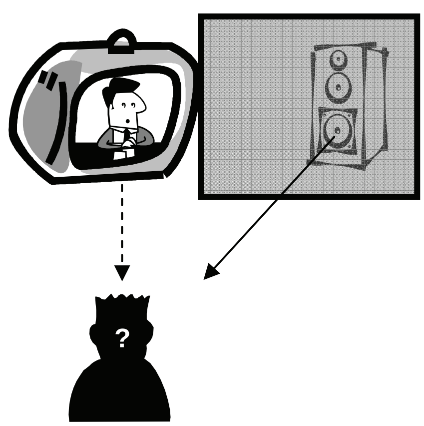


Figure 14-3. Schematic of typical demonstration of the ventriloquism effect. Listener is presented with visual stimulus along with an auditory stimulus for which the source is unseen. The location of the sound source is perceived to be collocated with the visual stimulus.

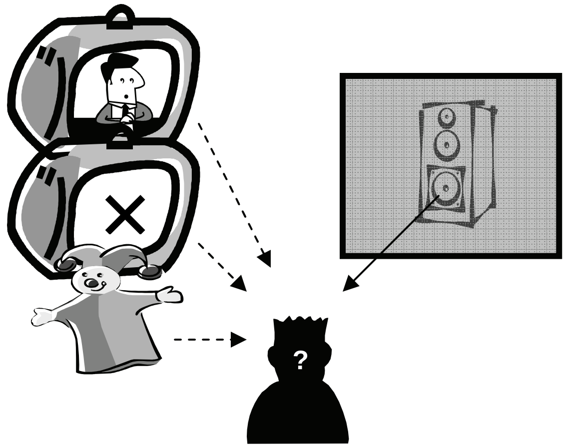


Figure 14-4. Schematic of experiment testing the effect of realism on the strength of the ventriloquism effect. In this example, the image of the human speaker is more likely to be fused with the perceived location of the sound than the video monitor with an X taped on it or the puppet. However, all three visual objects can bias the perceived location of the sound towards their location.

Vision can affect recognition of speech as well. Lip reading, called also speech reading, is used unconsciously to clarify ambiguous speech information. The posterior lateral surface of the superior temporal gyrus (located in the auditory cortex) has been found to be involved in the processing of audiovisual speech (Reale, et al., 2007). The fact that the auditory cortex has a region that processes visual information highlights the interdependence of the senses. The signal to noise ratio required for intelligibility is less for speech accompanied with a visual display of the speaker (Binnie, Montgomery and Jackson, 1974). However, in some cases, adding lip reading to auditory perception lip reading may also result in the change of the perceived sound. McGurk and MacDonald (1976) describe a phenomenon where phonemic categorization is biased depending on the visual display accompanying the auditory track. In their experiment listeners were asked to report the heard syllable (e.g., /aba/, /aga/, or /ada/). When the visual display and the auditory display were inconsistent, such as when a visual /aga/ was combined with a heard /aba/ it was often reported heard as /ada/.

If vision is biased by an auditory object, the condition is called *auditory capture*. Auditory capture occurs less frequently than visual capture. The instances where it does occur suggest that it only occurs when participants are given a reason to distrust the visual information. For example, 17 percent of the participants in Warren et al's (1981) experiment perceived the visual stimulus as shifted towards the sound source in the highly compelling condition. However, this effect is probably due to the instructions that suggested that the visual image was unreliable; they were given goggles and told that the image presented through the goggles might be "distorted." Radeau and Bertelson (1976) found auditory capture when the visual stimulus was a single light occurring in an otherwise dark environment. However, the effect disappeared as soon as the visual environment was enriched by a textured background.

Generally, vision dominates audition with respect to localization. Auditory capture of source location is most likely to occur only when visual information is ambiguous. Since vision is a more reliable source of location information, it usually results in a more compelling percept of location. Therefore, if it is necessary to convey spatial location information in a HMD display, auditory cues are best used to signal the general region of interest with a visual cue giving the precise location. However, when auditory- and visual-temporal information conflict, audition may capture vision. An elegant demonstration of this phenomenon can be found in an experiment conducted by Shams, Kamitani, and Shimojo (2000). They found that if participants were presented with stimuli consisting of a single flash and multiple auditory beeps (1-4), their perception of the number of flashes was captured by the auditory stimuli, and they consistently saw multiple flashes.

The term “capture” can be applied also to the perception of the direction of motion. In most cases, just as for stationary objects, visual motion will capture auditory objects and the sound will be heard as moving with the visual object even when the sound is stationary or moving opposite the visual motion (Kitajima and Yamashita, 1999; Mateeff, Hohnsbein and Noack, 1985). However, there is some evidence that auditory capture of motion can occur when the visual motion is ambiguous (Addie, 2003; Alais and Burr, 2004; Shimojo, Miyauchi and Hikosaka, 1997).

Outside the visual focal range, visual information is relied upon less. Audition has the advantage of being able to alert the listener to events occurring anywhere in the 360° range horizontally, as well as below and above the listener. Therefore, when visual information occurs outside one’s focal range in the periphery, auditory information is relied upon more than visual motion information (Strybel and Vatakis, 2004). Further, less synchrony is needed for fusion to occur when the objects are outside of one’s focal region (Noesselt et al., 2005). This may increase the risk that unrelated visual and auditory events will be perceived as a fused object. More likely, it can increase the overall uncertainty about the information presented because if the auditory information has terminated, it may be difficult to locate its source in space. This can be alleviated by presenting visual and auditory information jointly. This redundancy will allow the user to be alerted to the event occurring outside his visual range, and allow him to locate it and see it after its onset.

Whether capture is a factor for HMDs depends on the display and the situation. A 3-D or stereo audio display will provide fairly accurate location information for auditory information presented through the display and the visual and auditory information should match. If the system is monaural or bi-aural (same signal presented to both channels), the user will probably attribute the auditory information to one of the visual events in the display, but no spatial information is being given. However, capture can easily occur in events outside the display. A sound event triggered by an unknown source can be attributed to any plausible visual object in the vicinity. This can be the source of a tragic error. Capture can also be used to protect oneself. For example, if one has hidden a howitzer, noises made by it and the Soldiers operating it can be redirected to a decoy target placed in full view a few feet away.

Multisensory Attention

In addition to assessing processing of multisensory information, designers of HMDs also must be concerned with information that may be lost or missed, or that may cause other information to be lost or missed, particularly given the concerns about information overload. In order to address these concerns, we need to examine the role of attention within and across sensory modalities. With information available in multiple sensory modalities occasionally providing inconsistent information, it is sometimes necessary to select the information that is to be given additional processing, to the exclusion of the information not selected. While at other times, it is necessary to process both streams of information simultaneously. These two situations are commonly referred to as requiring *selective attention*, and *divided attention*, respectively.

We typically encounter more information from the environment than we can process at any one time (Johnston and Dark, 1986). This is usually true in any one sensory modality and is certainly true under normal circumstances where all manner of multimodal stimuli are present. As noted at the beginning of this chapter, most perception research has focused on a single sensory modality, and the same is true for research investigating attention. However, in the same way that perception in the natural world is multimodal in nature, attention is multimodal as well. Recognizing the multimodality of attention, recent work (within the last 10 years) has begun to investigate attention within and between individual sensory modalities.

Spence and Driver (e.g., 1994; 1996; 1997) have demonstrated extensive spatial links between touch, audition, and vision. Most of this work involves a variation of the spatial cuing task first used by Posner and colleagues (e.g., Posner, 1980). In the spatial cuing task, participants respond to a target. Prior to the appearance of the target, pre-cues appear that either correctly indicate the location of the subsequent target (a valid cue), indicate an incorrect location (an invalid cue), or provide no location information (a neutral cue). Posner and colleagues used visual targets and visual cues, but Spence and Driver have demonstrated the same cuing effects with all possible combinations of auditory, visual, and tactile targets and cues. These findings suggest that there exists a supramodal spatial attention system; such that spatial attention can be directed to an area of extrapersonal space (e.g., a portion of a visual display) by non-visual cues, and these cues will still facilitate processing of the visual information presented there. For example, Spence and Driver (1996) showed that observers more accurately localized an auditory or visual target as being above or below the midline of a display when a cue (either auditory or visual) directed them to attend to the side of the display on which the targets appeared. (For a more complete review, see Driver and Spence [1998] and Spence and McDonald [2004].)

Driver (1996) used a unique task to demonstrate an advantage for speech shadowing by the introduction of the ventriloquism effect (VE). The task was to shadow (i.e., repeat) target words presented from a loudspeaker. Distractor words were presented along with the target words from the same loudspeaker, and the task was to report only the target words while ignoring the distractor words. Above the loudspeaker was a television monitor, and an identical secondary pair of monitor and loudspeaker was positioned next to the first one. A video that showed a full frontal face view of a person speaking the target words was presented on one of the monitors. The video could appear in the monitor above the speaker that presented the words (same-side condition), or in the monitor opposite (different-side condition). Shadowing performance was significantly better in the different side condition than in the same side condition, suggesting that the presence of the visual information aided in the spatial separation of (and thus the selection of) the relevant auditory signal, the target words, from the irrelevant distractor words. The implication of this finding is that the integration of auditory and visual information can be used to functionally increase the signal-to-noise-ratio (SNR).

Consistent with Driver (1996), Santangelo and Spence (2007) found that auditory-visual cues captured attention under conditions of high and no perceptual load, however equivalent unimodal (e.g., auditory or visual) cues only captured attention in the no-load condition. They used the same spatial cuing task as Spence and Driver (1996) described above but in this study the cues were not informative as to which side of the display the subsequent target would appear. As a result, any effect of the cues on RT performance indexes the involuntary capture of attention. In addition they presented the cues under two conditions. In the high perceptual load condition, observers also had to monitor a rapidly presented stream of letters presented at fixation for occasionally presented target digits and in the no-load condition there was no centrally presented stream. The fact that such capture was not found for unimodal cues in the high load condition suggests that bimodal auditory-visual cues may be important for disengaging attention from a concurrent perceptually demanding stimulus. This may have important implications for warning signals.

Human Factors Issues in Auditory-Visual Displays

A commonly expressed argument for the inclusion of audio in a display system is that the visual system is overloaded and that additional information can be presented to the user via the auditory system. Unfortunately, this statement fails to take into account the costs of switching attention between modalities, the fit of the information to the modality, the resources shared by modalities, and the overall limits on attentional capacity. The Colavita effect illustrates this problem. When an auditory and a visual signal occur simultaneously, often the auditory signal will not be detected. However, the fact that they co-occurred increases the probability that at least one signal will be detected. Therefore, the advantage of adding an auditory component to a display system is by providing redundancy (Selcon, Taylor and McKenna, 1995) and facilitation to information presented visually. Otherwise, the auditory information is competing with the visual information for attention and may cause loss of information transfer.

The case for auditory warning signals

Redundancy is one technique to prevent or decrease information loss. Another way is to ensure that the modality used to convey information is a good match for the type of information to be conveyed. The goal is to present information in such a way that it requires very little attention and memory to recognize and respond to. For example, the fact that auditory information is dominant temporally makes it an ideal cue for observing visual changes in the environment. Morein-Zamir, Soto-Faraco, and Kingstone (2003) presented flashes from two lights placed vertically and asked participants to report which light (top or bottom) was the lagging light. For stimulus onset asynchronies shorter than the participants' visual temporal acuity, they found that an auditory cue presented just before and after the visual flashes captured the visual percept and allowed them to answer correctly. Humans are usually much quicker to detect changes in the auditory scene than in the visual scene – making sound cues ideal for alerting the user to situations requiring a fast response. Further, the visual range is limited to the frontal region of the surrounding environment, from -90° to 90° in azimuth. One's ability to focus is limited to only a small portion of that range. The auditory system, however, is able to hear sounds from the full 360° range. However, because sound is transient and spatial resolution is better for vision, visual display can serve as a redundant system, allowing an early auditory warning to be followed up by attention to the visual display.

Auditory signals by themselves should convey meaning and ideally, their meanings should be intuitive, rather than assigned (Patterson, 1990; Perry et al., 2007). For example, we have learned to expect that approaching objects will get louder. Therefore, a signal announcing an approaching aircraft should get louder as it approaches. High frequencies can indicate physical height or urgency. Urgency or severity can also be conveyed by increasing the repetition rate of a sound. Sounds in the real world are rarely tonal, and tones used in an auditory display need not be either. Timbre, the quality given to a sound by its overtones, is a natural way to convey meaning as well as to add a dimension to a signal. For example, each signal can be created from the sound inherent to the equipment it is informing the user about, and then the urgency for all can be conveyed using repetition rate or another dimension. If three-dimensional (3-D) auditory information is available, signals can be made even more meaningful by being co-located with the object of interest, drawing attention directly to the location requiring a response. It is important to consider the other auditory signals in a display and to be cautious about the meanings assigned to a dimension. If signals vary on an unimportant dimension, it will be more difficult to attend to the relevant ones (Pollack, 1970).

Earcons, auditory icons, auditory tactical signals and auditory warnings are all names given to auditory signals commonly included in a display system to represent specific events or objects. Earcons refer to arbitrary tones or tonal sequences used to convey a message in a user-computer interface (Blattner, Sumikawa and Greenberg, 1989; Gaver, 1994). An auditory icon is the mapping of a computer event to a sound, usually one with an intuitive mapping (Lucas, 1994). These are most easily understood if they are easily detected, understood and attended to. There are a number of things that should be considered when designing auditory signals for use in tactical displays.

First, the SNR should be sufficient to allow detection. Although detectability depends on a number of factors (Handel, 1989; Yost, 1994), a few basic guidelines are presented here. Ideally, the sound should be 15 dB higher than the ambient noise at all possible listening locations. However, it should not be so loud so as to cause hearing damage (Patterson, 1990). In cases where a 15 dB SNR is not possible, one must consider other factors that cause masking, such as the frequency content of the target and the background and factors that aid in sound segregation, e.g., grouping or spatial separation.

Knowledge of the way frequency components cause masking can help in the design of signals with a higher probability of detection. For example, if the noise in the environment consists primarily of speech, masking can be avoided by choosing frequency components or profiles outside the range of speech. Further, remember that low frequencies mask higher ones due to the upward spread of masking (Egan and Hake, 1950); therefore, very high frequencies should be avoided. It is also necessary to be conscious of the range of sensitivity of hearing (Fletcher and Munson, 1933). Humans are not very sensitive to sounds below 100 Hz. In addition, noise-induced and age-related hearing loss first occurs at approximately 4000 Hz and above. Thus, these frequency ranges should be avoided for allocation of signal energy. Finally, the probability that the noise will contain precisely the same frequency as the warning signal can be reduced by using a signal comprised of multiple frequencies (a complex signal). A pure tone is not a good choice for a warning tone. Instead, using a tone or complex signal that alternates between two fundamental frequencies improves the probability of detection by decreasing the probability that the auditory signal will share the same frequency content of the environmental noise.[4] The ideal choice of frequency for an auditory signal is a complex signal with a varying fundamental frequency that is lower than most of the ambient noise in the environment but with most of its spectral energy occurring outside the frequency range of the dominant ambient noise.

One can capitalize on the randomness of the noise in the environment by choosing a signal that repeats rhythmically (Patterson, 1990). This technique has several advantages. First, the regularity of the rhythm will draw attention, if the sounds in the background have irregular tempos. Second, humans are more sensitive to changing sounds than to steady state sounds. Therefore, a long continuous signal could be missed, while a repeating one will have multiple onsets to draw attention. Finally, the use of different repetition rates can add meaning and make the sound more memorable; this will be discussed later.

There should be a balance between the urgency of the sound and the annoyance and this balance should take into account the importance of the message conveyed by the sound. Further, the sound should not be so intrusive that the user is unable to respond to its message or perform other tasks. This may require a task analysis of the types of alarms that may co-occur and the kinds of tasks that will be required to respond to those alarms. Several features can make a sound seem more urgent: intensity, speed of repetitions, frequency content and envelope (onset, decay) (Edworthy, Loxley, and Dennis, 1991).

More intense sounds will seem more urgent. However, as stated before, intensity must be limited by safe presentation levels in order to avoid hearing damage. Further, a very loud sound may make communication difficult, making it difficult to respond to the emergency that triggered the sound. However, there are a couple of strategies that utilize intensity while attempting to avoid intrusiveness. For example, an auditory signal can start out at a normal level and increase in intensity if the problem is not resolved or if the problem severity increases. For example, a particular hotel clock alarm started soft, paused, and then got louder on the next repetition. Thus, if it didn't wake an individual the first time, it was more likely to be heard later. However, the individual had the option of shutting it off as soon as soon as it was heard, before it got louder. Another tactic is to present the signal initially at a very loud level, and then to drop it to a lower level in order to allow the listener time to respond. If no response is made, the signal can return to the loud level again, cycling between levels as needed. This strategy can be combined with increasing the speed of repetitions.

High priority signals can be made to sound more urgent by increasing the energy in the higher frequency components and by adding dissonance (Patterson, 1982). Unpleasant, dissonant sounds will stand out from the auditory environment and convey urgency. Dissonance refers to the lack of harmonicity of the spectral components in a sound. If the frequencies that make up a sound occur in multiples of the fundamental frequency, they will be harmonic and pleasant to the ear. If they don't, they will be dissonant and unpleasant. Three psychoacoustical properties describe how the amplitude envelope can make sounds unpleasant, sharpness, roughness and fluctuation strength. Sharpness refers to the proportion of higher frequency content in a sound. Increasing the level of frequency components above about 2700 Hz will increase the sharpness. Modulating the amplitude of a sound creates either roughness or fluctuation strength. Roughness refers to amplitude modulation between 15 and 300 Hz. Roughness is greatest at a modulation rate of 70 Hz. Roughness can be created by modulating the whole signal, but spectral variation can have a similar effect. Fluctuation strength refers to modulation below about 20 Hz. This effect is similar to that of a siren. Finally, abrupt onsets will also make the sound seem more urgent.

---

[4] For example, German police sirens alternate between two tones, in contrast to U.S. fire vehicles with a sinusoidal wailing sound.

In order for an individual auditory signal to be useful, the user must be able to remember quickly and accurately what the signal means. Pollack and his colleagues (Pollack, 1952, 1953, 1956, 1973, 1976; Pollack and Ficks, 1954; Sumby, Chambliss and Pollack, 1958) investigated the use of auditory signals for information transmission. Their findings are quite relevant to the design of memorable auditory signals. Despite the fact that listeners are able to discriminate different loudness levels and frequencies quite well, they aren't able to remember them well enough to identify the specific signals (Pollack, 1952). Therefore, designing an auditory display that uses different frequencies to distinguish between types of warning is a poor design. The listener may confuse a particular frequency with a neighboring one and misidentify the frequency. This is true even if the frequencies are spaced across a large frequency range (Pollack, 1953). At most, listeners were able to identify four or five levels of frequency; but it is recommended to limit the selection to two or three frequencies. This is true of other dimensions such as, loudness levels, repetition rate, and duration as long as they are discriminable (Pollack and Ficks, 1954). Memory for frequencies can be improved slightly if the cue frequency is combined with a reference frequency especially if the cue frequency is near to the reference frequency. It is likely that the reference is forming a salient interval that is recognizable, just as one recognizes the first few notes of a tune even if they cannot accurately identify the first note.

Despite the fact that one can identify five levels of a dimension, it is probably better to limit the set to two or three levels, especially since the user will need to be performing multiple tasks concurrently. In order to increase the number of recognizable signals, multiple dimensions should be combined. Pollack and Ficks (1954) tested listener ability to identify sounds based on levels of frequency, loudness, rate, continuity (percentage of the time "on"), duration, and spatial location. They found little improvement in the number of signals identified for sets divided into more than three levels per dimension but they could learn to identify signals distinguished by a large number of dimensions having binary values. Therefore, rather than having five different alarms that are assigned to different frequencies, they can be assigned to one or two frequencies but also vary in loudness, repetition rate, duration or location.

In an environment that has more than one signal present, care should be taken to avoid the requirement that the user have to memorize an extensive list of auditory signals. Signals should be designed to be inherently informative. One way to achieve this is to locate the sound source near the object requiring a response or the information it is cueing. If the "low battery" signal comes from the telephone, it is clear what the meaning is. Whenever possible, the sound should convey its own meaning. One way to make an auditory signal meaningful is to use a speech signal. Obviously, if the sound is, "the washer fluid is low", there's no need to memorize its meaning (Simpson, 1987). However, there are three potential problems with this approach. First, if there's already a lot of speech present in the environment, the auditory signal may be easily masked by informational masking. Further, speech can be easily susceptible to noise, especially if the spectral content is similar to or higher than the environmental noise. Finally, not all users may be as familiar with the language used and therefore may have trouble understanding the alarm.

If speech is to be used, consider the noise in the environment, the voice of the speaker, the vocabulary set and attention. Intelligibility of speech depends on the perception of its high frequency components, the consonants and these components are easily masked by noise. Speech can be preprocessed with a 3 dB/octave boost or peak clipped in order to reduce masking effects. Synthetic speech allows control of parameters such as pitch, speech rate, sex and accent, allowing it to be more easily perceived over noise. However, it is more difficult to understand and may require more attention for processing (Pisoni, 1982). Polysyllabic words are generally more intelligible than monosyllabic words. Similarly, sentences are more intelligible than single words, as they give a context that allows a listener to fill in masked information. However, since deciphering a long sentence is not recommended for time critical information it is recommended that sentences are limited to 4-8 syllables (Simpson et al., 1987). Depending on the context, a tonal alert signal, or a distinctive voice can serve to draw attention to the speech signal.

A warning should be given about the problem of excessive false alarms. Usually, a warning signal presented by a display system is a mechanical and automatic way of informing the human user of a problem or event. However, this may lead to a warning being triggered erroneously (a false alarm) or not at all (a miss). To some extent, it may be preferable for the system to err on the side of caution. However, if false alarms occur often, this may lead to a tendency by the user to ignore (Hancock, Parasuraman, and Byrne, 1996; Parasuraman, Hancock, and Olofinboba, 1997) or attempt to permanently shut off the signal (Sorkin, 1989) deeming it as an annoyance. Ideally, the system should be made as accurate as possible, with as few false alarms and misses as possible. Given that any system will have a certain amount of error, the number of false alarms can be controlled by setting the response criterion of the machinery that produces the alarm to a higher value. In some cases this will not raise the "miss rate" significantly. If this is not the case, the choice of a response criterion should be dependent on the potential danger incurred if the problem is not detected. Using an alarm that is incremental, that is one that varies in response to the changing probability that a problem exists, can reduce annoyance and increase compliance with the signal (Sorkin, Kantowitz, and Kantowitz, 1988). Finally, the user should be trained to understand the tradeoff between misses and false alarms. These considerations obviously apply not only to auditory signals but also to other types of warning signals including visual, tactile, and the signals of mixed modality.

Visual warning signals

Wickens, Gordon and Liu (1998) identify four features that are analogous to considerations for auditory warnings and should be considered when designing visual signals: visibility, discriminability, meaningfulness and location. Visibility is a concern for HMDS because not only do warnings need to be detected, but the user needs to interact with the environment while wearing it. If the display device is see-through (transparent) or monocular, care must be taken so that the display doesn't not carry so much information that it distracts the user from the real world around him.

The permanence of vision allows warnings that don't require immediate action to be postponed until the user is able to respond. In order to reduce visual clutter, information should be located in a window that can be minimized until desired. An icon can be used to remind the user of a message that is awaiting attention. If possible, the message can reside in a peripheral region of the display until it is retrieved.

Although omnidirectional auditory signals carry an advantage when it comes to quickly drawing the user's attention to information not necessarily in line of sight with the new information, they are transient and temporary in nature. Verbal messages that are longer or more complex should be transmitted visually so that the user can refer back to them and ensure that the full message is understood. If the message requires immediate action, an auditory cue can be used to call attention to the message and the message can be presented in both modalities, however, the primary mode should be visual. Care should be given that the visual message doesn't interfere with other tasks currently underway.

Some operational environments are simply too noisy for reliance on auditory displays. In others, a task analysis may reveal that the user is overburdened with auditory information. For example, a commander may be required to monitor multiple radio channels simultaneously. In these cases, visual indicators of information are preferable. Short messages that occur frequently and are part of the standard "vocabulary" can be represented by icons and other symbols. Just as with auditory warnings, care should be given to make signals as discriminable and meaningful as possible. Minimize the number of signals requiring memorization and maximize the meaningfulness of each icon.

A careful task analysis can highlight which messages are likely to be most important, and as with auditory warnings, visual warnings should be designed to reflect the urgency of the message. For example, there can be three levels of alerts: warnings, cautions, and advisories. Cautions and advisories can be presented visually because action can be postponed. When conditions are too noisy for urgent warnings to be heard, visual cues such as flashing lights can be use. Other non-essential display information can be dimmed and minimized. Estimates of likelihood can be used in order to avoid excessive false alarms (Sorkin, Kantowitz, and Kantowitz, 1988)

Indicators of locations, whether it is in form of overlays on the real world or on map displays, are best presented visually. An auditory signal can signal the general location, but a visual display has the benefit of occupying a location and remaining there as long as the information remains true or until the user is able to respond to it.

Table 14-1 summarizes some basic guidelines of when warnings should be visual and when they should be auditory. In many instances, as will be discussed in the next section, both modalities can be used effectively.

Table 14-1.
When to Use the Auditory Versus Visual Form of Presentation.

Use auditory presentation if:	Use visual presentation if:
1. The message is simple.	1. The message is complex.
2. The message is short.	2. The message is long.
3. The message will not be referred to later.	3. The message will be referred to later.
4. The message deals with events in time.	4. The message deals with location in space.
5. The message calls for immediate action.	5. The message does not call for immediate action.
6. The visual system of the person is overburdened.	6. The auditory system of the person is overburdened.
7. The receiving location is too bright or dark adaptation integrity is necessary.	7. The receiving location is too noisy.
8. The person's job requires him or her to move about continually.	8. The person's job allows him or her to remain in one position.

Source: Deatherage (1972: Table 4-1).

Auditory-visual warning signals

One way to increase meaning is to use redundant features. Sound can be combined with speech (Simpson and Williams, 1980) or visual icons to increase the probability of comprehension. We do not have to guess why our car is beeping at us in the morning because the seatbelt light is also on, and often is flashing with the same pattern as the tone. A visual cue can alert a listener to an impending auditory message. Auditory cues can signal a viewer to updates on a tactical display. An auditory cue can signal the arrival of a new message that is sent via both modalities so that if the user is busy, the message can be reviewed at a later time.

When considering the design or purchase of HMDs, one should consider the ways in which the visual and auditory displays interact with each other and with the environment in which they are used. Visual and auditory information should be consilient and thus redundant if at all possible. Rather than trying to increase information conveyed by presenting some information visually and other information auditorally, cognitive load should be decreased by coherent multimodal presentations that facilitate quick reactions. However, it is important to conduct a task analysis in order to determine when job tasks are likely to interfere with each other and incoming information from the display. The *multiple resource theory* framework consists of the following four dichotomies: stages (cognitive vs. response), sensory modalities (auditory vs. visual), codes (visual vs. spatial), and channels (focal vs. ambient) (Wickens, 2002). If two tasks are to be performed simultaneously, one task will usually suffer; however, the secondary task will usually be less difficult if they share fewer resources (Wickens, Dixon and Seppelt, 2005). Helleberg and Wickens (2001) demonstrated that verbal information presented auditorally interfered most with a visual scanning task, due to the need to writes notes – interference cause by competition for response resources. Performance was not necessarily improved for the redundant condition, perhaps because the auditory instructions disrupted focal attention and participants still relied on the visual instructions. In this case, performance was best when the information was presented visually. Multiple resource theory will be discussed in greater detail in Chapter 19, *The Potential of an Interactive HMD*. When redundancy is not feasible, care should be taken to present information via the most appropriate modality as suggested by Table 14-1.

It should also be stressed that transmission of information through auditory and visual channels must be synchronized because such synchrony facilitates the realism of the display, accurate attribution of percept to object and faster reaction times. The window during which asynchrony is undetectable depends partly on the mode and partly on the information presented, but can be conservatively defined as an visual lag of no more than 40 ms and a visual lead of no more than 100 ms.

It is desirable to have 3-D or at least, stereo sound presentation if feasible. The spatial separation of different events allows the user to attend to them better and to filter out irrelevant noise. If vision and sound are co-located in space, they are intuitively understood to be a single event and detection and response is quicker. Although capture allows us to tolerate some spatial dislocation between auditory and visual information, spatial dislocation reduces display fidelity. Further, dislocated auditory signals can, through capture, be attributed to the wrong visual events. However, a visual “master” signal located in the front of the system operator may be effectively used as a cue signal before an auditory warning signal presented in a 3-D space attracts operator’s attention to the specific location in space.

In summary, the inclusion of well-designed auditory displays in a multi-sensory HMD system can greatly reduce information loss and cognitive load. Careful considerations of the limitations of each modality allow the design of supplemental signals in the other modality that provide redundancy and prevent errors. By capitalizing on the temporal and spatial advantages of each modality, information can be easily understood and the correct responses quickly performed. This makes the auditory system an important consideration in the design or purchase of an HMD system.

References

Addie, J.D. (2003). *Spatial multi-sensory interaction: Auditory capture of visual apparent motion.* Unpublished master’s thesis, Arizona State University.

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 15, 257-262.

Allen, F., and Schwartz, M. (1940). The effect of stimulation of the senses of vision, hearing, taste, and smell upon the sensibility of the organs of vision. *Journal of General Physiology,* 24, 105-121.

Baddeley, A.D. (1982). *Your Memory: A User's Guide*. New York: Macmillan.

Bertelson, P., and Radeau M. (1981). Crossmodal bias and perceptual fusion with auditory – visual discordance. *Perception and Psychophysics*, 29, 578-584.

Binnie, C.A., Montgomery, A.A., and Jackson, P.L. (1974). Auditory and visual contributions to the perception of consonants. *Journal of Speech and Hearing Research,* 17(4), 619-630.

Blattner, M.M., Sumikawa, D.A., and Greenberg, R.M. (1989). Earcons and icons: Their structure and common design principles. *Human-Computer Interaction*, 4(1), 11-44.

Bolia, R.S., D’Angelo, W.R., and McKinley, R.L. (1999). Aurally aided visual search in three-dimensional space. *Human Factors*, 41(4), 664-669.

Broussard, I.G., Walker, R.Y., and Roberts, E.E. (1952). The influence of noise on the visual contrast threshold (No. 101). Fort Knox, KY: U. S. Army Medical Research Laboratory.

Bruce, V., Green, P.R., and Georgeson, M.A. (1996). *Visual Perception: Physiology, Psychology, and Ecology* (3rd Ed.). East Sussex, UK: Psychology Press.

Colavita, F.B. (1974). Human sensory dominance. *Perception and Psychophysics,* 16(2), 409-412.

Coleman, P.D., and Krauskopf, J. (1956). The influence of high intensity noise on visual thresholds (No. 222). Fort Knox, KY: U. S. Army Medical Research Laboratory.

Davis, E.T. (1966). Heteromodal effects upon visual thresholds. *Psychological Monographs,* 80, 24.

Deatherage, B.H. (1972). Auditory and other sensory forms of information presentation. In: van Cott, H.P., and Kinkade, R.G. (Eds.), *Human Engineering Guide to Equipment Design*. Washington, DC: Government Printing Office.

Dixon, N., and Spitz, L. (1980). The detection of audiovisual desynchrony, *Perception,* 9, 719-721.

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature,* 381, 66-68.

Driver, J., and Spence, C. (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society, Section B,* 353, 1319-1331.

Edworthy, J., Loxley, S., and Dennis, I. (1991). Improved auditory warning design: Relations between warning sound parameters and perceived urgency. *Human Factors*, 33, 205-231.

Egan, J.P., and Hake, H.W. (1950). On the masking pattern of a simple auditory stimulus. *Journal of the Acoustical Society of America,* 22, 622-630.

Fletcher, H., and Munson, W.A. (1933). Loudness, its definition, measurement and calculation. *Journal of the Acoustical Society of America,* 5, 82-108.

Gaver, W. (1994). Using and creating auditory icons. In: G. Kramer (Ed.), *Auditory Display: Sonification, Audification and Auditory Interfaces, SFI Studies in the Sciences of Complexity*, Vol. 18. Reading, MA: Addison Wesley.

Giard, M.N., and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience,* 11, 473-490.

Gondan, M., Niederhaus, B., Rösler, F., and Röder, B. (2005). Multisensory processing in the redundant target effect: A behavioral and event-related potential study *Perception and Psychophysics,* 67, 713-726.

Grant, K.W., and Greenberg, S. (2001). Speech intelligibility derived from asynchronous processing of auditory-visual information. Presented at the ISCA 2001 International Conference on Auditory-Visual Speech Processing, Aalborg, Denmark.

Grant, K.W., van Wassenhove, V., and Poeppel, D. (2003). Discrimination of auditory-visual synchrony. Presented at the ISCA 2003 International Conference on Auditory-Visual Speech Processing, St. Jorioz, France.

Hairston, W.D., Wallace, M.T., Vaughan, J.W., Stein, B.E., Norris, J.L., and Schirillo, J.A. (2003). Visual localization ability influences cross-modal bias. *Journal of Cognitive Neuroscience,* 15, 20-29.

Hancock, P.A., Parasuraman, R., and Byrne, E.A. (1996). Driver-centered issues in advanced automation for motor vehicles. In: Parasuraman, R., and Mouloua, M., (Eds.), *Automation and human performance: Theory and applications*. Mahwah, NJ: Lawrence Erlbaum.

Handel, S. (1989). *Listening: An Introduction to the Perception of Auditory Events*. Cambridge, MA: MIT Press.

Hartmann, G.W. (1933). Changes in visual acuity through simultaneous stimulation of other sense organs. *Journal of Experimental Psychology,* 16, 393-407.

Helleberg, J., and Wickens, C.D. (2001). *Auditory vs. visual data link: Relative effectiveness.* Paper presented at the *Human Factors and Ergonomics Society*, 45th Annual Meeting, Minneapolis, MN.

Howard, I.P., and Templeton, W.B. (1966). *Human Spatial Orientation.* New York: Wiley.

Ince, L.P. (1968). Effects of low-intensity acoustical stimulation on visual thresholds. *Perceptual and Motor Skills,* 26, 115-121.

Johnston, W.A., and Dark, V.J. (1986). Selective attention. *Annual Review of Psychology,* 37, 43-75.

Kitajima N, Yamashita Y. (1999). Dynamic capture of sound motion by light stimuli moving in three-dimensional space. *Perceptual and Motor Skills,* 89, 1139-1158.

Klatzky, R.L. (1975). *Human Memory: Structures and Processes*. San Francisco: Freeman.

King, A.J., and Palmer, A.R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus, *Experimental Brain Research,* 60, 492-500.

Koppen, C., and Spence, C. (2007). Seeing the light: exploring the Colavita visual dominance effect. *Experimental Brain Research,* 180(4), 737-754.

Kravkov, S.V. (1934). Changes of visual acuity in one eye under the influence of the illumination of the other or of acoustic stimuli. *Journal of Experimental Psychology,* 17, 805-812.

Kravkov, S.V. (1936). The influence of sound upon the light and colour sensitivity of the eye. *Acta Ophthalmologica,* 14, 348-360.

Kravkov, S.V. (1939). The influence of the loudness of the indirect sound stimulus on the colour sensitivity of the eye. *Acta Ophthalmologica,* 17, 324-331.

Krumbholz, K., Patterson, R., Nobbe A, and Fastl, H. (2003). *Journal of the Acoustical Society of America*, 113, 2790-2800.

Leshowitz, B. (1971). Measurement of the two-click threshold. *Journal of the Acoustical Society of America*, 49, 462-466.

Letourneau, J., and Zeidel, N.S. (1971). The effect of sound on the perception of color. *American Journal of Optometry and Archives of American Academy of Optometry,* 48, 133-137.

Letourneau, J.E. (1972). The effect of noise on vision. *The Eye, Ear, Nose and Throat Monthly,* 51, 441-444.

Lewkowicz, D.J. (1996). Perception of auditory-visual temporal synchrony in human infants, *Journal of Experimental Psychology: Human Perception and Performance,* 22, 1094-1106.

Lucas, P.A. (1994, November 7-9, 1994). *An evaluation of the communicative ability of auditory icons and earcons* Paper presented at the Second International Conference on Auditory Display, Santa Fe, New Mexico.

Maloney, D.M., and Welch, R.B. (1972). The effect of accessory auditory stimulation upon detection of visual signals. *Psychonomic Science*, 29, 345-347.

Maruyama, K. (1957). The effect of tone on the successive comparison of brightness. *Tohoku Psychologica Folia,* 15, 56-69.

Maruyama, K. (1959). Effect of intersensory tone stimulation on absolute light threshold. *Tohoku Psychologica Folia*, 17, 51-81.

Mateeff, S., Hohnsbein, J., and Noack T (1985). Dynamic visual capture: apparent auditory motion induced by a moving visual target. *Perception*, 14(6) 721-727.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature,* 264, 746-748.

Meredith, M.A. (2002). On the neuronal basis for multisensory convergence: A brief overview. *Cognitive Brain Research,* 14, 31-40.

Meredith, M.A., Nemitz, J.W., and Stein, B.E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors, *Journal of Neuroscience,* 7, 3215-3229.

Miller, J.O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology,* 14, 247-279.

Miller, J.O. (1986). Timecourse of coactivation in bimodal divided attention. *Perception and Psychophysics*, 40, 331-343.

Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., and Foxe, J.J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research,* 14, 115-128.

Morein-Zamir, S., Soto-Faraco, S., and Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research*, 17, 154-163.

Munhall, K.G., Gribble, P., Sacco, L., and Ward, M. (1996). Temporal constraints on the McGurk effect, *Perception and Psychophysics,* 58, 351-362.

Noesselt, T., Fendrich, R., Bonath, B., Tyll, S., and Heinze, H. (2005). Closer in time when farther in space: Spatial factors in audiovisual temporal integration. *Cognitive Brain Research*, 25, 443-458

Parasuraman, R., Hancock, P.A., and Olofinboba, O. (1997). Alarm effectiveness in driver-centered collision warning systems. *Ergonomics,* 40, 390-399.

Patterson, R.D. (1982). *Guidelines for Auditory Warning Signals on Civil Aircraft* (CAA Paper No. 82017). London: Civil Aviation Authority.

Patterson, R.D. (1990). Auditory warning sounds in the work environment. *The Philosophical Transactions of the Royal Society: Series B,* 327, 485-492.

Perry, N.C., Stevens, C.J., Wiggins, M.W., and Howell, C.E. (2007). Cough once for danger: Icons versus abstract warnings as informative alerts in civil aviation. *Human Factors,* 49, 1061-1071.

Pisoni, D.B. (1982). Perception of speech: The human listener as a cognitive interface. *Speech Technology,* 1, 10-23.

Pollack, I. (1952). The information of elementary auditory displays. *Journal of the Acoustical Society of America,* 24(6), 745-749.

Pollack, I. (1953). The information of elementary auditory displays. II. *Journal of the Acoustical Society of America,* 25(4), 765-769.

Pollack, I. (1956). Identification and discrimination of components of elementary auditory displays. *Journal of the Acoustical Society of America,* 28(5), 906-909.

Pollack, I. (1970). Depth of sequential auditory information processing. III. *Journal of the Acoustical Society of America,* 50(2 , Part 2), 549-554.

Pollack, I. (1973). Multidimensional encoding within the temporal microstructure of auditory displays. II. *Journal of the Acoustical Society of America,* 54(1), 22-28.

Pollack, I. (1976). Multidimensional encoding within the temporal micro-structure of auditory displays. III. Multistate displays. *Journal of the Acoustical Society of America,* 59(1), 148-152.

Pollack, I., and Ficks, L. (1954). Information of elementary multidi-mensional auditory displays. *Journal of the Acoustical Society of America,* 26(2), 155-158.

Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology,* 32, 3-25.

Radeau, M., and Bertelson, P. (1976). The effect of a textured visual field on modality dominance in a ventriloquism situation. *Perception and Psychophysics,* 20, 227-235.

Radeau, M., and Bertelson, P. (1977). Adaptation to auditory–visual discordance and ventriloquism in semirealistic situations. *Perception and Psychophysics,* 22, 137-146.

Reale, R.A., Calvert, G.A., Thesen, T., Jenison, R. L., Kawasaki, H., and Oya, H. (2007). Auditory-visual processing represented in the human superior temporal gyrus. *Neuroscience,* 145, 162-184.

Selcon, S.J., Taylor, R.M., and McKenna, F.P. (1995). Integrating multiple information sources: Using redundancy in the design of warnings. *Ergonomics,* 38, 2362-2370.

Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature,* 408, 788.

Shimojo, S., Miyauchi, S., and Hikosaka, O. (1997). Visual motion sensation yielded by non-visually driven attention. *Vision Research,* 37, 1575-1580.

Simpson, C. (1987). Speech controls and displays. In: Salvendy, G., (Ed.), *Handbook of human factors*. New York: Wiley.

Simpson, C., and Williams, D.H. (1980). Response time effects of alerting tone and semantic context for synthesized voice cockpit warnings. *Human Factors,* 22, 319-330.

Sinnett, S., Spence, C., and Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Perception and Psychophysics,* 69, 673-686.

Sorkin, R.D. (1989). Why are people turning off our alarms? *Human Factors Bulletin,* 32, 3-4.

Sorkin, R.D., Kantowitz, B.H., and Kantowitz, S.C. (1988). Likelihood alarm displays. *Human Factors,* 30, 445-460.

Spence, C.J., and Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms facilitate sound localization. *Journal of Experimental Psychology: Human Perception and Performance,* 20, 555-574.

Spence, C., and Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance,* 22, 1005-1030.

Spence, C., and Driver, J. (1997). Audiovisual links in exogenous covert spatial attention. *Perception and Psychophysics,* 59, 1-22.

Spence, C., and McDonald, J. (2004). The cross-modal consequences of the exogenous spatial orienting of attention. In: Calvert, G.A., Spence, C., and Stein, B.E. (Eds.), *The Handbook of Multisensory Processes*. Cambridge, MA: MIT Press.

Stein, B.E., London, N., Wilkinson, L.K., and Price, D.D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A Psychophysical analysis. *Journal of Cognitive Neuroscience,* 8, 497-506.

Stein, B., and Meredith, M.A. (1993). *The Merging of the Senses.* Cambridge, MA: MIT Press.

Stevens, J.C., and Marks, L.E. (1965). Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences of the United States of America,* 54(2), 407-411.

Stone, J.V., Hunkin, N.M., Porrill, J., Wood, R., Keeler, V., Beanland, M., Port, M., and Porter, N.R. (2001). When is now? Perception of simultaneity, *Proceedings of the Royal Society of London* 268B, 31-38.

Strybel, T.Z., and Vatakis, A. (2004). Effect of crossmodal distractors on auditory and visual apparent motion in the periphery. *Abstracts of 45th Annual Meeting of the Psychonomic Society,* 9, 97.

Sumby, W.H., Chambliss, D., and Pollack, I. (1958). Information transmission with elementary auditory displays. *Journal of the Acoustical Society of America,* 30(5), 425-429.

Symons, J.R. (1963). The effect of various heteromodal stimuli on visual sensitivity. *Quarterly Journal of Experimental Psychology,* 15, 243-251.

Thomas, G.J. (1941). Experimental study of the influence of vision on sound localization. *Journal of Experimental Psychology,* 28, 163-177.

Thurlow, W.R., and Jack, C.E. (1973). Certain determinants of “Ventriloquism Effect.” *Perceptual and Motor Skills,* 36, 1171-1184.

Thurlow, W.R., and Rosenthal, T.M. (1976). Further study of existence regions for the “Ventriloquism Effect.” *Journal of the American Audiology Society,* 1, 280-286.

Turatto, M., Benso, F., Galfano, G., and Umilta, C. (2002). Nonspatial attentional shifts between audition and vision. *Journal of Experimental Psychology-Human Perception and Performance,* 28(3), 628-639.

Vatakis, A., and Spence, C. (2006a). Audiovisual synchrony perception for music, speech, and object actions. *Brain Research,* 1111(1), 134-142.

Vatakis, A., and Spence, C. (2006b). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neuroscience Letters,* 393(1), 40-44.

Warren, D., Welch, R., and McCarthy, T. (1981). The role of visual-auditory “compellingness” in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception and Psychophysics*, 30, 557-564.

Watkins, W.H. (1964). Effect of certain noises upon detection of visual signals. *Journal of Experimental Psychology,* 67, 72-75.

Watkins, W.H., and Feehrer, C.E. (1964). *Investigations of acoustic effects upon visual signal detection.* (No. TR-64-577). Bedford, MA: United States Air Force Electronic Systems Division.

Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science,* 3, 159-177

Wickens, C.D., Dixon, S.R., and Seppelt, B. (2005). *Auditory preemption versus multiple resources: Who wins in interruption management?* Paper presented at the *Human Factors and Ergonomics Society* 49th Annual Meeting, Orlando, FL.

Wickens, C.D., Gordon, S.E., and Liu, Y. (1998). *An Introduction to Human Factors Engineering*. New York: Longman.

Yost, W.A. (1994). *Fundamentals of Hearing* (3rd Ed.). San Diego, CA: Academic Press.

15 COGNITIVE FACTORS

Gregory Francis
Clarence E. Rash

The opening chapter of this book noted that the primary goal of using helmet-mounted displays (HMDs) is to increase individual and unit performance. To meet such a goal, there must be an accurate transfer of information from the HMD to the user; and this transfer must occur at appropriate times. Ideally, an HMD would be designed to accommodate the abilities and limitations of users' cognitive processes. It is not enough for the information to be displayed (visually, auditorially, or tactually); the information must be perceived, attended, remembered, and organized in a way that guides appropriate decision-making, judgment, and action.

Cognitive science emphasizes the scientific study of human cognition through empirical measurements of human behavior. Although philosophers have been interested in human thought for thousands of years, the field of cognitive science is relatively new, barely more than 100 years old. Given that the field is in its infancy, it is not surprising that there are more questions than answers. Indeed, one of the main discoveries of the field is just how difficult human cognition is to explain. Despite tremendous advances and discoveries over the past 100 years, the major problems remain unsolved. Indeed, outside of very constrained situations, it is very difficult to predict the cognitive properties and capabilities of any given individual or group of individuals.

To appreciate the complexity of human cognition, consider the task of a Warfighter listening to auditory information with an HMD (this example is modified from a discussion in Willingham (2007), see Chapter 2, *The Human-Machine Interface Challenge*, for a similar description of processes involved in perception):

> BASE: Where are you?
> WARFIGHTER: I have just reached the top of the hill.

The whole "conversation" lasts maybe a few seconds, and it might appear that this simple question and answer process is trivial. Indeed, people frequently do this type of activity without any trouble. In reality, though, the processes involved in even this simple behavior are exceptionally complex. Figure 15-1 schematizes some of the processes that must be involved as the Warfighter answers the question.

First, the Warfighter must recognize the sounds coming from the HMD as speech rather than other kinds of sounds. Speech interpretation is quite complicated. For example, studies of speech show that there are no clear pauses between spoken words in normal speech. Instead, the end of one word flows in to the beginning of a following word. Nevertheless, the Warfighter interprets the stream of sounds as corresponding to individual words in a sentence. Once the words are recognized, the solider has to interpret the meaning of the sentence. This too is a complex process that depends on the context in which the words are presented. In some contexts, the question might not be a literal request for location, but a statement indicating that the Warfighter is not where he or she should be (e.g., Where *are* you?). In still other contexts, the word *you* might refer to a group of soldiers rather than an individual.

Once the Warfighter knows what is really being asked, a decision has to be made whether to answer. If stealth is currently required, it may be better for the Warfighter to remain quiet. If an answer should be given, the Warfighter has to decide on an appropriate answer. The Warfighter has to know whether to reply in latitude and longitude coordinates or, as in this case, in reference to local geography. In other situations an appropriate answer might have been "Almost there," or "Two minutes away."

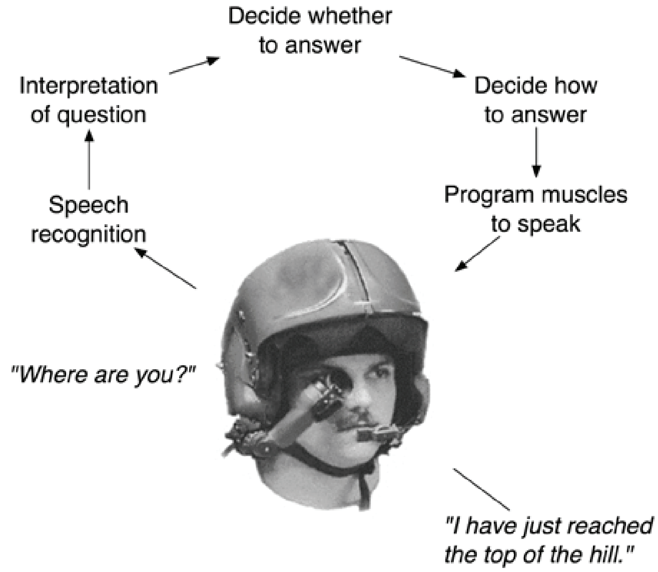


Figure 15-1. A few of the cognitive processes that may be involved in answering a simple question.

Once an appropriate answer is determined, the Warfighter has to send commands to muscles in the lips and tongue to form the speech sounds that are sent back to the base. These commands require exquisite timing to produce clear sounds.

Throughout the short conversation, the Warfighter is searching through memory for appropriate information. The soldier's memory contains all kinds of inappropriate information: details of the latest Spiderman movie, the taste of pancakes, and the name of his or her hometown. Somehow, all of this irrelevant information is not used, and instead the Warfighter selects the bits of information that are useful to the current situation.

As this brief example shows, even a simple conversation involves a complex set of processes. Cognitive scientists try to identify those processes and understand the details of each process. Each process itself can usually be broken down in to additional sub-processes that are also very complicated. Evidence for this complexity can be found in systems for artificial intelligence. There are still no computer algorithms that can interpret casual human speech, understand how to respond to simple questions, or generate speech that sounds quite like a human.

The problems in cognitive science are daunting, and even the best theories currently available are not going to give a complete description of how to analyze and design HMDs to take advantage of cognitive properties. However, such difficulties *do not* imply that studies of cognition have no advice to offer, as incomplete advice may still be better than no advice at all. There are two major contributions from cognitive science that can be applied to the design of HMDs.

The first contribution is the identification of different aspects of human cognition. University textbooks on cognitive science (e.g., Goldstein, 2005; Reed, 2004; Smith and Kosslyn, 2007; Willingham, 2007) generally organize and classify these aspects as: sensation and perception, attention, memory, knowledge, language, decision-making, and problem solving. All of these topics, and many specialized subtopics, are relevant to the use of HMDs, and they must be understood in order to optimize the usability of HMDs. If one can identify which aspect of cognition is influencing behavior, one can focus on designing the HMD to best match the known properties of that aspect. Much of this chapter is devoted to giving a brief introduction to these major topic areas and indicating how they might be related to HMD design and use.

The second major contribution from cognitive science is the development of empirical methods for studying cognition. The details of these methods are not trivial or obvious, as evidenced by their common misapplication.

Empirical reports of phenomena such as mind reading or extra-sensory perception can almost always be traced back to poor empirical measurements of human behavior and/or improper statistical control and analysis (e.g., Finegold and Flamm, 2006; Hinkle et al. 2003). Likewise, poor measurement of human cognition in the context of an HMD could lead to misunderstandings about how people will behave and interact with the system.

Many perceptual issues of HMDs have been explored in previous chapters. Here we try to focus on aspects of perception and cognition that have not already been discussed. Many researchers of cognitive science explicitly make a distinction between perception and cognition; although most researchers agree that the distinction is a fuzzy boundary with substantial overlap. Generally speaking, perception is about awareness of objects in the world, such as seeing a forest of trees fifty meters away or hearing a person walking through a forest. Cognition is about "higher-level" information processing, such as recognizing the particular forest as where you broke your arm when you fell from a tree two years ago. These processes are distinct in the sense that seeing a tree does not necessarily require committing knowledge of the tree to memory, using knowledge about the tree to guide decision-making, or attending to details of the tree's shape.

The following discussion highlights some important aspects of cognitive science as it relates to HMDs. In some cases, the discussion points out how important aspects of cognitive science have been used to better understand the design and use of HMDs. In other cases, the discussion explores where there appears to be an opportunity for future work. Many times, the cognitive relationships to HMDs are similar to the cognitive relationships for head-up displays (HUDs). Unless specifically mentioned otherwise, the following discussion generally applies to both HMDs and HUDs.

In this chapter, we will first describe methodological techniques for studying cognition, including experimental psychology, cognitive neuroscience, and computational modeling. We then discuss the general properties of cognition such as information processing and cognitive resources. Following this general overview, we explore specific subtopics of cognition, including perception, attention, memory, knowledge, decision-making, and problem solving. We then consider a variety of topics that have special interest for HMD design, including characterizations of human error, the effect of stressors on cognition, situation awareness, and workload. We then describe two case studies. One case study explores perceptual and cognitive issues of enhanced stereo-vision HMD designs. The other case study investigates visual phenomena related to a long-fielded aviation HMD, the Integrated Helmet and Display Sighting System (IHADSS). Finally, we discuss how properties of cognitive science can be applied to HMD design issues.

Methodological Techniques for Studying Cognition and Perception

Cognitive science utilizes three main techniques to study cognition: experimental psychology, cognitive neuroscience, and computational modeling. A brief discussion of each of these techniques will help set the stage for understanding how cognitive effects can be studied with regard to HMDs.

Experimental psychology

The field of experimental psychology uses scientific techniques and approaches to study behavior. The very idea of studying human behavior in a scientific way is relatively modern, dating to the 1800s (see Boring [1950] for a history of experimental psychology). Over the past 150 years, scientists have developed sophisticated techniques to isolate properties of human behavior. An important aspect of these techniques has been the development of statistical methods that analyze the experimental measurements. Most of our understanding of cognition comes from experimental studies of human behavior.

These empirical techniques reflect both the properties of the aspect of cognition that is being studied and the amount of control one has over an experiment. Most, if not all, cognitive processes are complicated and vary greatly across individuals and tasks. However, what varies and how it varies depends on what is being studied.

For example, studies of visual perception often use relatively few subjects and insist that data not be averaged across subjects. This emphasis reflects a general principle of visual perception that almost everyone behaves in roughly the same way to a carefully controlled stimulus. A key aspect of studies of perception is that a visual stimulus can be precisely defined and measured physically. This kind of control allows scientists to precisely measure differences between individual subjects. In many cases the differences are found to be quantitative rather than qualitative. That is, almost every subject behaves in a similar way (a more luminous stimulus appears brighter) but differ in the exact details (the absolute threshold for detecting a faint stimulus differs across subjects).

In contrast, studies of memory tend to use larger subject pools, and many memory phenomena are found only when data across many subjects are averaged together. This emphasis reflects the general principle that it is impossible to precisely control a memory "stimulus" because memory performance depends on many internal aspects of the subject, and these internal aspects may vary dramatically from one person to the next. Another difficulty in studying memory is that, unlike many studies of visual perception, one cannot repeat a stimulus and expect to get the same cognitive behavior. Thus, many effects can only be identified after averaging out individual differences from many observers.

Despite these (and many other) differences there are a number of empirical methods that are used in a variety of experimental studies. Table 15-1 (adapted from Smith and Kosslyn, 2007) summarizes some of the main experimental methods used in cognitive science.

Table 15-1.
Major behavioral methods used in cognitive science.

Method	Example	Advantages	Limitations
Response time	Searching for a visual target that appears on an HMD.	Objective measure of behavior; indicates the time needed for cognitive processing.	Sensitive to uncontrolled details of the experimental context; speed-accuracy trade-off.
Accuracy (percent correct)	Memory recall, such as trying to remember a radio frequency.	Objective measure of behavior.	Ceiling effects (task too easy); floor effects (task too difficult); speed-accuracy trade-off.
Judgments	Rating workload on a seven-point scale.	Easy and inexpensive to collect; assesses subjective reactions.	Participant may not know how to use scale; may not able to report on processes of interest; may not be honest.
Protocol collection (speaking aloud one's thoughts)	Talking with a pilot about how to hover a helicopter.	Can reveal a sequence of processing steps.	Cannot be used for most cognitive processes, which occur unconsciously and in fractions of a second.

Cognition inherently involves time. Although it may seem that people immediately respond to sensory inputs such as sounds and visual objects, scientific study demonstrates that such responses require time for information to be processed. One way of measuring temporal aspects of cognition is with a response time experiment.

In a response time experiment, a subject is given a task and asked to complete it as quickly as possible. A clock is started at the moment of task initiation and stopped at the moment of task completion. The time between the start and end of the task is the time needed for the subject to complete the task, i.e. the processing time. For example, a researcher may be interested in knowing how quickly a pilot can respond to a warning signal. By varying the properties of the signal, the context within which it appears, and other tasks the pilot might have to perform, one can gain insight into the cognitive mechanisms that are involved in processing the warning signal. Differences of even a few milliseconds can be important for identifying the underlying properties of cognition and in some situations can be operationally important. In general, a task that requires more cognitive processing will lead to longer response times.

One limitation of response time experiments involves the speed-accuracy trade-off. The speed accuracy trade-off refers to the general finding that errors go up when people have to respond more quickly. Giving people more time to process information generally leads to more accurate responses. This is important because it means that when comparing response times in two situations, you have to be certain that the accuracy is equivalent across the two situations.

Accuracy itself is a useful measure of behavior. Consider a memory task where a subject is shown a set of items and then later shown a test item. The subject's task is to judge whether the test item is one of the previously presented items or is a new item. The item in question could be either a visual or auditory object (e.g., a symbol or tone, respectively). A simple measure of human memory is to record the percentage of trials where the subject is correct on the task. A higher percentage indicates better memory. Such a measure can be recorded for a single subject across multiple trials of an experiment or for a single trial of an experiment across multiple subjects. Accuracy can likewise be used for any task where a correct/incorrect answer can be objectively identified.

A similar percentage statistic can also be used to measure behavior that does not have an objectively defined correct answer. For example, to measure the occurrence of visual afterimages (a percept of a visual pattern generated at the offset of a visual stimulus), a researcher would simply ask subjects to indicate whether or not they see an afterimage. There is no "correct" answer here; the subject must simply report what is seen.

Changes in percentage reports across varying conditions can be used to understand how mental mechanisms operate. For example, a researcher could measure percentage reports of afterimages with several different HMD systems. The researcher then could look for the HMD features that appear to be related to afterimage appearance and gain an understanding of what factors produce afterimages.

Two limitations of this kind of measurement are ceiling/floor effects and speed-accuracy trade offs. A ceiling effect occurs when performance is so good in all tested conditions that there is no evidence of any difference in cognitive processing. In a memory task where performance is 100% correct for all conditions, it is not possible to demonstrate that some items are more memorable than other items. This finding does not mean that there really is no difference, only that the task was so easy that the test does not demonstrate the differences. A floor effect is similar, but at the opposite extreme, where the task is so difficult that individuals are guessing.

The other measures in Table 15-1 are less objective than response time or accuracy. For judgments and protocol collection, the subject is asked to describe some aspect of their behavior or cognitive processes. These approaches are difficult to validate and depend on the subject knowing what to report. This is problematic because many aspects of cognitive processing are not consciously available (e.g., no one can describe how they remember the name of their home town, they simply "know" it).

The vast majority of investigations into cognitive factors of HMDs will use methods from experimental psychology. The techniques discussed above can be easily modified and adapted to a particular task or situation.

Cognitive neuroscience

Cognitive neuroscience tries to relate human cognition to properties of the brain. The ability to identify such relationships has blossomed over the past twenty years with the development of brain scanning techniques such as positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and evoked response potentials (ERPs). These techniques can measure brain activity in space and time while a person is performing a specific cognitive task. Prior to such techniques, neuroscience studies were limited to observations of brain-damaged patients, single cell recordings during brain surgery, and animal studies. See Gazzaniga et al. (1998) for an introduction to this topic.

Brain processes operate on many different scales, so there are many different methodological techniques for studying the brain and relating it to cognition. Figure 15-2 reproduces a graph from Churchland and Sejnowski (1988) that shows how several different experimental techniques differ in terms of temporal and spatial resolution. Notice that significant processes in the brain operate over 11 magnitudes in duration and 8 magnitudes in distance. While new technologies have improved dramatically over recent decades, there is still no single technique that is capable of covering the full range of cognitive processes in the brain.

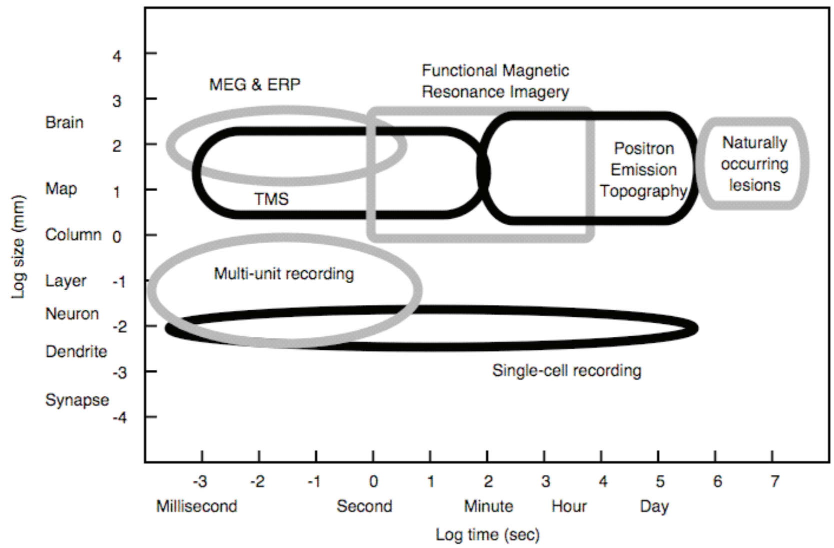


Figure 15-2. This plot shows the temporal and spatial scales of several different neuroscience techniques for studying cognitive neuroscience (adapted from Churchland and Sejnowski, 1988).

PET and FMRI track relative levels of blood flow and blood oxygen concentration (hemodynamics). PET traces a radioactive substance that is injected into the bloodstream. FMRI detects the properties of a radio wave in a strong magnetic field. The properties depend on the concentration of oxygen, which is related to blood concentration. For both of these techniques, higher blood flow to an area suggests that a brain region is involved in some cognitive task. The information from these brain scans is limited in spatial and temporal resolution. The techniques blur together responses from many thousands of individual neurons. (Neurons are a specialized type of cell that receive and send signals [messages] between the body and the brain and between different places in the brain.) Thus, one may know that a certain region of the brain is involved in a cognitive task, but be unable to identify which specific neurons in that region are involved. The same region of the brain (but different neurons) may also be involved in a quite different cognitive task. Temporal limitations are even more severe. Increases in blood flow occur in response to metabolic demands of neurons. However, changes in blood flow often lag neural responses. As a result, a sequence of mental events that occur faster than a few seconds (e.g., recalling an item from memory or responding to a warning signal) cannot be cleanly separated in the brain scan signals.

As an example where these imaging techniques have been used within the U.S. Army, a U.S. Army Medical Research and Materiel Command (USAMRMC)-sponsored PET study of complex cognitive task performance showed that during sleep deprivation there is decreased brain activity in several regions mediating higher cognitive functions and alertness; these include the prefrontal and posterior parietal cortices and thalamus (Thomas et al., 2000, 2003). Brain activity in specific subregions of the prefrontal cortex and thalamus, and an area of visual (occipital) cortex, were found to decrease across the 72 h sleep deprivation period and to correlate with the decreases in cognitive performance and slowing in saccadic velocity (Thomas et al., 2003).

Another well-known technique for measuring brain activity is the electroencephalogram (EEG), which tracks the electrical signals generated by neuron activity. These signals produce a pattern of activity across the scalp of a person. As a person engages in different cognitive tasks, the pattern of electrical activity changes across the scalp. This electrical activity also changes moment by moment during the processing of a cognitive task. This makes an EEG recording well suited to measure the temporal properties of brain events.

Unfortunately, the EEG signal is extremely noisy. The electrical signals must travel through various brain structures before reaching the scalp. To deal with the noise, researchers often average many EEG signals that are time-locked to the start of an environmental event (e.g., the appearance of a visual or auditory stimulus). The resulting averaged electrical signal is called an event related potential (ERP). The ERP signal has excellent temporal resolution and can have properties that appear to be related to certain cognitive events. On the other hand, it is often difficult to identify the spatial location of the signal that is read on the scalp. EEGs and ERPs generally have better temporal resolution than PET or FMRI, but poorer spatial resolution.

While the major techniques for measuring brain activity have been described, there are many others that are variations of these brain scanning techniques. It is not uncommon for researchers to combine several techniques to study a particular situation.

For the most part, cognitive neuroscience has yet to move beyond the research laboratories and directly influence the design or use of HMDs; this is not likely to change in the short term. The incorporation of the sciences of human factors and ergonomics into HMD design methods took more than a decade to come to fruition (some researchers would argue that the process is still continuing), and the progress of cognitive neuroscience will likely have to endure a similar progression. An understanding of which brain area is involved in a cognitive task is generally less important for HMD design than knowledge of the behavior itself. Long-term, theories and ideas from cognitive neuroscience will hopefully provide a more detailed understanding of the relationship between the brain and cognitive processing. With such knowledge, HMD design and use can be tailored to the properties of the brain. Future HMD designs may include feedback loops, where brain activity will be used to control the HMD's presentation content and duration via the neuroscience techniques described here, perhaps enhanced by other feedback signals such as oculomotor behavior.

Computational modeling

A third investigative technique for cognitive science is the use of quantitative and computational models. There is general agreement that cognition is the result of the processing of information. The goal of this approach is to identify the details of the computational basis of various cognitive mechanisms.

There is no general model of human cognition. The nature and structure of existing models differ dramatically depending on the topic that is being modeled. For example, models of some aspects of visual perception (e.g., Itti, Koch, and Niebur, 1998; Raizada and Grossberg, 2003) draw strongly from both experimental data about human perception and neurophysiological data on the brain's visual system. These complex models are often defined by thousands of mathematical equations.

In contrast, some models describe behavior without direct regard for the underlying neurophysiological mechanisms. One of the most successful models in psychology describes the time needed to make a rapid hand movement to a target of a given size (S) at a given distance (D). Fitts (1954) proposed that the following equation models the movement time (MT):

$$MT = a + b\log_2\left(\frac{2S}{D}\right). \qquad \text{Equation 15-1}$$

The terms *a* and *b* are free parameters that vary for different tasks. The term log_2 refers to the logarithm, base 2. While this equation ignores the vast complexity of the brain and cognitive processing, it nicely captures properties of human behavior and can be used to guide the design of systems for human-computer interaction (e.g., Guiarda and Beaudouin-Lafon, 2004; Francis and Oxtoby, 2006).

Still other types of models draw on ideas from computer science and artificial intelligence. For example, a cognitive architecture called Adaptive Control of Thought—Rational (ACT-R) is a system that describes how information is stored in memory and later retrieved from memory (Anderson, 1993). ACT-R tries to identify and model the procedures involved in how the brain is organized to produce cognition. Another model that combines ideas from artificial intelligence and psychology is the Executive-Process/Interactive Control (EPIC) model for human information processing (Kieras and Meyer, 1997). This cognitive architecture model tries to account for the detailed timing of human perceptual, cognitive, and motor activity. EPIC provides a framework for constructing models of human-system interaction. The model generates events (e.g., eye movements, key strokes, vocal utterances) whose timing is accurately predictive of human performance. For both of these models, a special-purpose version must be created for any given situation. A model created to explain details of reading would not be applicable to a model created for responding to warning sounds.

In principle, computational theories and models have great promise for contributing to the design and use of HMDs. A computational model that can accurately predict human behavior can reduce one of the biggest burdens on HMD design by substituting a computer model for a human subject during development. Indeed, there have been several efforts to use computational theories and models to guide computer interface design (e.g., Byrne et al., 2004; Card et al., 1983; Foyle et al. 2005; Kieras and Meyer, 1997; Liu et al., 2002). Many of these efforts have been successful in matching human data, although the models are usually not complex enough to apply outside of very limited domains. An excellent review of the successes and difficulties of applying theoretical ideas to human-computer interface design can be found in Rogers (2004).

In practice, it takes a substantial amount of work to identify what aspects of a situation need to be included in a model. In addition, one often discovers that a model cannot deal with some important details (e.g., a model of visual perception may have no stage for decision-making). There is often a difficult conundrum related to model complexity. Simple models fail to match empirical data or provide predictions of human behavior because they lack the sophistication and fluidity of human cognition. On the other hand, more complex models become mired down in issues of parameter settings. As the models become more complex, many different parts of the model contribute to many different behaviors. As a result, it becomes increasingly difficult to identify the relative contribution of any part of the model. Teasing apart the different model contributions requires an enormous amount of empirical work.

Much of the difficulty in computational modeling revolves around the fact that there is no generally agreed upon theoretical framework for how cognition operates. There is agreement that cognition involves the processing of information, but this leaves unspecified the details of how information is represented and the precise handling of the information. Without a general framework, the field of cognitive science has developed a variety of models that each deal with some particular aspect of cognition, but these models are often incompatible. For example, models of visual perception (e.g., Itti and Koch, 2001; Raizada and Grossberg, 2003) and of working memory (e.g., Baddeley, 2003) are so different that there does not appear to be a way to connect one to the other.

Cognitive Resources

The previous section suggests that experimental approaches provide the most information about cognitive processing, and that the neuroscience and computational techniques do not yet offer much additional insight to HMD design issues. While there is some truth to this suggestion, all three techniques do agree on a very important characteristic of cognition that is extremely important for HMD design- the concept of limited cognitive resources.

Cognitive resources refer to information-processing capabilities and knowledge that can be used to perform mental tasks. Different cognitive tasks seem to involve different information processing systems, and the resources and limits of these systems determine the cognitive capability to perform a given set of tasks. One of the main goals of cognitive science is to identify the properties of these systems and characterize their limits. This is true of experimental, cognitive neuroscience, and modeling approaches.

A number of cognitive science theories suggest that individuals have a limited processing capacity (e.g., Broadbent 1958; Kahneman, 1973; Lebiere et al, 2002; Posner, 1978; Wickens, 1984). The phrase *cognitive capacity* is often interchangeable with that of *cognitive resources* (Harris and Muir, 2006). Wickens (1992) disagrees with this, defining *capacity* as the maximum or upper limit of processing capability, while *resources* represent the mental effort supplied to improve processing efficiency.

One example of such a limitation can be seen in the resolution of the human eye. The best visual resolution of the eye is in a small area (approximately 1.5 millimeters diameter) of the retina known as the fovea. Visual discrimination tasks that require fine spatial detail (such as reading of small text) can only be accomplished with images that fall onto the fovea. As a result, individuals move their eyes in order to take in different parts of a scene with sufficient detail to complete the task.

A second example of processing limitations is revealed in reading, where only one thing can be read at a time. Consider the two sentences in Figure 15-3. If you focus on the x's in the middle and go from top to bottom, you can read either the sentence on the left or on the right. However, it is not possible to read both sentences at the same time. There is a fundamental limitation in the cognitive processes involved in reading that prevent dual reading.

You	x	Is
can	x	it
read	x	time
this	x	for
text	x	a
without	x	quick
any	x	snack
trouble.	x	yet?

Figure 15-3. The letters are large enough that you can read any individual word while looking at the central x. However, you cannot read the left and right sentences simultaneously (adapted from Wolfe et al., 2006).

There are similar limitations for other aspects of cognition. In a visual (or auditory) scene with several stimuli, a person can only attend to a relatively small number of stimuli simultaneously. There is some variability in estimates of the exact number, but it appears to be on the order of 1 to 4 (Cowan, 2001; Davis, 2004).

Likewise, human memory seems to be divided into several subsystems with each having its own processing limits. Long term memory (LTM) seems to have almost unlimited capacity to store new information, while short term memory (STM), or working memory, has a much smaller capacity to hold information, limited to 4 to 7 items (see Neath and Surprenant (2003) for an introduction to the properties of human memory). This limitation is easily demonstrated using Figure 15-4. Get a pencil or pen and cover the figure with a piece of paper so that the letter strings cannot be seen. Slide the paper down so that the first row can be seen. Study the letters for a few seconds, and then cover the letters with the paper. Now write down the letters in the row in exactly the same order they were given. Repeat this task for each of the next rows. Finally, check your memory performance. Most people have no trouble recalling all the items in the first few rows, but start to have difficulty recalling a list of items longer than 7 items. This limitation reflects the properties of a STM system that can only process a limited amount of information. When the list of items exceeds that system's limit, some information is forgotten.

SDK
TJPK
WZCML
CNBSKW
YLKDRWP
QSCVNTKF
BNCJWFHSL
XNBMJSGFHDK
RTDKSLCNMNBP

Figure 15-4. A demonstration of resource limitations for short term memory. See the text for details.

These cognitive limits have important consequences for the design and use of an HMD and related systems. The processing limits of human cognition emphasize that having information physically available to a person is not the same thing as ensuring that the person processes (or can process) the information. A system that presents too much information (visual, auditory, or both) may be worse than a system that leaves some information unavailable because the former overtaxes the processing capabilities of various cognitive systems.

One important aspect of cognitive processing involves assigning cognitive resources to different tasks. As described below, certain cognitive systems are involved in a variety of different tasks, and often the processing limits of those systems restrict how many tasks can be accomplished. Just as important as the processing limits is the need and ability to switch between different tasks. For example, McCann et al. (1993) found that it took effort and time to switch from processing information on a head-up display to processing information in the world.

Some cognitive behaviors seem to require very little effort. When a task is highly practiced it sometimes becomes *autonomized* and appears to require very little cognitive resources (Logan, 1988). A common example of automaticity is driving a car. A novice driver must expend a significant amount of cognitive resources to insure that many different factors are properly maintained (e.g., speed, distance from other cars, staying in the appropriate lane). With extensive practice, these monitoring activities become autonomized and happen so automatically that people are not even aware that such monitoring is occurring.

One the one hand, it is very beneficial to have certain behavior become autonomized because such behaviors are performed effortlessly and reliably. On the other hand, without conscious monitoring of behavior, autonomized behavior may be insensitive to small deviations from normal conditions and lead to inappropriate responses (Endsley, 1999).

Cognitive Functions

The previous sections of this chapter introduced some important concepts in cognitive science and mentioned some of the methods, approaches, and issues in the field. We now turn to a discussion of some of the major topic areas in cognitive science and discuss their relationship to HMDs. These topics include perception, attention, memory, knowledge, decision-making, and problem solving. Some topic areas are more important than others, and there are clear imbalances in the amount of research related to the different topic areas. Indeed, for some topics, such as visual attention, there are so many studies that it is not practical to review even a small minority of interesting findings. In contrast, for other topics there appears to be virtually no research activity.

Perception

Perception is conscious sensory experience. It is a combination of a stimulus signal producing transduction to neural receptors and cognitive mechanisms interpreting those signals. Perception deals with psychological awareness of objects in the world based on the effect of those objects on sensory systems. An integrated HMD must satisfy the user's need for visual and auditory perception. Cutting-edge systems also are incorporating haptic (touch) systems to transmit information.

Visual perception

The most basic requirement of an HMD with regard to visual perception is that the HMD needs to be able to generate light patterns that can be detected by the earliest stages of the visual system (i.e., the eye). The necessary intensity, contrast, field-of-view, spatial frequency, temporal responses, and spatial resolution for an HMD to generate appropriate stimuli for visual perception are fairly well understood. This is an important topic that has been dealt with in other chapters (2, 4, 6, 7, 10, 12, 14, 16) in this book, in previous edited books (see especially Chapters 5 and 6 in Rash [2000]), and in several reviews (Crawford and Neal, 2006; Edgar, 2007; Patterson et al., 2006). Rather than repeat this discussion, it will be fruitful to look at other aspects of perceptual experience beyond the visibility of stimuli.

Ultimately the perception of visual stimuli is an awareness of objects in the world rather than knowledge about patterns of light. Perception is *not* a copy of the retinal image. This is easily demonstrated by looking at Figure 15-5. Unless you have previously seen this image, it is quite challenging to identify how the different elements of the image group together to produce a coherent picture of an animal. Indeed, most viewers are unable to identify the object the first time they see this image.

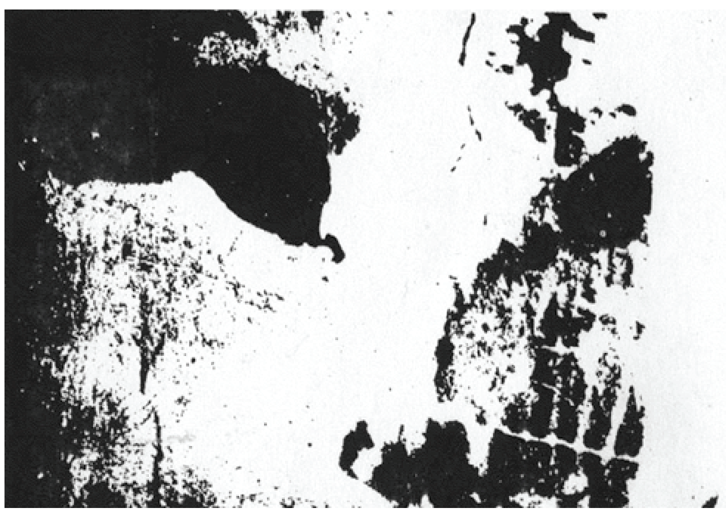


Figure 15-5. What shape do you see in this figure? If you cannot identify an animal after a few minutes, look at Figure 15-6 for clues about the shape.

An outline to help you identify the animal is given in Figure 15-6. After viewing that figure, return to Figure 15-5. It should now be fairly easy to see the object in the image. Your memory of how to organize the image elements influences your perceptual experience. In fact, you will probably never be able to see the image as it appeared the first time you saw it. Instead your memory will forever bias it to look like the object identified in Figure 15-6. Note that the retinal image has not changed at all from one viewing to the next.

One reason Figure 15-5 is difficult to interpret is because it is not clear how the black and white patterns group together. Grouping of image elements is a basic problem for visual perception. Due to occlusion from other objects, shadows from retinal veins, and noise in the physiological pathways, different parts of an object are often spatially disconnected. The visual system deals with this problem by grouping together separate parts of a visual scene to produce a coherent representation of groups of elements. This type of perceptual organization is critically important for understanding a visual scene. In many instances the process is so automatic and reliable that people do not realize that different parts of an image are being grouped together by the perceptual system.

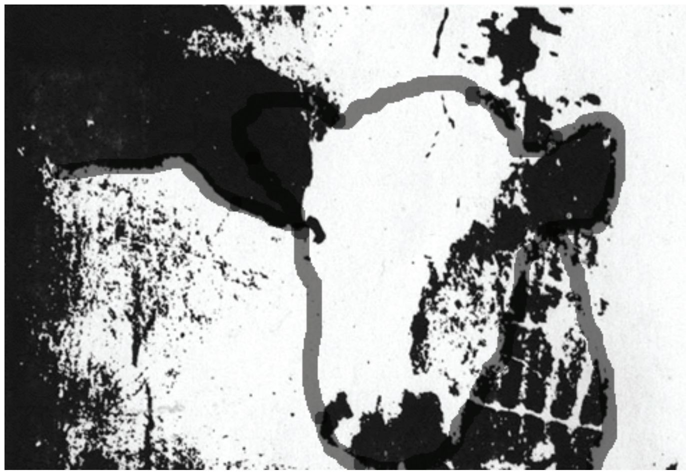


Figure 15-6. The gray lines outline the shape of a cow. Now looking at Figure 15-5 should cause you to see the cow shape.

For example, Figure 15-7a shows what appears to be a dark ink stain in front of variously oriented capital letter Bs (Bregman, 1981). Each letter B consists of multiple parts that are separated by the ink stain. The visual system is somehow able to link together the disparate parts of individual Bs to produce a coherent and meaningful perceptual experience. The presence of the ink stain appears to be an important part of this grouping process, because when it is absent, as in Figure 15-7b, the elements do not group together to from letter Bs. Even though the amount and pattern of light corresponding to the B's is the same in both images, the differences in grouping change the perceived objects in the scene.

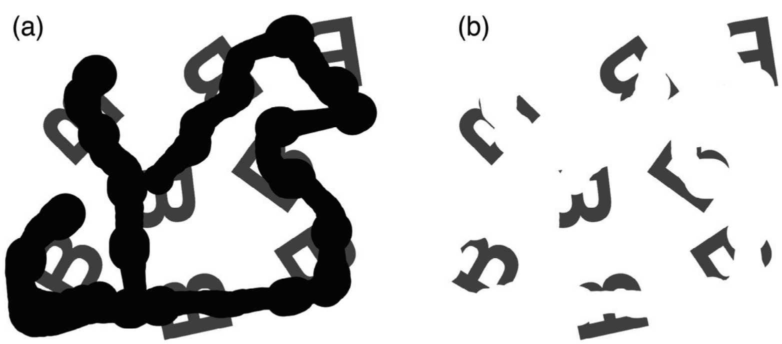


Figure 15-7. The rotated letter B's are visible in (a) when occluded by dark ink. In (b) the dark ink is replaced by the background color, which makes the B's more difficult to recognize.

Figure 15-8 shows other examples of perceptual grouping (Kanizsa, 1979; Wertheimer, 1923). In Figure 15-8a, the dots can appear to form vertical columns or horizontal rows (left) depending on the spatial proximity of the dots or their color similarity. In Figure 15-8b, there is a grouping among the slices of the Pac-Man cutouts that produces an illusory white triangle that appears to float above the other elements.

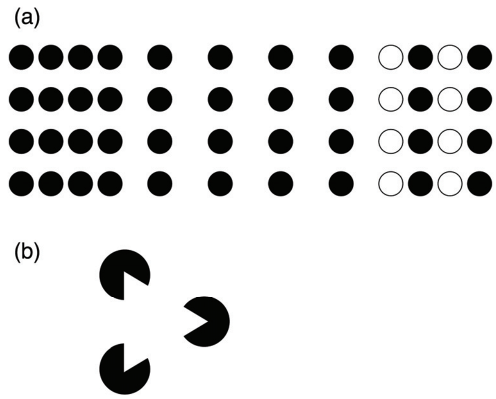


Figure 15-8. The dots in (a) can be perceived to group in to horizontal rows (left) or as vertical columns (middle and right), depending on the proximity of the dots and their colors. In (b) an illusory triangle is perceived in front of the discs.

Grouping effects such as these are very important for HMD displays. Many displays have sparse and disconnected elements that must be grouped together to form a coherent percept. Generally speaking, designers can quickly recognize when things do not group together properly because nearly every one's grouping process operates in a similar way. Still, there is the potential for inappropriate grouping if display systems are altered or used in ways that were not anticipated. Symbols that group together well under one condition, may not group together in a similar way under other conditions.

There are several computational theories of perceptual organization. Perhaps the most sophisticated theory is the neural network model proposed by Grossberg (1997). In this model the visual scene is analyzed by parallel processing streams that process boundary information (edges) and surface information (colors, brightness) in a complimentary way to identify objects in depth. A key part of this analysis involves grouping together edges in an appropriate way. However, even this theory is unable to take an arbitrary complex scene and predict how elements will be grouped. This is because the process of perceptual organization (and the model) is very sensitive to many details of the scene. A small change of color, contrast, or position can lead to a radical reorganization of the grouping of elements.

Auditory perception

In addition to visual information, many HMDs provide users with auditory information. Analogous to the presentation of visual stimuli, a basic requirement of an HMD is for the device to reliably present auditory stimuli that can be detected by the earliest stages of the auditory system (outer and inner ear). Appropriate intensities, frequencies, and durations of sounds are discussed in other chapters in this book (5, 8, 9, 11, 13, 14) and chapter 8 in Rash (2000). Rather than repeat this discussion, it will be fruitful to look at other aspects of perceptual experience beyond the detectability of stimuli.

Ultimately the perception of auditory stimuli is an awareness of sound sources in the world rather than knowledge about patterns of air pressure. A variety of auditory cues are used to derive an understanding of the location and properties of various sources in an environment. This kind of auditory scene analysis allows individuals to segregate different auditory streams. As for visual perception, these processes are often automatic and are so reliable that people do not realize that there are specific cues involved in tracking and sorting auditory streams.

For example, localizing a sound in three dimensions involves several different cues. Azimuth (horizontal location) is largely based on interaural differences, where a property of sound is different across the two ears. One such cue is an interaural time difference. A sound on your left side will reach your left ear before it reaches the right ear. Similarly, a sound on your left side will have a higher intensity in your left ear than in your right ear because the head casts an acoustic shadow that leads to an interaural level difference. Identification of a sound source's elevation is largely based on frequency cues. The head and ear decrease the intensity of some sound frequencies and increase others. These effects depend on the shape of the head and ear between the inner ear and the sound source. Sounds in different locations are influenced by different parts of the head and ear folds, and these differences change the content of the sound in a way that reveals the sound's location. Further details can be found in textbooks on perception (e.g., Goldstein, 2002; Wolfe et al., 2006) and in other chapters in this book. Many HMDs include ear coverings that interfere with normal auditory perception and only allow for verbal communication. Systems that include 3D audio earphones reintroduce many of the auditory cues described above. 3D audio systems can also be used to introduce entirely new kinds of information, so simulated sound sources at different perceived locations can provide multiple types of information. Despite substantial technological developments and enthusiasm for the idea, 3D audio systems have had limited impact in aviation cockpits (Johnson and Dell, 2003).

Auditory perception differs in many ways from visual perception. In some respects they provide complimentary information about a complex environment. For example, visual perception can provide information about objects that are too far away to be heard, while auditory cues can provide information about objects that are hidden from view. Several cockpits use these differences to insure information processing in various scenarios. For example, Martin et al. (2000) discuss how 3D audio displays operate normally even with hypoxia (introduced at a simulated high altitude). Such audio displays can thus continue to provide sound localization in situations where visual cues may start to fail. Given the differences between the two systems, it is not surprising to learn that other aspects of cognition (attention, memory, and knowledge) treat auditory and visual information differently. Some of these differences are discussed below.

Tactile perception

A third source of perceptual information in some HMDs comes from tactile interfaces that send information through the sense of touch. Chapter 18 (*Exploring the Tactile Modality for HMDs*) discusses the physiological basis of tactile perception in depth, and describes how vibrotactile interfaces might be applied to HMDs.

The primary motivation to explore the use of tactile perception is to provide a means of avoiding processing limitations imposed by the visual and auditory modalities. As discussed above, there are limits on how much information can be processed by any cognitive system. The hope is that the tactile system will complement the visual and auditory systems and provide an independent source for additional information. Such a source is not likely, however, to overcome processing limitations of higher-level (non-perceptual) cognitive systems.

One of the challenges faced by HMD designers exploring tactile perception is to identify what kinds of information can be conveyed through the tactile perceptual system. By the nature of its physiology, touch is closely tied to other perceptual experiences such as pain, temperature, and pressure. Likewise, individuals rarely use touch as a static detector but instead make specific movements to explore properties of their environment (Lederman and Klatzky, 1987). For example, lateral motion across a surface is used to reveal the texture of a surface, while static contact is used to reveal the temperature of an object.

Attention

Because all cognitive systems have limited processing capability, there is a distinction between the full spectrum of environmental stimuli and the amount of information that is actually processed. The mental processes that are involved in producing (or resulting from) this distinction are referred to as attention. The very same physical stimulus can be processed very differently when attended compared to when unattended. If someone asks you a question while you are busily thinking about something else, you may not even hear the question. The person may have to nudge you to draw your attention.

In casual conversation, people tend to use the term attention to refer to a voluntary focusing of attention. There is a feeling that one can direct attention to different aspects of the environment. In reality, attention is not based on a unitary mechanism, but involves the properties of many different cognitive systems.

Cognitive scientists make a distinction between voluntary (top-down) and involuntary (bottom-up) attention (Pashler, 1997; Posner, 1980). Voluntary attention occurs when a person makes a noticeable cognitive effort to remain focused on a particular task. Involuntary attention is often related to some environmental stimuli (such as loud sounds or flashing lights) that seem to automatically draw a person's attention.

Attention effects can occur for many cognitive processes. If someone reads a phone number to you, you may need to mentally rehearse it in order to remember it for a short period of time. If someone interrupts you to ask a question, the memory resources that you would have exerted on rehearsing the phone number are now allocated to the conversation. As a result of this reallocation, the phone number may be forgotten. When making decisions or solving problems, individuals often attend to some kinds of information more than to other kinds of information. The attended information plays a larger role in the characteristics of the decision or the arrived-at solution.

Thus, attention is a multi-faceted term that applies to many different aspects of cognitive processing. This idea is part of many views of cognition, and it plays a central role in Wickens' (1980, 1992) model of human information processing. In the model there are limited amounts of attentional resources that must be distributed effectively to complete a given task.

Attention effects can have large (and startling) impacts on behavior, and they are present at many stages of cognition. As a result, attentional effects are the most commonly studied aspect of cognition in relation to HMDs. Ideally Wickens' theory would identify how to work within the limited capabilities of each cognitive system and would predict bottlenecks in the flow of information from one system to the next. Indeed, the theory has been used for just this purpose (Wickens et al., 2005). However, the allocation of attentional resources cannot be directly measured, so it is often difficult to judge which cognitive system is ultimately limiting performance on a task.

It is not practical to consider all possible ways attention can interact with an HMD; however, this section will discuss three topics related to attention effects with HMDs: attentional allocation and information redundancy, visual search, and change blindness and cognitive tunneling. These particular topics were chosen because they apply to many different situations and highlight notable relationships between attention and HMDs.

Attention allocation and information redundancy

One of the earliest decisions that must be made in the design of a display system is what modality to present a specific piece of information. Visual images and sound are the two most commonly used modalities, but it is not always clear which is best for a given situation.

Wickens' multiple resource model (Wickens, 1980, 1992) suggests that the best modality depends on how the modalities are being used for other tasks. If the visual system is busy with many other tasks, then a visually presented stimulus may overtax the resources of the visual system and thereby lead to errors or poor performance. In such a situation, it may be better to convert some of the processing load to the auditory domain. The more general goal is to avoid resource competition, where multiple stimuli and tasks effectively compete for cognitive resources. By distributing the stimuli and tasks across separate systems, resource competition can be reduced.

On the other hand, stimuli in some modalities have bottom-up attention properties that preempt other cognitive systems. For example, an auditory stimulus seems to interfere with the processing of visual stimuli more than the other way around (Helleberg and Wickens, 2003). This is perhaps because an auditory stimulus is necessarily transient and must be acted on before being forgotten. Such preemptive effects can introduce difficulties in completing other tasks. In contrast, a static visual presentation of a stimulus will remain visible for a longer period of time. An individual can complete a current task and then investigate the visual stimulus when it will not interfere with other tasks.

Helleberg and Wickens (2003) explored modality effects for the presentation of simulated data link air traffic control (ATC) instructions. The instructions were presented either visually, auditorially, or both, while subjects flew simulated cross-country flights. The subjects in this study did not use an HMD, but the issues are relevant for both situations. At various times in the flight, ATC instructions would appear and subjects had to perform a task using the instructions.

The influence of processing the ATC instruction was measured by tracking errors in a prescribed flight path. Larger errors indicated greater difficulty in dealing with the ATC instructions. Helleberg and Wickens (2003) expected that performance would be best when the instructions were redundantly presented with both visual and auditory modalities. The auditory stimulus would be processed by a separate cognitive system from the systems involved in maintaining the flight path (largely a visual task). At the same time, the permanence of the redundant visual presentation would allow subjects to continue with a given visual task and then transfer their cognitive resources to the ATC instructions at the first available opportunity.

The empirical measures did not match the expected pattern. The best performance was for the visual presentation of the ATC instructions. The worst performance was for the auditory presentation of instructions. The performance for the redundant modality presentation was in between the other two.

This conclusion is notable because it demonstrates a common pattern in this kind of research. First, it is very difficult to use a model to predict what will happen in any particular situation. There are almost always multiple effects that work in opposite directions, and which one dominates a particular conclusion is sensitive to a great many factors. Second, the conclusion is almost always limited to the details of the experiment. The conclusions from this study are valid for the particular ATC instruction set, the flight paths, and the simulated aircraft. If any of those variables changed, the conclusion may be altered. One can easily imagine scenarios where the auditory presentation would lead to better performance than the visual presentation of ATC instructions. Moreover, as Helleberg and Wickens (2003) noted in their conclusion section, with appropriate training their subjects might have been able to learn to utilize the redundant display in a more efficient way.

Visual search

An HMD usually provides more than one piece of information at a time. A user who interacts with the visual presentation on an HMD must often search the display to identify a specific item that is relevant for a current task. This type of search is ubiquitous throughout daily experience, and there have been thousands of empirical studies that investigated the details of how such a search is performed. There are many varieties of visual search experiments, but most require the subject to observe a scene and either report when they have found a target item or to decide that the target item is not present. Measures of human performance generally include percentage correct and reaction time.

Cognitive scientists use the visual search paradigm to gain an understanding of the mechanisms and principles of cognitive systems (e.g., Treisman and Gelade, 1980; Wolfe, 1994). Both bottom-up and top-down attentional components play an important role in visual search tasks. Different display designs can alter the bottom-up attentional effects of different targets. A well-designed display will lead to bottom-up attentional effects that guide the user's attention to needed information. Likewise, top-down knowledge of the target properties can modulate the bottom-up information effects (Itti and Koch, 2001; Wolfe, Cave and Franzel, 1989).

Visual search is such a basic part of many tasks that it is often used to judge the quality of various display systems. For example, Hollands et al. (2002) used a visual search task to compare cathode ray tube (CRT) and liquid crystal display (LCD) monitors for possible use in military aircraft. They concluded that the degradation of LCD pixels with off-axis viewing made them unsuitable for some situations. (Note: Off-axis luminance and contrast in LCD monitors has greatly improved since this study.)

A potential problem for HMDs is that so much information can be placed on the visual display that it becomes difficult to find needed information. Studies of visual search suggest that the solution is to make an item-of-interest very distinct from other items (Wolfe, 1998). This solution is the basis for keeping some visual and auditory properties reserved exclusively for warnings (Smith and Mosier, 1986) or to use redundant multi-modal alarms (Nelson and Bolia, 2005). A distinct item-of-interest can be quickly found regardless of how many other items are on the display. In contrast, an item-of-interest that shares features with other items on the display may be difficult to detect and become increasingly difficult to find as more other items are present. However, in real world use of an HMD, what is labeled as an item-of-interest in one context may be clutter in a different context and vice-versa. Thus, it is difficult to make all items sufficiently distinct from other items.

Yeh and Wickens (1998) investigated a cueing approach to the clutter problem, where potential target items were cued with an arrow drawn on the HMD. Cueing led to faster reaction times and higher accuracy than if cuing was not used. Such cueing did come with a cost, however. The subjects were also asked to complete a secondary task (jamming enemy radio signals when necessary); accuracy at this secondary task was poorer when the targets were cued on the display. Presumably, the attentional pull of the cue hindered resource allocation to the secondary task. Similar results also were found for a hand-held display with similar information.

Other attempts to improve visual search include decluttering techniques (e.g., Schultz et al., 1985). With decluttering, items that are deemed to be irrelevant to the current task (e.g., commercial aircraft that are far away) are given reduced visibility or removed from the display. For this approach to be successful, one must be able to identify an algorithm for selecting irrelevant items in a way that fits the user's intuitions and expectations. St. John et al. (2005) used a decluttering heuristic for the display of a simulated naval air defense task. They found that response times to important events on the display were faster for decluttered displays than for a no-declutter display.

A common limitation of these (and many other studies) is that it is uncertain how well the results generalize to other situations. The decluttering algorithm used by St. John et al. (2005) was specially crafted for the display and task. Some displays and tasks may be more difficult to declutter. Likewise, the benefits of cueing surely depend on the task and details of the items that are being searched as well as the abilities and cognitive style of the operator.

Perhaps the most important lesson from studies of visual search is that there are multiple effects of adding information to a display. In addition to giving the user more information, the added information includes a potential cost for the user trying to find items-of-interest on the display. This conclusion echoes the experiences of HMD designers. As Newman and Greeley (1997) noted, "…there is an absolute need to keep the amount of information to the minimum necessary for the task. The reason is simple; the reason for a see-through display is to see through it."

Change blindness and cognitive tunneling

Attentional effects can be so strong that subjects will report not seeing otherwise very salient stimuli when the subjects are engaged in a demanding task (Simon and Levin, 1997). Large changes in a visual scene that co-occur with other elements appearing or disappearing, eye blinks, or movie cuts can be unnoticed even when subjects know to look for some change (Rensink et al., 1997). This effect is known as change blindness. A version of this effect can be seen in Figure 15-9, where two similar images are shown. There is a significant difference between the two images, but it is rather difficult to locate and identify the difference.[1]

Similar difficulties can be found for many situations that are relevant to environments that use HMDs. In general, individuals are not very good at noticing changes in a scene unless they are attending the object that changes. Other changes in a scene (such as gun flashes) can misdirect attention from a scene and lead to a failure to detect a significant change. Should such effects occur on an HMD during critical phases of a maneuver, the results could be devastating.

On the other hand, Triesch et al. (2003) used an HMD to set up a virtual reality situation where subjects moved (virtual) tall or small bricks to conveyer belts. On ten percent of the movements, the bricks changed height. The height change was scheduled to co-occur with an eye saccade. Subjects usually did not notice the change when the task simply involved placing the bricks on the belts regardless of brick height. As the task changed to make brick height more significant, subjects were more likely to report noticing the change in brick height. This finding suggests that the impact of change blindness is modulated by the task being performed by the subject.

Cognitive tunneling refers to a difficulty in dividing attention between two superimposed fields of information (e.g., HMD symbology as one field and see-through images as another field). It is also sometimes called attentional tunneling or cognitive capture. Some of these effects are similar to effects categorized as change blindness. In the aviation environment, such effects can lead to serious problems. Fischer et al. (1980) and Wickens and Long (1995) found that pilots sometimes did not detect an airplane on a runway when landing while using a HUD system. Clearly, the importance of the detection task is not enough, by itself, to overcome some change blindness effects. Cognitive tunneling is an extreme form of a trade-off between attending to displays and attending to the outside world. Brickner (1989) and Foyle et al. (1991) noted that a HUD improved monitoring of altitude information in a simulated flight, but at the expense of maintaining flight path. Shelden et al. (1997) suggested that cognitive tunneling can be avoided by having HUD symbology be linked to the outside world. The meta-analysis on cognitive tunneling by Fadden et al. (1998) is a good starting point for further exploration.

Memory

Human memory interacts with attention and perception effects. Indeed, many failures of attention are described as breakdowns in memory for recent events. Cognitive scientists have identified many components of memory (Neath and Surprenant, 2003). Figure 15-10 describes some of the different types of memory and their properties. One major distinction between memory systems is between short term memory (STM) and long term memory (LTM). As its name implies, short term memory deals with memory of items for relatively short periods of time (a few seconds). Generally, STM has a relatively small capacity, meaning that it can hold only a few items before some forgetting takes place. A more elaborated view of this system sometimes goes by the term *working memory* (Baddeley, 1986, 2003), which has been broken down in to a variety of subsystems that process information in a variety of ways. Different subsystems are hypothesized to deal with different types of information.

The visuospatial sketchpad is hypothesized to deal with visual short term memory (VSTM). VSTM would play an important role in, for example, monitoring a variety of potential threats on a display. Duncan et al. (1997) found that judgments of target features were faster when the features were on a common object rather than on different objects. On this basis, they suggested that only one item can be attended and held in VSTM at any moment in time. In contrast, Trick and Pylyshyn (1993) noted that subjects could reliably track three or four moving targets among a field of non-targets. This result (and others) suggests that VSTM can hold around four objects (Cowan, 2001). However, there is some debate (e.g., Davis, 2004) about the validity of these conclusions and their meaning.

[1] Look at the engine on the wing.

Figure 15-9. These two similar images have a significant difference that is surprisingly difficult to find.

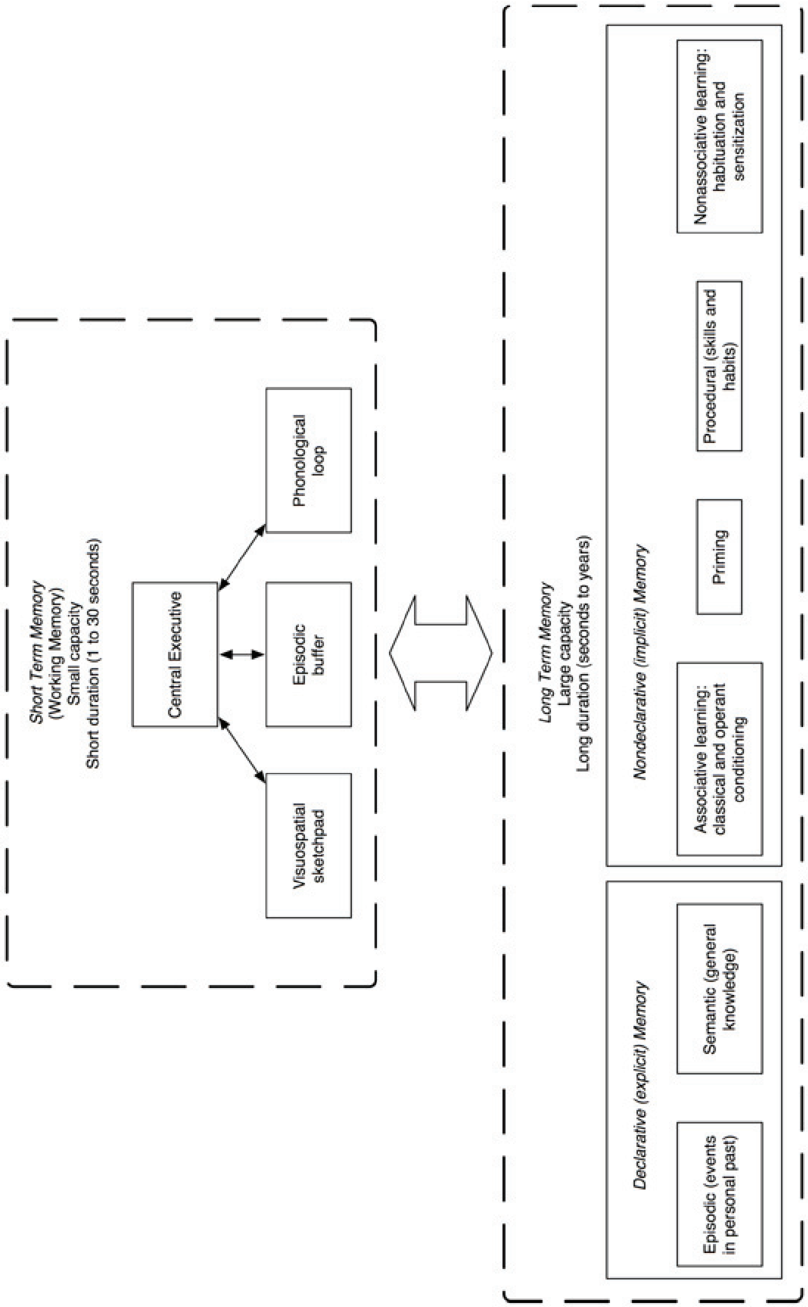


Figure 15-10. Some hypothesized memory systems.

The phonological loop is hypothesized to deal with speech and language information in working memory. Any spoken or read information that needs to be remembered for a short period of time would be held in the phonological loop. The content in the phonological loop appears to be speech sounds. This can include both spoken words and visual words that are read and then converted to speech sounds in the phonological loop. The phonological loop includes a subsystem that stores sound information for a few seconds and a system that mentally rehearses items in the loop. Forgetting often occurs because items are not rehearsed.

The working memory theory is not intended as simply a description of memory, but as a system for manipulation of information in complex tasks that involve memory. One implication of the properties of working memory is that tasks will be easier to perform if information processing can be distributed across the different components of working memory (e.g., visual and spoken information is involved) rather than weighted exclusively on one component. Wickens (1980, 1992) has made a similar observation with his multiple-resources model.

Working memory interacts with long term memory (LTM). LTM holds some information for very long periods of time (essentially a lifetime) and has a very large capacity that does not seem to be exhausted with an average human lifespan. Studies of memory suggest that LTM can be broken down in to a variety of subsystems. One major split is between declarative (explicit) and nondeclarative (implicit) memory. Declarative memory refers to memory experiences that can be explicitly recollected or declared. This includes episodic memory of particular events in your life and semantic memory, which refers to general knowledge. When you recollect what you had for breakfast this morning, you are probably recalling the memory from episodic memory. You recall the information and part of the memory involves the context in which the event occurred. On the other hand, when you recall your mother’s name, you probably recall the information from semantic memory. Here you recall the memory, but it (probably) does not include knowledge about the context in which you learned that information.

Nondeclarative memory refers to nonconscious forms of LTM that influence behavior but are not explicitly recalled. This includes knowledge that is implied rather than directly known. For example, a practiced driver knows how to hold the steering wheel to appropriately direct a car, but the driver may not be able to explain to someone else how to perform this behavior. Nondeclarative memory is often a “feeling” of knowledge.

HMD devices may alter how people remember information. In a certain sense, the HMD can become another source of memory that can be tapped to get information about past events. Hoisko (2003) describes how an off-the-shelf system of a camera, microphone, and HMD can be used as a memory prosthesis. As people rely on the HMD for representing information, they may not feel a need to remember every detail.

Knowledge

Knowledge is information in LTM and takes a variety of forms. For example, some visual information retains its spatial and temporal properties. Other visual information is converted in to a semantic form, where only the meaning of an event is recalled and not the specific details. Often a memory includes both types of information. Still other types of knowledge hold information about procedures and rules for behavior in specific situations.

Visuospatial knowledge

Mental images are one example of visuospatial knowledge. If asked to count the number of windows in their house or apartment, many individuals will form a mental image of their house and (mentally) move from room to room and count the windows. This ability suggests that some information in LTM maintains the visual and spatial characteristics of the stimuli that engendered the knowledge.

Psychologists have discovered that these kinds of mental images have many of the properties and limitations of real images. For example, Shepard and Metzler (1971) asked subjects to look at a pair of block shapes similar to those in Figure 15-11. The shapes in Figure 15-11 a have the same structure, but one is rotated relative to the

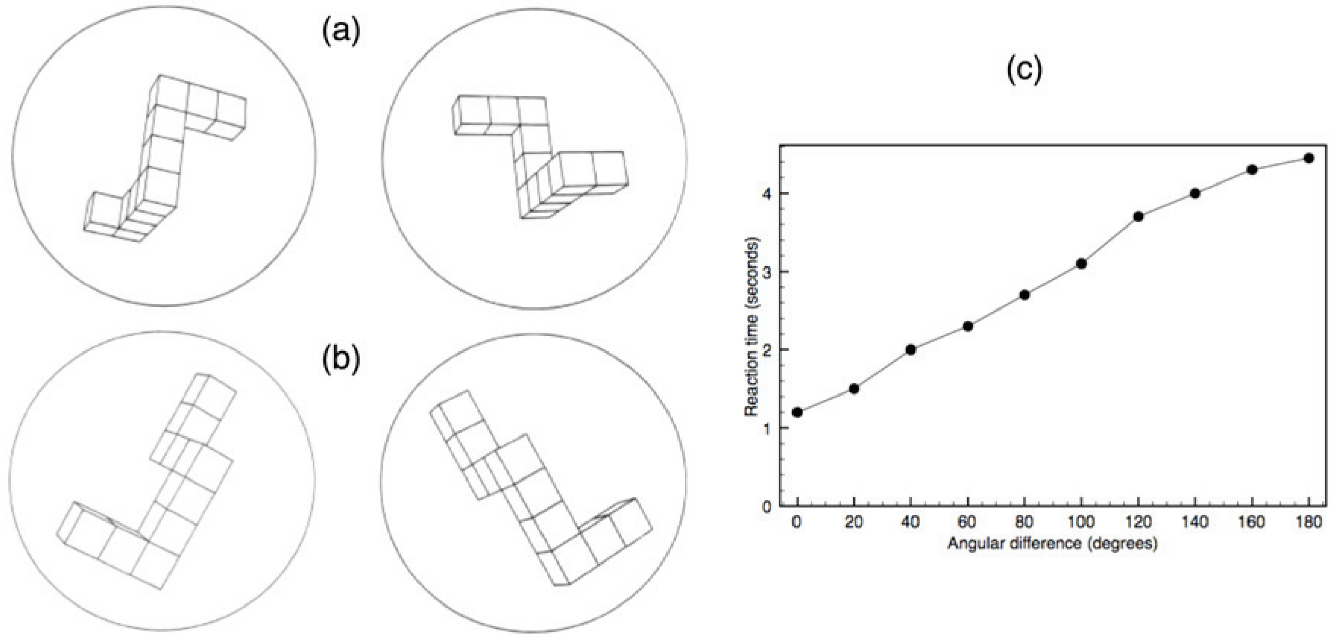


Figure 15-11. Stimuli and data for a mental rotation experiment (Shepard and Metzler, 1971). The shapes in (a) have the same structure, but one is rotated relative to the other. The shapes in (b) have different structures. The plot in (c) shows the reaction time for identifying same shapes as a function of the angular difference between the shapes.

other. The shapes in Figure 15-11b have different structures. For each pair the subject was to quickly decide whether the shapes were the same (but one rotated) or different. Many people report that they make their decision by mentally rotating one shape to match up with the other shape. Shepard and Metzler (1971) hypothesized that if the mental representation of the shape included the properties of a real image, then rotating a mental shape a given angle would require rotation through the intervening space as well. Thus, bigger angles of rotation should lead to longer delays before subjects make their decision. Figure 15-11c plots subjects' response time as a function of the angular difference between the two shapes. The results suggest that it takes about one second to rotate a shape through 50 degrees. Clearly there are similarities between the mental representation of these images and normal perception.

Semantic knowledge

Semantic knowledge refers to representations of meaningful concepts and categories. The properties of mental concepts and categories are very important to understanding other aspects of cognition. If a person has knowledge that an animal they see is a cat, that person immediately has knowledge about the concept of cats that (probably) apply to this specific cat. Thus, a person can expect that the cat likes certain kinds of foods, has a certain type of relationship with people, catches mice, has teeth, and so on. All of this knowledge can be applied without much observation of this specific cat because the information is stored as a "cat" concept based on past experience.

Studies of semantic knowledge seek to understand how this past experience is represented as a concept. One key finding from cognitive science is that many mental concepts are based on a *prototype* element. A prototype is a standard representation of items corresponding to a concept or category. It is often a conglomeration of several different examples of a category. For example, the category "birds" corresponds to a prototype that is similar to a robin. Unusual birds such as penguins and ostriches are quite different from the prototypical bird. Consistent with this representation, individuals are much faster at classifying sparrows as birds than classifying penguins as birds (Rosch, 1975).

Rosch et al. (1976) argued that there are three levels of categories. The superordinate level refers to a broad class of everyday objects, such as a transport device. The basic level corresponds to items of common use, such as a tank (in military settings). The subordinate level corresponds to a still more specific type of item, such as an M1-A1 Abrams main battle tank. Basic level categories are hypothesized to have special status in knowledge systems. Individuals can categorize basic level items faster than items at the superordinate or subordinate levels. Not surprisingly, the basic level categories for an expert of a topic (say, a tank commander) would be different from the basic level categories for a novice (Tanaka and Taylor, 1991).

Schemas

A schema is a cognitive structure that contains a sort of mental model of how the world operates within a particular situation. Schemas allow people to adapt to new situations by using knowledge about other similar situations.

If an American attempts to drive a car in the UK, the schemas involved in driving a car will both help and hinder his/her efforts. The schemas will help because nearly all cars work in a similar kind of way: turning the steering wheel changes the direction of the car, pushing the right most foot pedal accelerates the car, and so on. This kind of general knowledge transfers from one case (driving an American car) to another (driving a British car). The schemas will hinder because an American driver is used to driving on the right hand side of the road, while the British drive on the left hand side of the road. When pulling onto a road, an American driver has a tendency to immediately go to the right-hand lane, because the American driving schema indicates that this is the appropriate behavior.

Schemas are an integral part of daily life. When we encounter a new gadget and cannot figure out how to use it, the problem is usually that the way the device works is different from the schema we have in mind on how it should work. Thus, an important issue for HMDs is to insure that either the HMD is designed to match the schemas that people bring to the device, or that people can be trained to develop appropriate schemas for the device.

Along these lines, Yeh et al. (2003) investigated how people modified their attentional strategies as a function of the precision of a target cue. In their HMD a cue indicated a particular target for a user to focus on. Sometimes the cue precisely indicated where the target was located; other times the cue was less precise and only gave a general idea of where the target might be located. The precise and imprecise cues were drawn differently, so the subject could tell which precision condition they were operating with. Over time, the user developed schemas regarding how to behave with regard to these differing cues.

More generally, a user always adapts to the behavior of a system. Sometimes these adaptations are inconsistent with the expected uses of the system (Norman, 2002). An example related to HMDs involves head tracking. In some HMD systems the user's head is tracked and the image is updated appropriately to correspond to the head's orientation. In some cases, the system may appreciably lag behind the user's head movements. In normal environments, a person will make head movements in order to produce optic flow fields that contain information about the visual environment. If HMD latency interferes with the optic flow field, users will slow down their head movements in order to minimize this interference. Such a strategy involves creation of a new schema for how to extract information from the optic flow field.

There is no simple formula or rule that insures that a device's properties will match a user's schemas. The design of an HMD must include subject matter experts to understand what features the user needs and how to structure the user's interaction with the system.

Decision-making

One of the benefits of HMDs is that the user has access to a vast amount of information. Such a benefit should enable the user to make better decisions. Indeed, making better decisions, such as how to fly an aircraft or identifying where the enemy might be located, is exactly what HMDs are intended to support. As for many other cognitive issues, until an HMD is put into use, there is no easy way to be certain that it will actually lead to better (or even good) decision-making.

There is often an implicit bias to believe that individuals are, or can be trained to be, rational decision makers. From this view, the goal of an HMD is to provide the best information so that individuals can make the best choices. However, this view is incorrect. While individuals can make rational decisions, rationalism is not always what guides decision-making behavior. It should be emphasized that this is not a matter of emotional biases undermining rationality. Emotions play an important role in decision-making by characterizing the value of different options and identifying what the decider *wants*. The problem is that people can have quite reasonable and consistent emotional judgments but still make non-rational decisions. We briefly discuss a few properties of human decision-making. There is nothing special about HMDs that would influence individuals to exaggerate many of these biases, but their existence may explain why individuals behave as they do. Further details can be found in Kahneman and Tversky (1982).

Loss aversion

Individuals are generally more sensitive to the loss of a thing of value than to the gain of the very same thing. Most individuals will not take an even bet (e.g., a coin is flipped and if it comes up heads you win $5 but if it comes up tails you lose $5) because the possible loss is more aversive than the possible gain is alluring.

Loss aversion can have large impacts on an individual's behavior. For example, when choosing from a variety of possibilities that each contain positive and negative consequences, individuals tend to select the option that minimizes the perceived negative outcomes. Such a choice may not actually be the best decision, defined as giving the greatest satisfaction with the outcome.

Loss aversion can have very subtle effects on decision-making. Consider the following scenario:

> *Context A*: Suppose you are piloting a helicopter on a scouting mission. Your current position has an excellent view for observing enemy movements. However, there is a strong crosswind that pushes you uncomfortably close to trees. You decide that you need to find a new location and identify two possibilities:
>
> *Position 1*. Adequate view; little crosswind.
> *Position 2*. Good, but not excellent, view; moderate crosswind.

Most individuals will choose Position 2 rather than Position 1. The reason is found by comparing the gains and losses for the two choices relative to the current location of the aircraft. With Position 1 there is a substantial loss of view quality and a substantial gain in safety from the crosswind. Loss aversion makes the loss seem more important than the gain. For Position 2, there are similar gains and losses, but none are as extreme as for Position 1. A choice between the two positions tends to be dominated by a comparison of the relative losses. Position 1 involves more severe losses than Position 2, so most individuals prefer Position 2. The fact that Position 1 involves more gain than Position 2 is less important.

Now, consider a second scenario:

> *Context B*: Suppose you are piloting a helicopter on a scouting mission. Your current position has no crosswind, so it is relatively easy to avoid the nearby trees. However, the position provides a poor view for observing enemy movements. You decide that you need to find a new location and identify two possibilities:
>
> *Position 1*. Adequate view; little crosswind.
> *Position 2*. Good, but not excellent, view; moderate crosswind.

As in Context A, the decision-making process is dominated by the perceived losses of any potential switch. In this case, individuals tend to choose Position 1. The loss from the current location is relatively small (no crosswind to little crosswind). In contrast, Position 2 produces a larger loss (from no crosswind to moderate crosswind). The positions have similar effects on gains, with Position 2 having a larger gain than Position 1. But since losses dominate gains in decision-making, most individuals prefer Position 1.

Significantly, the two options are identical across both scenarios. For a rational decision maker, the current position of the helicopter should not make a difference in deciding between the two options. After all, the pilot wants to keep the aircraft safe and observe enemy movements. If the pilot is leaving the current position for a new one, it might seem that the properties of the current position are not relevant for judging which alternative is best. The conclusion is that humans are not rational decision makers.

Comparing alternatives

When choosing from a variety of options, individuals tend to compare pairs of options against each other. When coupled with loss aversion effects, this can lead to very unusual behavior, where an option that no one ever selects dramatically influences other selections.

For example, consider the following two decision-making contexts:

> *Context C*: Suppose you are piloting a helicopter back to an airfield. You need to return as quickly as possible, but you also need to minimize strain on the engine. Given the terrain, there are three possible routes:
>
> *Route 1*: Little engine strain, 30 minutes.
> *Route 2*: Little engine strain, 40 minutes.
> *Route 3*: Moderate engine strain, 20 minutes.

Given such a scenario, most people quickly discount Route 2 because Route 1 is a better choice. It is less clear whether Route 1 or Route 3 is the best choice over all; it depends on the chooser's personal preference and (unspecified) details of the situation. Given this set of choices, most people chose Route 1.

Now consider a second decision-making context, which differs only in the nature of Route 2:

> *Context D*: Suppose you are piloting a helicopter back to an airfield. You need to return as quickly as possible, but you also need to minimize strain on the engine. Given the terrain, there are three possible routes:
>
> *Route 1*: Little engine strain, 30 minutes.
> *Route 2:* Much engine strain, 20 minutes.
> *Route 3*: Moderate engine strain, 20 minutes.

Once again, most individuals quickly discount Route 2, but this time because Route 3 is a better choice. Once again, it is less clear whether Route 1 or Route 3 is the best choice over all; it depends on the chooser's personal preference and details of the situation. However, given this set of choices, most people chose Route 3 rather than Route 1.

Thus, the properties of Route 2, which hardly anyone ever selects, can bias individuals to choose one of the remaining alternatives. It appears that the clear advantage of Route 1 over Route 2 in context C and Route 3 over Route 2 in context D biases the decider to prefer the option with the obvious advantage.

This effect has some important implications for decision-making while using HMDs. An HMD can display a wide variety of information, and adding information to an HMD can have an influence on decisions that might not be expected. The option to display information that no user would ever chose may, nevertheless, bias the user to make selections that are not necessarily optimal.

Risk

When the choices available to people have probabilistic outcomes, they are making risky decisions. Individual's intuitions regarding the properties of probability are often incorrect, especially for small probabilities. Moreover, individuals deal with risk differently depending on whether the options available appear to be losses or gains. When the choices available to them are presented as gains or benefits, individuals tend to exhibit risk-avoiding behavior. For example, most individuals prefer option 1 from the following:

> *Context E*: You are leading a group of 600 Warfighters after completing a mission when you suddenly spot a much larger group of enemy fighters. If you stay where you are, you will be overrun and everyone will die. Your advisors identify two possible choices of action:

Option 1: Take a route that will expose part of your group to enemy fire; the best estimate is that 200 Warfighters from your group will be saved.
Option 2: Take a route that has a 1/3 probability of having no one be detected by the enemy; thereby saving all 600 Warfighters. However, there is also a 2/3 probability that the enemy will detect everyone and no one will be saved.

Both choices have the same expected value (if repeated many times an average of 200 Warfighters would be saved), but for a particular choice, individuals tend to prefer option 1 with the certain saving of 200 Warfighters.

The situation is reversed for perceived losses. Here individuals tend to be risk-seeking.

Context F: You are leading a group of 600 Warfighters after completing a mission when you suddenly spot a much larger group of enemy fighters. If you stay where you are, you will be overrun and everyone will die. Your advisors identify two possible choices of action:

Option 1: Take a route that will expose part of your group to enemy fire; the best estimate is that 400 Warfighters from your group will be killed.
Option 2: Take a route that has a 1/3 probability of having no one be detected by the enemy; thereby none of the Warfighters will be killed. However, there is also a 2/3 probability that the enemy will detect everyone and all 600 Warfighters will be killed.

Once again, both choices have the same expected value (if repeated many times an average of 400 Warfighters would be killed), but for a particular choice, individuals tend to prefer option 2 with the possibility of none of the Warfighters being killed.

The results are interesting because the choices in the two situations are identical. With 600 Warfighters in the group, saving 200 Warfighters is the same thing as having 400 Warfighters killed. What is significant is that phrasing these options as gains (Context E) or losses (Context F) biases the decision-making of individuals.

Risk-avoiding and risk-seeking behaviors are not absolute rules of decision-making. Some individuals are more prone to take risks than others, and some individuals placed in Context C may decide to go with the second option. Nevertheless, these effects tend to bias individuals in a variety of important ways in many different contexts.

In the context of HMDs, these effects indicate that great care needs to be taken in how possibilities are presented to a user. If options are presented in a way that emphasizes the gains (benefits) of different possibilities, users will tend to make decisions in a way that avoids risk. On the other hand, when options are presented in a way that emphasizes the losses associated with different possibilities, users will tend to make decisions in a way that seeks risk.

Problem solving

A problem to be solved refers to an obstacle between a present state and a goal that is not immediately obvious how to get around (Lovett, 2002). The ability to solve a problem is related to many previously discussed cognitive properties. Some problems are difficult because their solution requires keeping in mind more information than can be held by working memory. Other problems are difficult because loss aversion effects bias a person to consider only some types of possible solutions. Still other problems are difficult because a person lacks the appropriate schemas to characterize and analyze the important issues of a problem.

One important aspect of problem solving is to identify the differences between expert and novice problem solvers. Warfighters are specially trained for their duties, and are thus experts at solving certain types of problems. As a result of their training, experts in a particular field solve problems faster and with a higher success rate than novices. The key difference between expert and novice problem solvers seems to be that experts have schemas for solving problems that better fit their specialized topic area.

Experts generally have more knowledge about their field of specialization than novices. The knowledge they have is also organized differently than novices. In particular, experts often organize their knowledge in a way that indicates the fundamental aspects of solving a class of problems. One significant side effect of the differences between experts and novices is that an expert's problem solving ability tends to be restricted to a particular domain. When asked to solve problems outside of his or her area of expertise, an expert often does no better than novices (Bedard and Chi, 1992).

Special Topics

Human error

Reason (1990) describes several primary types of errors as corresponding to different cognitive stages. *Slips* and *lapses* correspond to errors in execution and/or storage of an action sequence. Here a person intends to perform an action but actually does something else. Errors of this type include forgetting to flip a switch, or shutting off the wrong engine during an emergency (Wildzunas, 1997). Slips and lapses usually occur when attentional resources are insufficient for a task or are overwhelmed by other events. For example, in the Three Mile Island accident the attentional system of a worker was overwhelmed by over 100 simultaneous warnings signals. *Mistakes* correspond to incorrect intentions or plans. Reason (1990) suggests that there are two types of mistakes, rule-based and knowledge-based. These correspond to the schemas and knowledge systems discussed above.

With advancements in technology, computer systems have replaced humans for many types of tasks. Such replacement can be beneficial because it removes the possibility for human error in a variety of circumstances. However, the computer system can only function within the range of situations that have been considered by its designers. When circumstances fall outside that range, it is necessary for a human to intervene. Significantly, the circumstances outside a device's range are inevitably situations where humans are not particularly adept at solving problems. If the problems could be easily characterized and solved, their solutions would have been built into the computer system. Thus, with advances in technology, people are increasingly asked to deal with situations for which they are not well suited. As mentioned previously, expert problem solvers are experts because they have experience and practice that create appropriate schemas to solve new problems. Errors in these kinds of systems generally occur because a sequence of unforeseen circumstances causes an unanticipated problem. There is no opportunity for an individual to become an expert at solving these kinds of problems, because the only crises that occur are those that cannot be practiced.

There are many other important issues that relate human error to properties of cognition and system management. Interested readers are advised to start with Reason (1990) for a useful introduction to the topic. Shappell and Wiegmann (2000) introduced the Human Factors Analysis and Classification System to characterize data at four levels of human-related failure: unsafe acts, preconditions for unsafe acts, unsafe supervision, and organizational influences. Each of these levels then expanded into a total of 17 causal categories that help identify how to address the appearance of error. For an example of how such a system applies to Army aviation, see Manning et al. (2004), who applied this system to an analysis of errors in military unmanned aerial vehicle accidents.

Effects of stressors

Cognitive processes are influenced by a wide variety of factors. The descriptions of cognitive factors given above generally apply to many different situations. However, different subsystems for a cognitive process may respond differently to various stressful situations. For example, Walker et al. (2005) showed that sleep is necessary for information to be encoded in long term memory. Lack of sleep can lead to poor memory performance and skill acquisition (Walker and Stickgold, 2005). Lack of sleep also affects a variety of other cognitive and perceptual systems in aviation environments (Russo et al., 2005).

Working memory is sensitive to the presence of background noise, especially if it contains phonological information that interferes with the rehearsal of information in the phonological loop (Baddeley, 1986). Gomes et al. (1999) found that exposure to large pressure amplitude low frequency noise negatively impacted memory performance of aircraft technicians, but did not significantly affect performance on an attention task.

Lieberman et al. (2005) tested several aspects of cognition under combat-like stress. Their subjects were U. S. Army Rangers and U. S. Navy SEALS engaged in relatively brief, high-intensity training missions. As a result of the training, the subjects experienced sleep deprivation, high levels of physical activity, physiological, environmental, and psychological stress, and simulated combat activities. Lieberman et al. (2005) found that all cognitive measures showed a striking decrement compared to baseline measures. The cognitive functions affected included simple behaviors such as reaction time or vigilance and more complex behaviors such as memory and logical reasoning.

Chapter 16, *Performance Effects Due to Adverse Operational Factors*, discusses the effect of stressors on perception and cognition for HMDs in more detail.

Situation awareness

Situation awareness (SA) refers to an internalized model of the current state of an environment. This internal model is believed to be the basis of decision-making, planning, and problem solving. Thus, any problems with SA will impact almost every other aspect of performance.

SA involves much more than simple perception of the world. Information in the world must be perceived, properly interpreted, analyzed for significance, and integrated with appropriate schemas that allow for a predictive understanding of the current state of the system, the system's likely future states, and appropriate behaviors from individuals within the system. A breakdown at any of the cognitive functions described above can contribute to a loss of SA.

Endsley (1999) suggests that SA involves three levels:

- Level 1: Perception of the elements in the environment: Important and relevant items in the environment must be perceived and recognized. This analysis includes elements in an aircraft (e.g., system status, warning lights) and elements external to an aircraft (e.g., other aircraft, terrain).
- Level 2: Comprehension of the current situation: Here the items from Level 1 are synthesized to produce a holistic representation of the environment. This type of synthesis requires background knowledge (schemas) that can interpret the Level 1 items to identify the relative importance of the system's current state.
- Level 3: Projection of future status: With sufficient comprehension of the system and appropriate understanding of its behavior, an individual can predict (at least in the near term) how the system will behave. Such understanding is important for identifying appropriate actions and their consequences.

In the study of Endsley (1999) perceptual issues accounted for around 80% of SA errors, while comprehension and projection issues accounted for 17% and 3% of SA errors, respectively. That the distribution of errors is skewed to the perceptual issues likely reflects the fact that errors at Levels 2 and 3 will lead to behaviors (e.g., misdirection of attentional resources) that produce Level 1 errors.

St. John et al. (in press) noted that SA is negatively affected by interruptions and multi-tasking. One of the difficulties of maintaining SA is to recover from a reallocation of cognitive resources as tasks and responsibilities change in a dynamic environment. In many respects, interruptions and multi-tasking introduce conditions for change blindness. To aid recovery of SA from these types of interruptions, St. John et al. (in press) proposed four principles on how to communicate changes in a system:

1. Automatic change detection: Since an individual will often fail to detect a change, the system should indicate when a change has happened.
2. Unobtrusive notification: An indicated change should provide information in a way that is available to the user, but not by forcing an interruption of its own (e.g., by not using a pop-up window that itself must be clicked away).
3. Overview prioritization: Changes should be listed in a way that allows the user to identify what kinds of changes are most important.
4. Access on demand: A user should be able to control how much change information is displayed.

The use of an HMD introduces both solutions and problems for SA. One the one hand, an HMD allows for information from new types of sensors and algorithms that can help guide the user’s understanding of the environment. If organized properly, such information will tend to increase SA. On the other hand, if the information is organized improperly, this information will decrease SA. Moreover, even properly organized information can lead to a deterioration of SA if there is too much data. The National Research Council (1995), in an analysis of HMDs for the Land Warrior program, identified some of the cognitive factors and their potential benefits and costs with regard to SA. Table 15-2 describes the cognitive factors likely to be affected by HMDs and the benefits and costs of such effects with regard to SA.

What this analysis makes clear is that an HMD provides a rich set of possibilities for influencing SA, both positively and negatively.

Cognitive workload

Cognitive (or mental) workload can be defined generally as the amount of cognitive processing that is required for an individual to perform a set of tasks at a given time. This is a concept that goes beyond the processing resources of cognition and is intimately related to desired performance. One cannot talk about workload unless one has a goal of what a person should accomplish. Attempts to define or study workload always have an implicit baseline of performance and attempt to identify the cognitive processes and limitations that influence performance. Workload affects performance by affecting response time (e.g., time to acknowledge and initiate a task), task completion time, throughput (how much work is accomplished during a period of time), and error rate.

Workload is both task-specific and individual-specific (Rouse et al., 1993). The amount of cognitive workload associated with a given task is affected by such factors as whether or not the task is internally (self) or externally paced, whether the task demand is constant or always changing, the presence of other simultaneous tasks, the level of consequences of task failure (internal stress), and the presence of external stressors (e.g., heat, cold, noise, etc.).

Scribner et al. (2007) measured workload in a task involving shot accuracy. They concluded that an HMD was not recommended for a shooting task because the display clutters the visual field and requires manual dexterity to interact with the display. They did note, however, that an HMD was well suited for other kinds of tasks.

Previously, attention was considered the single cognitive resource that had to be divided between multiple tasks (Wickens et al., 1988). Now, it is generally recognized that the extent to which multiple tasks can be performed simultaneously depends on whether they draw from the same resource (Navon and Gopher, 1979; Wickens, 1992; Harris and Muir, 2006).

Cognitive workload for specific tasks is measured by various approaches, including subjective ratings, analytic measures, task performance, physiological measures (e.g., heart rate, galvanic skin response, pupil diameter, blood pressure, and respiratory rate). One practical problem with the workload concept is that it is not precisely defined, and so different measures of workload do not necessarily agree with each other or tap into the effects that are fundamentally related to task performance.

Table 15-2.
Factors of HMDs affecting situation awareness.
(Based on Table 3-2 from National Research Council [1995])

Factor	Benefit	Cost
Pre-attentive processing	Salient cueing of important information.	Distraction from critical environmental cues that may flag the need to fixate attention on the environment.
Attention	Cueing to attend to important information in HMD.	Limited attention degrades effective simultaneous intake of information through similar channels.
	Integration of HMD cues with external events providing information fusion.	Attentional narrowing under high task load or stress may result in fixation on displays, interrupting attention switching to environment.
	Expansion of area and time frame over which attention is distributed.	Trained information sampling strategies and scan patterns may be disrupted by stress and high task load.
		Attention to some elements of situation may result in decrease in SA on other elements.
Working memory	Direct presentation of needed information may support limited working memory.	Extra cognitive tasks and task complexity imposed by system can seriously overload limited working memory, restricting SA and decision-making, particularly under stress.
		Information overload may occur wherein the amount of information present exceeds the amount the user can take in, threatening appropriate prioritization of information.
Information	Provides more accurate, up-to-date information to soldiers in field, and back to headquarters from field.	Information overload will pose new sorting and processing demands.
	Provides information in a different format that may be more compatible with user needs.	Information presented that is not consistent with soldier needs will slow down processing of important information.
	Enhanced sensory information.	Information that must be integrated or processed to put in needed form will slow down processing.
	Provides more accurate information on location of self and others.	

Case Study 1: Hyperstereo Helmet-Mounted Displays (HMDs)

This chapter has described key perceptual and cognitive factors that are integral to human performance, frequently alluding to their relationship to HMDs. In this section, these factors are discussed as they apply to an HMD design approach that, while not new, is rapidly becoming a leading candidate for a number of programs that incorporate an HMD as the primary display.

The defining characteristic of this specific design approach is the movement of visual inputs from directly in front of the eyes to locations on the sides of the head/helmet. The motivation for this approach includes improved center-of-mass and expanded imagery capability. This new technology introduces a perceptual illusion called hyperstereopsis, where depth perception is dramatically modified.

In order to conduct operations 24/7 and in all-weather environments, militaries have adopted two major imaging technologies: image intensification (I^2) and thermal imaging (usually referred to as forward-looking infrared [FLIR]). These two technologies operate on different physical principles: I^2-based systems require a minimum level of ambient light and operate via the principle of light amplification (McLean et al., 1998); FLIR systems produce images of the outside scene by detecting small temperature differences between objects and the background (Rash et al., 1998). I^2 and thermal FLIR imagery offer the Warfighter views of the outside world that are substantially different from normal viewing and from each other. Each technology has its advantages and disadvantages and offers functional images of the outside scene under defined lighting and thermal environments.

I^2-based devices make up the most common night imaging technology within the military. While numerous variations in these devices exist, they are collectively referred to as night vision goggles (NVGs). NVGs are heavily utilized by dismounted and mounted Warfighters. The most currently fielded version of these devices is the Aviator's Night Vision Imaging System (ANVIS), which uses enhanced 3rd generation (GEN III+) image intensifier tubes (Figure 15-12, left). The U.S. Army's next most established HMD is the Integrated Helmet and Display Sighting System (IHADSS) fielded on the AH-64 Apache helicopter (Figure 15-12, right).

Figure 15-12. Pilots wearing ANVIS (left) and IHADSS (right).

An obvious approach for the next generation of HMDs is to provide Warfighters with the capability to view both I^2 *and* FLIR imagery, either in alternation (via selective switching) or as fused imagery, with the inclusion of symbology. In addition, recent advances in synthetic imagery make it desirable to have an HMD design that also allows its presentation.

However, while a host of optical issues must be addressed, any HMD designed to explore dual sensor (and synthetic) imagery presentations must still contend with the important biodynamic characteristics of head-supported weight and center-of-mass, as well as the conflicting optical requirements. It is mostly these concerns that have forced a decision between the two imaging technologies in the past.

Over the past two decades, in an attempt to improve center-of-mass, several HMD designs have been developed that move the I^2 sensors from directly in front of the eyes to positions on the sides of the helmet. Other proposed designs have coupled this relocation of the I^2 sensors with the added capability of presenting FLIR (and synthetic) imagery via miniature displays. One optical design accomplishes this by reflecting imagery off of the visor. The ability to provide the Warfighter with multiple versions of the outside scene is a leap in HMD design that could significantly improve user performance and situation awareness. A recent study investigating the use of both I^2 and FLIR sensors in the AH-64 Apache showed that each sensor provides unique capabilities (Heinecke et al., 2007).

Recognizing these advantages, virtually all of the major avionics manufacturers have explored this design approach. The majority of these efforts, although involving comprehensive developmental programs, never progressed to full production. Several of these manufacturers are already fielding HMD designs that relocate the I^2 tubes to the sides of the helmet and provide the capability of presenting both I^2 and FLIR imagery (as well as synthetic imagery). Most of these designs were first developed for fixed-wing applications. Kalich et al. (2007) summarize many of these hyperstereo designs. Two representative systems are the Integrated Night Vision System (INVS), which is built by Honeywell, Inc., Minneapolis, Minnesota, and commercially known as the Monolithic Afocal Relay Combiner (MONARC), and the TopOwl® system, which is manufactured by Thales, France. Each system is shown in Figure 15-13.

While improving center-of-mass issues and expanding imagery capability, these designs come with certain compromises. One perceptual consequence is a phenomenon referred to as "hyperstereo vision" or "hyperstereopsis" (see Chapter 12, *Visual Perceptual Conflicts and Illusions*). Stereopsis is a cue to depth perception based upon differences in the scene projected to the two eyes. The spatial separation of the eyes means that each eye has a slightly different view of the world. These differences lead to systematic shifts in image contours that correspond to items in depth relative to where the two eyes converge to a point of focus. The calculation of relative depth depends on the lateral separation of the eyes. In hyperstereo systems, the sensors receiving the visual information are placed substantially farther apart than the user's eyes. As a result, the image shifts across the two inputs are more substantial than for normal vision.

Figure 15-13. The MONARC (Honeywell, Inc.) (left) and the TopOwl® (Thales) (right) hyperstereo HMD designs.

Hyperstereopsis manifests itself as exaggerated depth perception, which is characterized by intermediate and near objects appearing closer than normal. At close distances (< 20 feet/6 meters), the ground appears to slope upward. Because the user's body and very nearby objects can be perceived under the goggles and by non-visual cues, a user sometimes experiences a "crater" effect, where the ground seems to rise up to chest level. Hyperstereopsis effects weaken for much longer distances because objects at longer distances introduce very small differences between the two eyes and sensors.

Hyperstereopsis effects are particularly problematic in the rotary-wing environment, where the most critical maneuvers are performed at very low altitudes and near the ground. For example, a pilot will perceive the near ground as rising up. When a helicopter pilot is sitting in the aircraft on the ground, it will look as if the ground level outside the cockpit is at chest level, causing some pilots to say it looks like they are sitting in a hole (Figure 15-14). However, distant objects will appear normal.

Figure 15-14. Depiction of illusion of ground position due to hyperstereo vision. The lines represent the level of the ground as perceived by the pilot.

Hyperstereopsis can also affect other aspects of perception. Objects can appear to be closer than reality and horizontal motion can be exaggerated. The horizontal and oblique velocity and acceleration vectors will be distorted differently, making shipboard landings, nap-of-the-Earth (NOE) flight, quick-in/quick-out maneuvers, motion parallax, and flow-field interpretation problematic. It likely will be very difficult to train for dynamic environments that involve the avoidance of obstacles near the helicopter.

The effects of hyperstereopsis need not always be negative. Some atypical hyperstereo configurations (based on camera pairs with extremely wide baselines or temporal delays with a single camera) have been investigated for their possible use in aerial search and rescue, target detection, and traversing drop-off terrain tasks (e.g., Cheung and Milgram, 2000; Schneider and Moraglia, 1994; Watkins 1997). And, as presented above, hyperstereo HMD designs allow for added operational capability by allowing the option of adding FLIR and synthetic imagery presentation.

Aware of the hyperstereopsis issue, the French, German and U.S. militaries have evaluated and conducted a number of limited operational evaluations on several of these designs in an attempt to determine its impact on performance (German Air Force Test Center, 1998; Kimberly and Mueck, 1991; Krass and Kolletzki, 2001; Leger et al., 1998). These studies primarily have investigated pilot performance and have resulted in mixed findings. Consistently, the reported hyperstereo effects were characterized by intermediate and near objects appearing distorted and closer than normal. The ground appeared to slope upwards towards the observer and regions beneath the aircraft appeared closer than normal; a tendency to fly higher than normal during terrain flight was noted.

The designers of these systems claim that users can “overcome” or “train out” the hyperstereo effects. The fielding of the Thales TOPOWL® hyperstereo HMD systems by the French and German armies supports this claim. A threshold period of 8-10 hours has been suggested (Kalich et al. 2007). However, the HMD community, collectively, has not fully accepted this position.

An underlying issue in “overcoming” the perceptual effects associated with the use of hyperstereo HMDs is whether this is achieved through perceptual adaptation or cognitive compensation. Perceptual adaptation refers to changes in perceptual experience. If users actually adapt to the use of a hyperstereo HMD, then they would eventually perceive the world to be at the veridical depth (i.e., coinciding with reality). In contrast, cognitive compensation implies that the world looks non-veridical but users develop strategies for successfully interacting with the modified appearance.

Studies that may be relevant to this adaptation vs. compensation conundrum are those that have investigated the use of prisms and mirrors to manipulate and produce unusual visual inputs. In these studies, images were shifted or inverted on the retina, producing a stimulus effect on the visual system not too dissimilar from systems that produce hyperstereopsis. These studies show that initially these changes cause major disruptions in visual-motor coordination and visual perception, followed by gradual "adaptation." This alleged adaptation is accompanied by a performance recovery that approaches, but does not equal, premodification performance (Welch, 1986; Wildzunas, 1997a). Most of these studies have involved tasks such as walking, ball tossing and other close-in eye-hand coordination activities. However, none of the studies involved tasks and working distances that are congruent with those associated with helicopter flight (CuQlock-Knopp et al., 2001; Judge and Bradford, 1988; Wildzunas, 1997b). Notably, these studies do not usually distinguish between perceptual adaptation or cognitive compensation, as either adjustment would lead to improved performance in a variety of tasks.

However, there is at least one scenario where there is strong evidence that perceptual adaptation can occur. When new glasses are prescribed, moderate levels of distortion may be present. But after a period of time, the wearer adapts and perceives the world as normal. In contrast, Lindin et al. (1999) analyzed the effects of wearing inverting prisms and determined that subjects did not "see" the world as "up-right;" rather, they learned to compensate for the inversion. This implies that major changes in the visual image are dealt with through cognitive compensation rather than perceptual adaptation.

Studies that have investigated hyperstereo in real aviation environments have not attempted to differentiate between adaptation and compensation. With few exceptions, most military investigations have been trial flights or flight tests with an engineering emphasis (German Air Force Test Center, 1998; Kimberly and Mueck, 1991; Krass and Kolletzki, 2001). Consequently, while hyperstereo HMD designs have been available for several decades, and several of these systems have been flight-tested, the high cost of flight tests has limited the study of long-term visual effects, especially the determination of an adaptation performance curve. This lack of data has prevented gaining a good understanding of whether the change in depth perception can be adapted to, or compensated for, with increasing exposure, which is critical to establishing sufficient training requirements of these systems.

However, in a joint flight study between Canada, Australia and the United States, conducted in August 2008, but not yet reported, pilot interviews following an average cumulative flight time of 9 hours using the Thales Aerospace TopOwl™ HMD, indicated that some level of adaptation to the hyperstereo effect may be achievable. With the exception of within 2-3 feet of the aircraft, the previously described "hole" effect seemed to no longer be experienced. This is a promising finding, but final analysis of the data has not been completed.

Identifying the mechanisms responsible for improved use of a hyperstereo system is important because the two possibilities have different implications and limitations. If users perceptually adapt to the changes in the visual image, then they would be able to operate with the system in virtually any situation. At the same time, switching from using the system to not (and vice-versa) may require some time for perceptual adaptation to take effect. Thus, transitions from normal vision to a hyperstereo system may be particularly important. Notably, if performance is based on perceptual adaptation, influence of memory, attention, and knowledge organization are unlikely to be a key part of these transitions.

In contrast to perceptual adaptation, if improved performance is due to cognitive compensation, then users have learned to change their behavior in response to perceptual experiences that they recognize as being different from normal. This strategy means that novel situations may not be dealt with properly because users have not learned the appropriate type of compensatory response. It might also be expected that with cognitive compensation a user can fairly easily transition between use and non-use of a hyperstereo HMD because all that changes are the schemas for operating with the system. Notably, if performance is based on cognitive compensation, transitions between normal and hyperstereo systems will be strongly dependent on the properties of memory, attention, and knowledge systems.

This case study emphasizes that an understanding of perceptual and cognitive systems are critical for judging the usability of systems such as HMDs that modify how humans interact with the outside world.

Case Study 2: The Integrated Helmet and Display Sighting System (IHADSS)

This second case study discusses some of the perceptual and cognitive issues associated with the Integrated Helmet and Display Sighting System (IHADSS) (Figure 15-15). This system, the U.S. Army's only fielded integrated HMD, is flown on the AH-64 Apache attack helicopter, and has been field-tested and used by over seven countries.

The IHADSS presents both pilotage visual imagery and aircraft flight symbology (e.g., airspeed, altitude, and heading) to the pilot. Pilotage imagery originates from a nose-mounted forward-looking infrared (FLIR) sensor known as the Pilot's Night Vision System (PNVS). This sensor is located approximately 9 feet (3 meters) forward of and 3 feet (1 meter) below the pilot's eye position.

The IHADSS consists of a miniature, 1-inch diameter, cathode-ray-tube (CRT) and an optical relay assembly, the Helmet Display Unit (HDU) (Fig. 15-16). The electronic image of the external scene is captured by the FLIR sensor and through a series of processes is presented as a luminance pattern image on the face of the CRT. This image is relayed optically through the HDU and reflected off a beamsplitter, also known as a combiner, into the pilot's eye (Rash and Verona, 1992). (See Chapter 3, *Introduction to Helmet-Mounted Displays*, for a more complete description of the IHADSS).

The pilotage imagery is presented monocularly (right eye only). The pilot's unaided (left) eye is available for viewing cockpit panel-mounted displays, reading maps, and observing lights, flares, and enemy fire outside the cockpit. This situation of presenting separate images to each of the two eyes is referred to as dichoptic viewing, a condition considered in the early design phases of the IHADSS as a potential source of visual problems.

Figure 15-15. The Integrated Helmet and Display Sighting System (IHADSS).

The HMD is designed so the image of the 30° vertical by 40° horizontal field-of-view (FOV) of the FLIR sensor subtends an identical 30° x 40° FOV at the pilot's eye. This provides unity magnification, which is necessary for piloting the aircraft. At nighttime, the pilot flies the aircraft using predominately the sensor imagery presented exclusively to the right eye via the HDU.

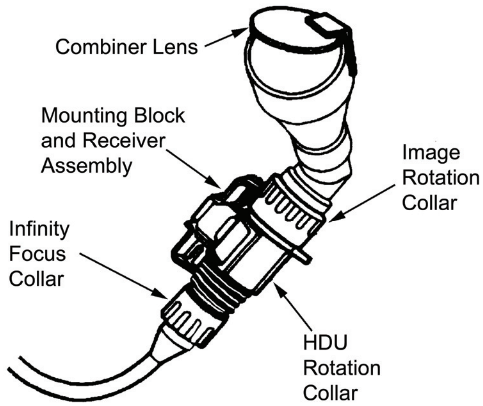


Figure 15-16. The Helmet Display Unit (HDU).

The AH-64 attack helicopter with its FLIR sensor and IHADSS HMD is a very challenging aircraft to fly. The pilot is expected to control and fly this tremendously sophisticated piece of machinery and perform combat missions using a reduced FOV picture of the outside world that is presented with visual cues from a completely different spectral range. The human visual system is designed to process information from natural visual scenes, but the IHADSS FLIR-based imagery is far from natural. In addition, the quality of the HMD imagery often is severely degraded in both contrast and resolution. As a result, this imagery has great potential to be misperceived (or misinterpreted).

In this section we consider several perceptual and cognitive issues that either are, or were expected to be, problematic for the use of the IHADSS. As part of this analysis, we try to highlight those perceptual and cognitive components that are likely to play a significant role in the use of the IHADSS. While this discussion focuses on real and potential problems with the IHADSS, this is not intended to undermine the many positive characteristics of the system. The IHADSS FLIR-based imagery provides information and opportunities for pilots to perform missions in situations that would otherwise be unmanageable, and as such is an enabler of missions, albeit with obviously increased risk compared to more normal conditions.

Depth perception

The IHADSS degrades a variety of visual cues for spatial depth. The most obvious missing cue is the loss of binocular stereopsis. The scene on the IHADSS display typically does not correspond to the scene for the unaided eye. As a result, the different views available to the two eyes do not contain the disparity cues that are normally used to judge relative distances of objects. The disparity cues that support stereopsis are most effective up to about 30 meters (100 feet) (Cutting and Vishton, 1995); a range that is quite important for tactical helicopter flight.

Fortunately, monocular cues to depth such as retinal size, occlusion, motion parallax, and perspective, generally provide cues to relative depth that compensate for the absence of the normal binocular disparity cues. This is easily verified by noting that the visual world does not appear flat when you close one eye. However, other aspects of the IHADSS can degrade some of these monocular cues as well.

The monocular cue of retinal size can be used to judge relative depth if the object is recognized and can be compared to its reference size in memory. The reduced resolution of the FLIR sensor/IHADSS display means that the visible scene lacks the crispness and detail of normal vision and it may be difficult to use the retinal size cue in some situations. Occlusion cues refer to the fact that parts of a closer object can cover and conceal parts of a farther object. Such cues will still be present in an IHADSS image, but their effectiveness may be reduced by the limited FOV. In general, a larger view of the scene will provide more information about the relative depths of objects in the scene. Perspective cues (which are based on changes in light as it passes through the atmosphere) may be entirely absent in FLIR imagery. Motion parallax refers to the fact that nearby objects will appear to move faster than farther away objects. In normal viewing, individuals often move their heads to produce these motion cues and thereby judge relative depth. Such efforts may be difficult in the IHADSS because there is a lag between the user’s head movement and the system’s updated imagery. Overall, AH-64 pilots have reported a reduction in monocular cues; most likely due to the reduced resolution of the FLIR sensor/IHADSS display (Crowley, 1991; Hale and Piccione, 1989).

During development of the IHADSS, there was substantial concern that the IHADSS’ monocular design would produce a depth-related phenomenon known as the Pulfrich effect. The Pulfrich effect occurs when both eyes view the same scene but one eye receives a higher level of light intensity than the other eye. This intensity difference leads to an interocular difference in the time needed for neural signals to reach those areas of the brain involved in depth perception. Thus, the intensity difference can lead to something similar to the motion parallax cues. The effect is that an object moving in a frontal plane appears to move out of the plane and approach toward or recede from the viewer. The difference in intensity could occur in two situations with the IHADSS. First, in nighttime viewing, the FLIR imagery may reveal the same objects and contours that are visible to the unaided eye, but at a higher intensity. Second, during daytime viewing the IHADSS monocle provides see-through capability, but is tinted to insure that symbology and other information is visible. This tinting means that the unaided eye views objects and contours with a higher intensity than the aided eye.

Despite these concerns, pilots have not reported experiences that would be consistent with the presence of the Pulfrich effect (Rash, in press). For nighttime viewing this can be explained by the fact that the two eyes rarely see the same scene. The unaided eye usually views the interior of the cockpit, while the aided eye views the FLIR imagery of the outside world. Thus, there is no opportunity for the information from the two eyes to combine and produce an illusory depth experience. It is less clear why the Pulfrich effect does not occur during daytime viewing, but perhaps the tint of the lens is not strong enough to produce the effect (Lit, 1949), or because most scenes have other depth cues that work against the effect.

One remaining influence on depth perception is the position of the PNVS sensor on the aircraft. Interpretations of unaided vision are tightly tied to the position of the eyes on the head. When flying the AH-64, the primary visual input for night and foul weather flight is the PNVS sensor. This sensor is located in a nose turret approximately 9 feet (3 meters) forward and 3 feet (1 meter) below the pilot's design eye position (Figure 15-17). Such positioning has the advantage of providing an unobstructed view of areas below the physical aircraft, which is definitely useful when landing in cluttered areas. On the other hand, this exocentric positioning of the sensor can introduce problems of parallax, motion estimation, and distance estimation (Hale and Piccione, 1989). Pilots also must learn how to manipulate the aircraft from a point of view that is different than their visual system is used to. Such learning faces issues similar to those involved in learning to fly with a hyperstereo system, as described above.

Binocular rivalry and attention switching

As discussed above, the monocular display format of the IHADSS means that the two eyes often view different scenes. If the two scenes are dramatically different and do not allow for an interpretation of a scene in depth, the percept tends to be related to only one scene, with the view in one eye suppressing the other (Bishop, 1981). This type of binocular rivalry was one of the biggest concerns about the development of the monocular IHADSS (Rash et al., 2008).

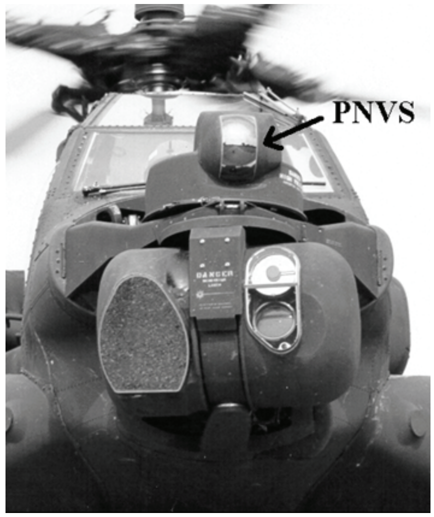


Figure 15-17. The position of the Pilot's Night Vision System (PNVS) and other imaging systems on the nose of the AH-64.

A variety of factors are known to influence binocular rivalry including brightness, timing, spatial detail, and color differences. For example, the relatively bright green phosphor in front of the right eye can make it difficult to attend to a darker visual scene in front of the left (unaided) eye. Conversely, if there are bright city lights in view, it may be difficult to shift attention away to the right (HDU) eye (Hale and Piccione, 1989). AH-64 pilots report occasional difficulty in adjusting to one dark-adapted eye and one light-adapted eye (Crowley, 1991). Most pilots have developed strategies to overcome rivalry effects (Rash, 2000).

While attentional focus can influence binocular rivalry (Chong et al., 2005) it is not the only important factor. In some situations, efforts to attend to items in one eye have only a slight effect on preventing a "flip" to the other eye (Meng and Tong, 2004). Such flips can be especially dangerous when a pilot tries to acquire information with one eye, but the scene from the other eye intrudes. For example, it may be hard to read instruments or maps inside the cockpit with the unaided eye, because the IHADSS eye "sees" through the instrument panel or floor of the aircraft, continuously presenting the pilot with a conflicting outside view. In addition, attending to the unaided eye may be difficult if the symbology presented to the right eye is changing or jittering (Crowley, 1991).

Moreover, attention switching between the eyes can be difficult, particularly as mission time progresses (Bennett and Hart, 1987). Some pilots resort to flying for short intervals with one eye closed, which is extremely fatiguing (Bennett and Hart, 1987; Hale and Piccione, 1989). User surveys indicate that the problems of binocular rivalry tend to ease with practice, but that rivalry is a recurrent pilot stressor, especially during a long, fatiguing mission and when there are other difficulties such as problems with display focus, flicker, or poor FLIR imagery (Bennett and Hart, 1987; Hale and Piccione, 1989).

More generally, these types of systems are, in principle, susceptible to the change blindness and cognitive tunneling phenomena discussed previously. It is difficult to attend to multiple scenes, and important information may be missed while attention is focused elsewhere.

Future designs of these kinds of systems may introduce even more opportunities for binocular rivalry. The design for the next generation U.S. Army helicopter calls for the integration of FLIR and I^2 sensor imagery. Due to the weight and size characteristics of FLIR technology, the FLIR's position will remain exocentric. However, the I^2 sensor(s) has two location options. It may be collocated with the FLIR sensor on the nose of the aircraft, or it may be helmet-mounted. If both sensors are exocentrically located, only the basic concerns of this mode of location, as listed above, require consideration. However, if the I^2 sensor is helmet-mounted, there may be problems associated with the mixed location modes and the resultant switching of visual reference points.

Head movement strategies

By virtue of their design, HMDs are mounted totally, or in part, on the user's helmet. In the IHADSS, the display section is helmet-mounted. The sensor section is nose-mounted on the aircraft and is integrated with the helmet in such a way that head movements control the direction of the sensor's line-of-sight. While head movements are a natural part of normal viewing, eye movements are also an important part of natural viewing, but eye movements are not captured and used with the IHADSS. Eye movements can be used to focus on particular parts of the IHADSS display, but, unlike normal vision, the external visual scene does not change in response to an eye movement.

Helmet-mounted imaging systems, such as the PNVS/IHADSS, use the pilot's head as a control device. Head position is employed to produce drive signals that slave the sensor's gimbaled platform to pilot head movements. Infrared detectors mounted on the helmet continuously monitor the head position of the pilot. Processing electronics of the IHADSS convert this information into drive signals for the PNVS gimbal. This type of control system is called a visually coupled system (VCS). It is a closed-loop servo-system that uses the natural visual and motor skills of the pilot to remotely control the pilotage and targeting sensors and/or weapons.

One important operating parameter of VCSs is the sensor's maximum slew rate. The inability of the sensor to slew at velocities equal to those present in unrestricted pilot head movements would result in 1) significant errors between where the pilot thinks he is looking and where the sensor actually is looking and 2) time lags between the head and sensor lines-of-sight. Medical studies of head movements have shown that normal adults can rotate their heads +/-90 degrees in azimuth (with neck participation) and -10 to + 25 degrees in elevation without neck participation. These same studies showed peak head velocity is a function of movement displacement, i.e., the greater the displacement, the greater the peak velocity, with an upper limit of 352 degrees/second (Alien and Webb, 1983; Zangemeister and Stark, 1981). However, these studies were laboratory-based and do not reflect the velocities and accelerations indicative of a helmeted head in military flight scenarios. In support of the AH-64 PNVS development, Verona et al. (1986) investigated single pilot head movements in an U.S. Army JUH-1M utility helicopter. In this study, head position data were collected during a simulated mission where four JUH-1M pilot subjects, fitted with a prototype IHADSS, were tasked with searching for a threat aircraft while flying a contour flight course (50 to 150 feet [15 to 46 meters] above ground level). The acquired head position data were used to construct frequency histograms of azimuth and elevation head velocities. Although velocities as high as 160 and 200 degrees/second in elevation and azimuth, respectively, were measured, approximately 97 percent of the velocities were found to fall between 0 to 120 degrees/second. This conclusion supported the PNVS design specification of a maximum slew rate of 120 degrees/second. It also lends validity to pilot complaints that the target acquisition and designation system sensor (with a maximum slew rate of 60 degrees/second) is too slow.

Even in the IHADSS, there are anecdotal reports that pilots complain they must slow down their head movements to effectively use the system. The problem seems to be related not to the slew rate of the PNVS, but to a lag for the system to detect changes in head acceleration (e.g., to start and stop a movement). Some lags are inevitable in a system, such as the IHADSS, where the FLIR sensor is physically separated from the head. In the IHADSS the VCS must continually calculate the user's head position, translate it into sensor motor commands, and route a command to the sensor gimbal; the gimbal must move the slew to the new position; finally, the display must be updated with a new image. Lags of this sort can produce a variety of deficits and image artifacts (Moffit, 1997; Kalawsky, 1993; Biocca, 1992). It appears that pilots have learned how to minimize these problems by restricting their head movements to fit within the limits of the system.

Head movement strategies are particularly important in the IHADSS (and most HMD systems) because the system provides a greatly restricted field of view (FOV) of a visual scene, as schematized in Figure 15-18. FOV in the IHADSS system is limited by two primary factors. The first factor is the weight of the helmet. A larger FOV invariably leads to larger optics and more weight. The second factor is the FOV of the sensor. For piloting an aircraft, the FOV of the sensor needs to be mapped with no change in magnification to the HMD display (else image quality quickly deteriorates).

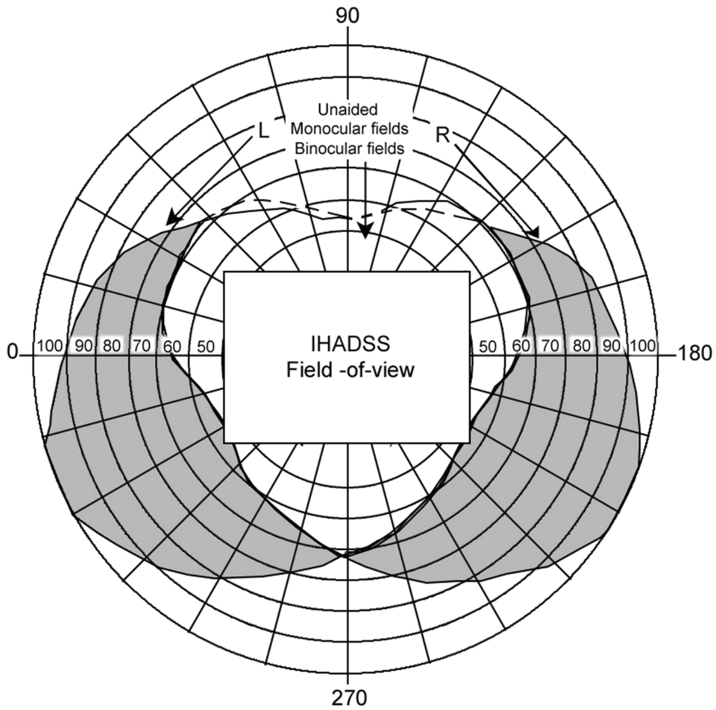


Figure 15-18. Pictorial representation of the IHADSS' 30- x 40-degree field-of-view as compared to that of the normal human field-of-view.

There are two aspects of the sensor's FOV. The first aspect is the amount of the visual field that can be covered in a single image. This is largely limited by the physical properties of the sensor. As shown in Figure 15-18, the IHADSS' 30 x 40 degree FOV appears small when compared to the FOV of the unaided eye. However, this reduced size is not so significant when one considers the multiple visual obstructions (i.e., armor, support struts, glare shield) that are normally present in military aircraft (Rash et al., 1990). With its external placement, the PNVS avoids many of these obstructions. The second aspect of FOV is the range of movement for the sensor. The IHADSS system provides an unimpeded external view throughout the range of the PNVS' movement (+/- 90 degrees in azimuth and +20 to -45 degrees in elevation).

As with pilots flying NVGs, AH-64 pilots are trained to use continuous scanning head movements to compensate for the limited FOV. Potentially disorienting effects occur when the pilot's head movements exceed the PNVS' range of movement. When this happens the head continues moving but the image remains unchanged. This situation could be misinterpreted by the pilot as a sudden aircraft pitch or yaw in the opposite direction of the head movement.

Not all of the effects of reduced FOV on pilot performance are fully understood. The task of determining a minimum FOV required to fly is not a simple one. The minimal FOV required is highly task-dependent. A high-speed flight across a desert floor with few obstacles can be accomplished with sensory cues that can be identified with a rather narrow FOV. On the other hand, performing a hovering turn in a confined area can only be accomplished with visual cues that need a wide FOV. Similarly, information is not processed equally across a display. Very fine visual details are only effectively processed for the parts of the display that fall on the fovea of the eye. Expanding the FOV to include more periphery information would not likely provide any benefit for some fine discrimination tasks, although eye movements across the display complicate matters substantially.

FOV problems in IHADSS are further complicated by the dual use of the IHADSS display. It is used to provide a view of the external world through the PNVS and also to provide flight symbology. Fight symbology information is placed on the edges of the CRT display, so as not to interfere with views of the external world. However, by being on the edge of the display, the symbols are difficult to resolve without an eye movement to place the symbol image on the fovea of the eye. Such eye movements require planning and attention that distract from other flight duties. To avoid this problem, some AH-64 pilots use the CRT horizontal and vertical size controls to reduce the overall size of the image (Hale and Piccione, 1989). This allows the pilot to view all of the imagery and symbology without difficult eye movements. Critically, though, the PNVS' FOV now occupies less area on the combiner and no longer maintains an accurate angular size of the scene. Since this minified image can cause problems with distance and size perception, it is strongly discouraged.

Interpreting sensor information

Normal vision has evolved to work with light energy over a specific part of the electromagnetic spectrum. As Figure 15-19 shows, the visible portion of this spectrum falls roughly between 400 and 700 nanometers (0.4 to 0.7 microns). Within this range, light behaves in certain ways as it reflects off of objects of different properties. Within the human visual system, different wavelengths of light that hit the eye are, to a first approximation, interpreted as different perceptual colors that identify properties of object surfaces. The visual system makes several assumptions about the properties of light and how it interacts with objects. For example, there is a bias to interpret illumination of a scene as coming from above (Ramachandran, 1988), which can have a strong effect on interpretations of cast shadows, relative depth perception, and figure-ground distinctions. Chapter 2, *The Human-Machine Interface Challenge*, discusses some other assumptions of perceptual and cognitive systems.

The images on the IHADSS that are generated by the PNVS do not necessarily obey the assumptions of the visual system. As Figure 15-19 shows, the PNVS thermal sensor captures electromagnetic energy from the infrared region with wavelengths between 8000 to 12000 nanometers. It is this ability to create images from long wavelength sources (heat energy) that allows the PNVS to provide nighttime vision. All physical objects emit some infrared energy. The PNVS sensors can detect emitted energy (of the right wavelength) from objects that are at temperatures of approximately -35 C° or higher. The IHADSS display shows a "heat map" of a visible scene.

Figure 15-20 shows a scene with a photograph taken by a normal (visible light) camera on the left and with an infrared camera on the right. There are many similarities between the two images. In each image, many of the major buildings are detected, and contrast between adjacent objects is notable. On the other hand, there are many significant differences in the two images. For example, the electrical wires visible in the normal image are almost invisible in the infrared image. Similarly, the writing on the railcars is visible in the normal image but washed out in the infrared image. The emissions from the buildings are nearly invisible in the normal image but are quite clear in the infrared image. Most of these differences are due to the properties of the sensors. They detect different types of electromagnetic energy and so are sensitive to different parts of the scene.

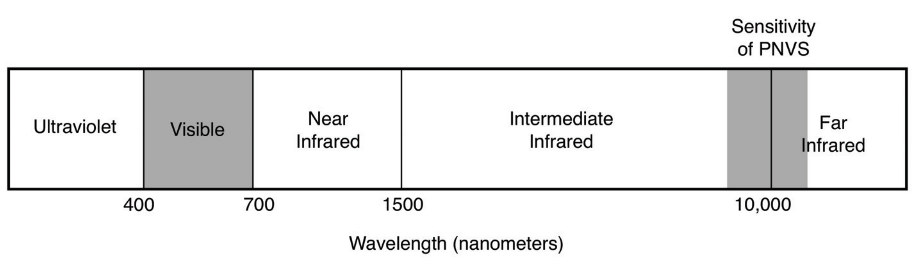


Figure 15-19. The electromagnetic spectrum. The gray areas indicate the wavelengths for normal vision and for the PNVS sensors.

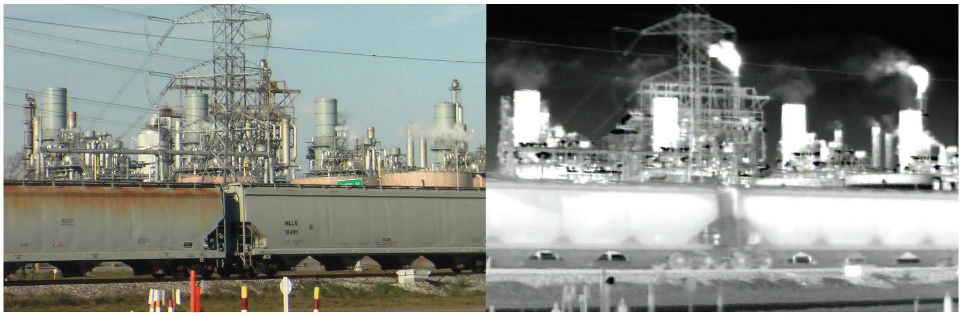


Figure 15-20. Pictures of a scene taken with a normal camera (left) and with an infrared camera (right).

One major difference from normal viewing of a scene is that the display is monochromatic in most devices. A normal image displays a variety of wavelengths of light that humans interpret as different colors. When printed on a black and white printer, both of the images in Figure 15-20 fail to show these colors. In the original (color) photo of the scene, wavelength differences allow a viewer to identify that the railcar on the left is rusty while the railcar on the right is not. Neither of the monochromatic images captures this difference and cannot distinguish the color differences of the railcars.

Using the PNVS requires learning how different surface properties correspond to different heat intensities. These relationships differ from those for light in the visible spectrum. Many surfaces that appear black in the visible spectrum may be very bright in the infrared spectrum. Likewise, surfaces that are bright in the visible spectrum may be relatively dark in the infrared spectrum (e.g., the sky in Figure 15-20).

The visual system's assumptions also can cause misunderstandings of an infrared image. For example, without the normal photo as a reference, the infrared image in Figure 15-20 does not seem to contain railcars at all. Instead, the cooler spaces between the wheels of the railcars appear to be some kind of bump (perhaps tents) in front of a wall. The railroad tracks appear to be another wall, perhaps in front of a body of water.

One challenge to learning how to interpret an infrared image is that the reflectance properties of surfaces depend on thermal properties that have unexpected consequences on the imagery. For example, twice a day (generally at midmorning and late afternoon) there is a thermal crossover, where temperature conditions are such that there is a near total loss of contrast between two adjacent objects in the infrared imagery. During this thermal crossover, the polarity of contrast reverses. In early morning the background temperature may be greater than a target's temperature, and the target will have a lower intensity on the infrared image. After thermal cross over, the target's temperature may be greater than the background and will have a higher intensity on the infrared image. Further, these attributes may change with the exposure history of the viewed objects, since objects will absorb more or less heat during the day depending on meteorological conditions. These are not changes that are easily interpreted by the human visual system, so they must be learned through cognitive strategies.

AH-64 accident rate

Immediately following the initial fielding of the AH-64A (and the IHADSS), numerous anecdotal reports of various physical and psychological/sensory problems surfaced. Hale and Piccione (1988) conducted the first user survey of AH-64 pilots and found evidence of increased pilot fatigue and, predominate among other complaints, headaches. They cited as possible causes almost all of the IHADSS-related factors discussed above plus additional hardware design issues (e.g., inadequate eye relief) and overall discomfort. Over the next 25+ years of fielding, numerous user surveys have documented consistent, but varying, rates of fatigue and other symptoms generally attributed to the IHADSS HMD (Behar et al., 1990; Crowley, 1991; Rash et al., 2001).

Informally, there has been a long-standing question within the aviation community as to whether there may be a connection between AH-64 accidents and the use of the IHADSS HMD (in combination with the FLIR pilotage sensor). The investigation of such a possible role was the primary objective of a study by Rash et al. (2003). This study analyzed accident data obtained from the U.S. Army Risk Management Information System (RMIS) database that was created in 1972 and is maintained by the U.S. Army Combat Readiness Center (USACRC) (formerly the U.S. Army Safety Center), Fort Rucker, Alabama.

Each AH-64A/D accident between October 1985 and March 2002 was reviewed by a panel of vision scientists and pilots that assessed the role of the HMD and/or FLIR sensor in the accident. Out of the 98 accidents that used the IHADSS only 2 accidents were identified as having the IHADSS/FLIR as a major contributing factor to the accident (meaning that without this component, any other factors could have been overcome without mishap). Thus, one important conclusion from this study is that the IHADSS/FLIR is not a major factor in the vast majority of AH-64 Apache accidents. This finding suggests that despite the difficulties of using the IHADSS, pilots have adapted to the needs and limitations of the system to effectively fly their aircraft.

An additional 19 accidents were revealed to have the IHADSS/FLIR as a subsidiary component of the accident (meaning that other factors would have lead to an accident in any case, but the IHADSS/FLIR made the accident sequence more difficult to deal with or the outcome more severe).

Table 15-3 lists causal factors related to the IHADSS/FLIR that were involved in the 21 accidents where the system was a major or subsidiary component of the accident. Some accidents involved multiple factors, so the numbers do not add up to 21 accidents or 100%. The most frequent causal factor in all of the accidents studied was dynamic illusions (91%), with undetected drift being the most common type. As an example, in one accident, the aircraft was allowed to drift into a tree because the student pilot failed to adequately monitor instruments, and the instructor pilot misjudged the position of the aircraft in relation to the trees (height judgment, 24%).

The second most frequent causal factor was degraded visual cues (62%), which was distributed across multiple sub factors with poor FLIR sensor conditions (19%) and impaired depth perception (19%) being most common. This was exemplified in one accident where the crew was operating under poor FLIR sensor conditions (following 4 days of rain). While trying to maintain a hover, the poor FLIR sensor visual cues, in conjunction with a lack of depth perception, prevented the crew from detecting the presence of trees and aircraft drift. As a result, the main rotor blades made contact with the trees.

The presence and frequency of the causal factors in the AH-64 accidents studied are consistent with the findings of Crowley (1991) and Rash et al. (2001). Both studies listed pilot reported problems associated with dynamic illusions, particularly undetected drift. Hale and Piccione (1988) and Rash et al. (2001) also raised concerns about poor sensor performance.

The HMD accident study concluded that while the presence and use of the IHADSS HMD present a very unique situation in the AH-64 Apache cockpit, it does not seem to be a major contributor to accidents. However, the study did suggest that the use of the IHADSS HMD was one more factor that increases workload and requires increased crew coordination. The study also concluded that the inability of the legacy FLIR sensor performance to provide pilots with sufficient resolution had an impact on safety. This poor performance is greatly increased during and following periods of environmental conditions that render the FLIR sensor ineffectual. The resulting lack of image quality significantly increases visual workload.

Applying Knowledge about Cognition to HMD Designs

The field of cognitive science has identified many aspects of cognition that are relevant to the development and use of HMDs. Indeed, HMDs provide such a rich variety of stimuli in challenging and important situations that they tap into properties of nearly every cognitive system. An understanding and appreciation of the properties of these cognitive systems will help focus designer and user expectations on what can and cannot be accomplished with HMDs.

Table 15-3.
Summary of accident factors (Rash et al., 2003).

Accident Factor	Number of accidents where factor was present or contributing	Totals (%) by accident factor
Display-related		**7 (33%)**
-Physiological causes	0	
-HDU impact on visual field	1	
-Alternation/rivalry	1	
-Degraded (insufficient) resolution	5	
Degraded visual cues		**13 (62%)**
-Poor FLIR conditions	2	
-Loss of visual contact with ground	2	
-Impaired depth perception	4	
-Limited FLIR sensor FOV	1	
-Inadvertent IMC	2	
Static illusions		**5 (24%)**
-Faulty height judgment	5	
-Trouble with lights	0	
Dynamic illusions		**19 (91%)**
-Undetected drift	11	
-Illusionary drift	0	
-Faulty closure judgment	5	
-Disorientation (vertigo)	3	
Hardware-related problems		**10 (48%)**
-FLIR sensor failure	5	
-IHADSS display/HDU failure	0	
-Design limitation	5	
Crew coordination related to IHADSS/FLIR sensor	12	**12 (57%)**

The problem with guidelines

It is common at the end of a chapter such as this one to provide a list of guidelines for designers to follow as they build and test HMDs (National Research Council, 1995, 1997; Patterson et al, 2006; Wickens et al., 1998). We are resisting the urge to create this kind of list because we do not feel that such guidelines are actually very useful in their current forms. Instead, we want to look at why these kinds of guidelines are not particularly useful and identify how information from cognitive science could be used in a different way.

To illustrate some of the problems with guidelines, consider a commonly cited guideline (e.g., Holley and Busbridge, 1995; Svensson et al. 1997):

- *Avoid overtaxing the user's short term memory capacity. Chunk items together so that a user does not have to remember more than seven items.*

The reference to seven items refers to a classic cognitive psychology paper by Miller (1956) that reported that individuals could remember (on average) about seven items for immediate recall (also see Figure 15-4). Longer lists of items produced some forgetting. The guideline is certainly correct that some environments can overtax a user's memory; however, there are several problems with this kind of guideline.

1. The implied statement in the second sentence is out of date. Miller's paper was a breakthrough finding at the time, but in the intervening 70 years the seven-item limit has been shown to be wrong. The number of items that can be kept for immediate recall varies dramatically. For items that sound very similar it is often much less than seven items (Conrad and Hull, 1964). With substantial amounts of training, some individuals can learn to use long term memory for immediate recall, and can recall lists of nearly 100 items (Chase and Ericsson, 1982).
2. Even if the second part of the statement were true, it is not clear how to satisfy the guideline. In particular, what counts as an *item*? Is a word a single item, or is it made of items defined by letters, syllables, or phonemes? Without a way to count the number of items it is impossible to determine if a task is over the user's limit.
3. The guideline does not make sense in isolation and cannot be treated as absolute. There may be some tasks where overtaxing the user's memory is not a problem. In particular, if information is displayed visually, a user can simply refer back to the display to examine information that might be forgotten. There may be some contexts in which this guideline is important, but the guideline itself cannot (and does not) indicate when it is important.
4. Satisfying this particular guideline can introduce a display that violates other guidelines. One way to avoid overtaxing short term memory might be to keep all necessary information visible on a display. But then the user must select which information should be displayed and must search the screen for the particular information needed. These new requirements probably violate other guidelines. It is not at all clear which guideline should dominate when several guidelines conflict with each other.

The second and fourth criticisms apply to most guidelines. Human cognition is sufficiently complex that it is very difficult to predict what a person will do in any specific circumstance. By their very structure, guidelines can only give general suggestions about what a designer should consider.

As a second example, the National Research Council (1997) analyzed HMDs for the Land Warrior System. The end of almost every chapter includes design guidelines. The executive summary emphasized four guidelines (page 6):

1. Minimize the degree to which the display is a physical barrier to acquiring information about the environment.
2. Provide integrated information in a task-oriented sequence, minimizing extraneous information and memory requirements.
3. Use graphics that have been well learned by the soldier. Simplify the presentation of data entry and system control options.

At first glance, these all seem like reasonable guidelines for an HMD design, but careful thought reveals that they are not really guidelines but *goals*. A guideline should indicate how to build a system, but these "guidelines" do not generally do that. For example, the second guideline gives no indication of how to minimize memory requirements. Many documents with guidelines do try to distinguish goals from guidelines (e.g., Toms and Williamson, 1998), but because it is not clear how to satisfy the guidelines, they are goals for all practical purposes.

This vagueness is a general problem with guidelines of this sort. On the one hand, many guidelines on HMD design simply identify desired properties of the system with little indication on how to achieve (or even measure) such properties. On the other hand, guidelines that do give specific advice (such as to avoid lists of more than seven items) are not generally true nor are broadly applicable across all situations.

There *are* some properties of cognition that are more universal, but the nature of these properties does not usually help guide system designs. For example, Tulving and Thompson (1973) proposed an encoding specificity principle of memory that states that the ability to remember an item depends on the similarity between the way the item is processed when it is encoded and the way it is processed when it is tested. There is substantial evidence that this statement is generally true for many different situations. However, this principle lacks sufficient detail to provide much guidance on how to design an HMD. In particular, one has to define how similarity is measured, but this term probably changes across individuals, tasks, and contexts.

We are not suggesting that the current use of guidelines is totally without merit. Although improperly named, guidelines do function as a set of goals for a design. Every design project needs goals of this type. Moreover, guidelines as they are currently used can push designers to consider issues that they might not have considered otherwise. Consider these two guidelines from Wickens et al. (1998):

1. Input modes, response devices, and tasks should be combined such that they are as dissimilar as possible in terms of processing stages, input modalities and processing codes.
2. The greater the automation of any particular task, the better the time-sharing capability. Information should be provided so that the person knows the importance of each task and therefore how to allocate resources between tasks.

These guidelines have many of the limitations and problems discussed above, but they do refer to specific topics in cognitive psychology that a designer might otherwise not consider. For example, the first guideline might motivate a designer to reconsider the system input modes and try to come up with a better approach. The guideline does not really indicate how this can be accomplished (or measured), but at least it does point to a potential need. In a similar way, a guideline that emphasizes limits to human memory may cause a designer to realize that users are struggling because of memory problems. In general, more thought given to the design process should lead to a better overall design.

Ultimately, a good design of a human-machine interface (HMI) system requires two things. First, the designer must be intimately aware of the needs and abilities of the user and must spend substantial time and effort to insure that the design satisfies those needs and takes advantage of those abilities. Second, the design must be tested, redesigned, and re-tested in a cycle that often repeats many times, taking as a criterion the final performance of the combined user/system in the test scenarios rather than simply the technical performance of the engineered system. Human factors (neuroergonomics) must be included at the beginning of a design process (e.g., Sheridan and Parasurman, 2006). Guidelines, in their current form, do not offer much meaningful guidance on how to accomplish these requirements.

An alternative to guidelines

Rather than providing guidelines that do not actually offer guidance, we propose that it would be better to simply list the main properties of cognition that are likely to be important for HMI design. For example, it *is* important to know that human working memory has a limited capacity and that a design that expects too much from working memory is going to be problematic. Note that such a statement does not suggest any guidance on how to solve the problem; that is the job of the designer. In many respects, such a list may not be much different from the guidelines that currently exist. Nevertheless, we think that such a list would indicate that these issues are starting points for HMI design rather than guidance. This is an important distinction.

A more fundamental break from guidelines is to explore areas where quantitative theories and models of cognition can predict human behavior. A quantitative model of, say, stimulus visibility, can make a precise statement about the visibility of a stimulus and how visibility may change for a variety of situations. Thus, rather than telling a designer to consider the influence that other items' colors may have on the visibility of a target stimulus, a quantitative model can predict the effect of color variations.

When a quantitative model exists that can predict an aspect of human behavior that is important for HMD design (visibility, usability, memory, etc.), then there is a standard way of utilizing that model to guide the design process. Namely, one can use standard optimization approaches (hill-climbing, genetic algorithms, etc.) to build a variety of designs that are shaped or measured relative to the model-predicted human behavior. Such an optimization approach can also easily include multiple models that may focus on different aspects of the HMD design. Thus, part of the design process can be simulated on a computer, which leads to a savings in time and money by reducing the need for empirical measurements. Such simulations also free the designer to consider a wider variety of designs because details of the designs can be tested more quickly.

This approach has been successfully applied to several situations including multifunction displays (Francis and Rash, 2002), keyboard designs (Francis and Oxtoby, 2006; Francis and Rash, 2005; Li et al., 2006), and computer menus (Liu et al., 2002). The main limit of this approach is that current models of cognition are unable to accurately predict human behavior in the situations that are relevant to many HMI design projects (Rogers, 2004). In some cases the models cannot be applied to real-world situations because they make assumptions that cannot be satisfied. In other cases a model is simply wrong. However, as part of a larger program of modeling, identification of model limitations and mistakes can be used to promote model development; something that is quite difficult for a set of guidelines. We anticipate that a vigorous use of models of cognition would lead to dramatic refinement of models that would improve their ability to predict human behavior.

Perhaps the clearest lesson for HMD designers to appreciate about cognition is that all cognitive functions are context and task dependent. Thus there are few simple solutions to the complex demands of HMD design because so much of HMD use depends on the complex capabilities and limitations of cognition.

References

Alien, J.H., and Webb, R.C. (1983). Helmet mounted display feasibility model. NAVTRAEQUIPCEN IH-338, Naval Training Equipment Center, Orlando, FL, 1983.

Anderson, J. (1993). *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Baddeley, A.D. (1986). *Working Memory*. New York: Clarendon Press/Oxford University Press.

Baddeley, A.D. (2003). Working Memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829-839.

Bedard, J., and Chi, M.T.H. (1992). Expertise. *Current Directions in Psychological Science*, 1, 135-139.

Behar, I., Wiley, R.W., Levine, R.R., Rash, C.E., Walsh, D.J., and Cornum, R.L.S. (1990). Visual survey of Apache aviators (VISAA). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 90-15.

Bennett, C.T., and Hart, S.G. (1987). PNVS-related problems: pilots reflections on visually coupled systems, 1976-1987. Working paper. Moffett Field, CA, NASA-Ames Research Center.

Biocca, F. (1992). Will simulation sickness slow down the diffusion of virtual environment technology? *Presence*, 1, 334-343.

Bishop, P.O. (1981). Binocular vision. In: Moses, R.A. (Ed.), *Adler's Physiology of the Eye*, 7th Ed. St. Louis: C.V. Mosby Company.

Boring, E.G. (1950). *A History of Experimental Psychology*. New York: Appleton Century-Crofts, Inc.

Bregman, A.L. (1981). Asking the "what for" question in auditory perception. In: Kubovy, M., and Pomerantz, I.R. (Eds.), *Perceptual Organization*. Hillsdale, NI: Erlbaum.

Brickner, M.S. (1989). Apparent limitations of head-up displays and thermal imaging. *Proceedings of the Fifth International Symposium on Aviation Psychology,* 703-707. Columbus, OH: Ohio State University.

Broadbent, D.E. (1958). *Perception and communication*. London: Pergamon.

Byrne, M.D., Kirlik, A., Fleetwood, M.D., Huss, D.G., Kosorukoff, A., Lin, R., and Fick, C.S. (2004). A closed-loop, ACT-R approach to modeling approach and landing with and without synthetic vision system (SVS) technology. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, 2111-2115). Santa Monica, CA: Human Factors and Ergonomics Society.

Card, S.K., Newell, A. and Moran, T.P. (1983). *The Psychology of Human-Computer Interaction*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Chase, W.G., and Ericsson, K.A. (1982). Skill and working memory. In: Bower, G.H. (Ed.), *The Psychology of Learning and Motivation* (Vol. 16). New York: Academic Press.

Cheung, K.M., and Milgram, P. (2000). Visual detection with hyperstereo video for aerial search and rescue. *Proceedings of the IEA 2000/HFES 2000 Congress* (Vol. 2), Santa Monica, CA: Human Factors and Ergonomics Society, 3, 472-475.

Chong, S.C., Tadin, D., and Blake, R. (2005). Endogenous attention prolongs dominance durations in binocular rivalry. *Journal of Vision*, 5(11):6, 1004-1012, http://journalofvision.org/5/11/6/, doi:10.1167/5.11.6.

Churchland, P.S., and Sejnowski, T.J. (1988). Perspectives on cognitive neuroscience. *Science*, 242, 741-745.

Conrad, R., and Hull, A.J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology,* 55, 429-432.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 1, 87-114.

Crawford, J., and Neal, A. (2006). A review of the perceptual and cognitive issues associated with the use of head-up displays in commercial aviation. *The International Journal of Aviation Psychology*, 16, 1-19.

Crowley, J.S. (1991). Human factors of night vision devices: Anecdotes from the field concerning visual illusions and other effects. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 91-15.

CuQlock-Knopp, V.G., Myles, K.P., Malkin, F.J., and Bender, E. (2001). The effects of viewpoint offsets of night vision goggles on human performance in simulated grenade-throwing task. Aberdeen, Maryland: U.S. Army Research Laboratory. ARL-TR-2407.

Cutting, J.E., and Vishton, P.M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In: Epstein, W., and Rogers, S. (Eds.), *Handbook of Perception and Cognition: Perception of Space and Motion*. New York: Academic Press.

Davis, G. (2004). Characteristics of attention and visual short-term memory: implications for visual interface design. *Philosophical Transactions, Series A, Mathematical, Physical, and Engineering Sciences*, 362, 2741-2759.

Duncan, J., Humphreys, G.W., and Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, **7**, 255-261.

Edgar, G.K. (2007). Accommodation, cognition, and virtual image displays: A review of the literature. *Displays*, 28, 45-49.

Endsley, M.R. (1999). Situation awareness in aviation systems. In: Garland, D.J., Wise, A.J., and Hopkin, V.D. (Eds.), *Handbook of Aviation of Human Factors*. Mahwah, NJ: Erlbaum Associates.

Fadden, S., Ververs, P.M., and Wickens, C.D. (1998). Costs and benefits of head-up display use: A meta-analytic approach. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 16-20). Santa Monica, CA: Human Factors and Ergonomics Society.

Finegold, L., and Flamm, B.L. (2006). Magnet therapy. *British Medical Journal*, 332, doi:10.1136/bmj.332.7532.4

Fischer, E., Haines, R.F., and Price, T.A. (1980). Cognitive issues in head-up displays. NASA Technical Paper 1711, NASA Ames Res. Ctr., Moffett Field, CA.

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381-391.

Foyle, D.C., Hooey, B.L., Byrne, M.D., Corker, K.M., Deutsch, S., Lebiere, C., Leiden, K. and Wickens, C.D. (2005). Human performance models of pilot behavior. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 1109-1113). Santa Monica, CA: Human Factors and Ergonomics Society.

Foyle, D.C., Sanford, B.D., and McCann, R.S. (1991). Attentional issues in superimposed flight symbology. In: Jensen, R.S. (Ed.), *Proceedings of the Sixth International Symposium on Aviation Psychology* (pp. 577-582). Columbus, OH: Ohio State University.

Francis, G., and Oxtoby, C. (2006). Building and testing optimized keyboards for specific text entry. *Human Factors*, 48, 279-287.

Francis, G., and Rash, C.E. (2002). MFDTool (Version 1.3): A software tool for optimizing hierarchical information on multifunction displays. Fort Rucker, AL. U.S. Army Aeromedical Research Laboratory. USAARL Report, No. 2002-22.

Francis, G., and Rash, C.E. (2005). Analysis and design of keyboards for the AH-64D helicopter. Fort Rucker, Alabama: US Army Aeromedical Research Laboratory. *USAARL Report* No. 2005-11.

Gazzaniga, M.S. Ivry, R., and Mangun, G.R. (1998). *Fundamentals of Cognitive Neuroscience*, W.W. Norton.

German Air Force Test Center (WTD). (1998). Vorbereitung und HIS Prufhunsept, Teilabshift 3 der Flugversuche. WTD 61, Manchung.

Goldstein, E.B. (2002). *Sensation and Perception, 6th Edition*. Pacific Grove, CA : Thomson Wadsworth.

Goldstein, E.B. (2005). *Cognitive Psychology: Connecting Mind, Research, and Everyday Experience*. Belmont, CA : Thomson Wadsworth.

Gomes, L.M.P., Martinho Pimenta, A.J.F., and Castelo Branco, N.A.A. (1999). Effects of occupational exposure to low frequency noise on cognition, *Aviation, Space, and Environmental Medicine*, 70, A155-A118.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review*, 104, 618-658.

Guiarda, Y., and Beaudouin-Lafon, M. (2004). Fitts' law 50 years later: Applications and contributions from human–computer interaction, *International Journal of Human-Computer Studies*, 61, 747-750.

Hale, S., and Piccione, D. (1989). Pilot performance assessment of the AH-64 helmet display unit. Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory. Technical Note 1-90.

Harris, D., and Muir, C. (2006). *Contemporary Issues in Human Factors and Aviation Safety*. Burlington, VT: Ashgate.

Heinecke, J.K., Ranaudo, R.J., Rash, C.E. and Hiatt, K.L. (2007). Dual-sensor use in the AH-64 crew station for urban combat in operation Iraqi freedom. *Proceedings of the American Helicopter Society 63rd Annual Forum.* On CD-ROM.

Helleberg, J.R., and Wickens, C.D. (2003). Effects of data-link modality and display redundancy on pilot performance: An attentional perspective, *The International Journal of Aviation Psychology*, 13, 189–210.

Hinkle, D., Wiersma, W., and Jurs, S. (2003). *Applied Statistics for the Behavioral Sciences, 5th Ed*. Boston: Houghton Mifflin Company.

Hoisko, J. (2003). Early Experiences of Visual Memory Prosthesis for Supporting Episodic Memory. *International Journal of Human-Computer Interaction,* 15, 209–230.

Hollands, J.G., Parker, H.A., McFadden, S., and Boothby, R. (2002). LCD versus CRT displays: A comparison of visual search performance for colored symbols. *Human Factors*, 44, 210-221.

Holley, C., and Busbridge, M. (1995). Evolution of the Venom variant of the AH-1 W Supercockpit. *Proceedings of the American Heliconter Society 51st Annual Forum*. 1436-1449.

Itti, L, and Koch, C. (2001). Computational modeling of visual attention, *Nature Reviews Neuroscience*, 2,194-203.

Itti, L., Koch, C., and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254-1259.

Johnson, C.W., and Dell, W. (2003). Limitations of 3D audio to improve auditory cues in aircraft cockpits. In: Einarsson, G., and Fletcher, B. (Eds.), *International Systems Safety Conference*. Unionville, VA: International Systems Safety Society.

Judge, S.J., and Bradford, C.M. (1988). Adaptation to telestereoscopic viewing measured by one-handed ball catching performance. *Perception*, 7, 783-802.

Kahneman, D., and Tversky, A. (1982). On the study of statistical intuitions. *Cognition*, 11, 123-141.

Kahneman, D. (1973). *Attention and effort*. New Jersey: Prentice Hall.

Kalawasky, R.S. (1993). *The science of virtual reality and virtual environments*. Wokingham, England: Addision-Welsey.

Kalich, M.E., Rash, C.E., McLean, W.E., and Ramiccio, J.G. (2007). A limited flight study for investigating hyperstereo vision. *Proceedings of SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications,* 6557(0I), 1-14.

Kanizsa, G. (1979). *Organization of vision*. New York: Praeger.

Kieras, D., and Meyer, D.E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12, 391-438.

Kimberly, J., and Mueck, S. (1991). *Integrated helmet display system (INVIS) flight assessment*. Fort Belvoir, VA: Airborne Electronics Research Detachment. Report No. NV-1-92.

Krass, and Kolletzki, D. (2001). Erfahrungsbericht der Fluge mit HMD/D der Fa. Sexant bei GenHFig/GWE. Buckburg.

Lebiere, C., Biefeld, E. Archer, R., Archer, S., Allender, L., and Kelley, T.D. (2002). Imprint/ACT-R: Integration of a task network modeling architecture with a cognitive architecture and its application to human error modeling. In M. Chinni (Ed.) *Military, Government and Aerospace Simulation* (pp. 13-19). San Diego: Society for Modeling and Simulation International.

Lederman, S.J., and Klatzky, R.L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19, 342-368.

Leger, A., Roumes, C., Bergeaud, J.M., Dareoux, P. and Gardelle, C. (1998). Flight testing of a binocular biosensor HMD for helicopter: Some human factors aspects. *Proceedings of SPIE, Helmet- and Head-Mounted Displays III*, 3362, 136-143.

Li, Y., Chen, L., and Goonetilleke, R.S. (2006). A heuristic-based approach to optimize keyboard design for single-finger keying applications. *International Journal of Industrial Ergonomics*, 36, 695-704.

Lieberman, H.R., Bathalon, G.P., Falco, C.M., Morgan, C.A., Niro, P.J., and Tharion, W.J. (2005). The fog of war: Decrements in cognitive performance and mood associated with combat-like stress. *Aviation, Space, and Environmental Medicine*, 76, C7-14.

Lindin, D.E., Kallenbach, U., Heinecke, A., Singer, W., and Goebel, R. (1999). The myth of upright vision: A psychological and functional study of adaptation to inverting spectacles. *Perception*, 28, 469-481.

Lit, A. (1949). The magnitude of the Pulfrich stereophenomenon as a function of binocular differences of intensity at various levels of illumination. *American Journal of Psychology*, 62, 159-181.

Liu, B., Francis, G., and Salvendy, G. (2002). Applying models of visual search to menu design. *International Journal of Human-Computer Studies*, 56, 307–330.

Logan, G.D. (1988). Automaticity, resources and memory: Theoretical controversies and practical implications. *Human Factors*, 30, 583-598.

Lovett, M.C. (2002). Problem solving. In: Medin, D.L. (Ed.), *Steven's Handbook of Experimental Psychology* (3rd Ed., pp. 317-362). New York: Wiley.

Manning, S.D., Rash, C.E., LeDuc, P.A., Noback. R.K., and McKeon, J. (2004). *The role of human causal factors in U.S. Army unmanned aerial vehicle accidents*. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2004-11.

Martin, R.L., Watson, D.B., Smith, S.E., McAnally, K.I. and Emonson, D.L. (2000). Effect of normobaric hypoxia on sound localization. *Aviation, Space, and Environmental Medicine*, 71, 991-995.

McCann, R.S., Foyle, D.C., and Johnston, J.C. (1993). Attentional limitations with head-up displays. In: Jensen, R.S. (Ed.), *Proceedings of the Seventh International Symposium on Aviation Psychology*. Columbus, OH: Ohio State University.

McLean, W.E., Rash, C.E., McEntire, B.J., Braithwaite, M.G., and Mora, J.C. (1998). A performance history of An/PVS-5 and ANVIS image intensification systems in U.S. Army aviation. *Proceedings of SPIE, Head-Mounted Displays I*, 3058, 264-298.

Meng, M., and Tong, F. (2004). Can attention selectively bias bistable perception? Difference between binocular rivalry and ambiguous figures. Retrieved 15 May 2007 from: *Journal of Vision*, 4(7), 539–551, http://www.journalofvision.org / 4/7/2/, doi:10.1167/4.7.2

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.

Moffitt, K. (1997). Designing HMDs for viewing comfort. *Head mounted displays: Designing for the user*. New York: McGraw-Hill.

National Research Council. (1995). *Human Factors in the Design of Tactical Display Systems for the Individual Soldier*. Washington, D.C: National Academy Press.

National Research Council. (1997). *Tactical Display Systems for Soldiers*. Washington, D.C: National Academy Press.

Navon, D., and Gopher, D. (1979). On the economy of the human processing systems. *Psychological Review*, 86, 254-255.

Neath, I., and Surprenant, A. M. (2003). *Human memory: An introduction to research, data, and theory* (2nd Ed.). Belmont, CA: Wadsworth.

Nelson, W.T., and Bolia, R.S. (2005). Battle management command and control (BMC2) human machine interface (HMI) design guide. Wright-Patterson AFB, OH: Air Force Research Laboratory. AFRL-HE-WP-TR-2005-0151.

Newman, R.L., and Greely, K.W. (1997) Integrating head-mounted displays into a cockpit. *Proceedings of SPIE, Head-Mounted Displays II*, 3058, 46-56.

Norman, D. (2002). *The Design of Everyday Things*. New York: Basic Books (Perseus).

Pashler, H.E. (1997). *The Psychology of Attention*. Cambridge, MA: MIT Press.

Patterson, R., Winterbottom, M.D., and Pierce, B.J. (2006). Perceptual issues in the use of head-mounted visual displays. *Human Factors*, 48, 555-573.

Posner, M.I. (1978). *Chronometric Explorations of Mind*. Hillsdale: Erlbaum.

Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.

Raizada, R., and Grossberg, S. (2003). Towards a theory of the laminar architecture of cerebral cortex: Computational clues from the visual system. *Cerebral Cortex*, 13, 100-113

Ramachandran, V.S. (1988). Perception of shape from shading. *Nature*, 331, 163-166.

Rash, C.E. (2008). A 25-year retrospective review of visual complaints and illusions associated with a monocular helmet-mounted display, *Displays*, 29(2), 70-80.

Rash, C.E. (Ed.) (2000). *Helmet Mounted Displays: Design Issues for Rotary-Wing Aircraft*. Bellingham, WA: SPIE Press.

Rash, C.E., McLean, W.E., Mora, J.C., Ledford, M.H., Mozo, B.T., Licina, J.R., and McEntire, B.J. (1998). Design issues for helmet-mounted display systems for rotary-wing aviation. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 98-32.

Rash, C.E., Reynolds, B.S, Stelle, J.A, Peterson, R.D., and LeDuc, P.A. (2003). The role of the Pilot's Night Vision System (PNVS) and Integrated Helmet and Display System (IHADSS) in AH-64 accidents. Fort Rucker, AL: U.S. Army Aeromedical research Laboratory. USAARL Report No. 2003-08.

Rash, C.E., Suggs, C.L., Mora, J.C, Adam, G.E., van de Pol, C., and Reynolds, B.S. (2001). Visual issues survey of AH-64 Apache aviators (Year 2000). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2002-02.

Rash, C.E., van de Pol, C., Harris, E.S., McGilberry, W.H., and Crowley, J.S., King, R.P., Braithwaite, M.G. (in press). The Effect of a Monocular Helmet-Mounted Display on Aircrew Health: A Cohort Study of Apache AH Mk 1 Pilots, Four-Year Review. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Rash, C.E., and Verona, R.W. (1992). The human factor considerations of image intensification and thermal imaging systems. In: Karim, M. (Ed.), *Electro-Optical Displays.* New York: Marcel Dekker, Inc.

Rash, C.E., Verona, R.W., and Crowley, J.S. (1990). Human Factors and Safety Considerations of Night Vision Systems Using Thermal Imaging Systems. *Proceedings of SPIE*, Helmet-Mounted Displays II, 1290, 142-164.

Reason, J. (1990). *Human Error*. Cambridge, UK: Cambridge University Press.

Reed, S.K. (2004). *Cognition: Theory and Applications, 6th Edition*. Belmont, CA: Thomson Wadsworth.

Rensink R.A., O'Regan J.K., and Clark J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368-373.

Rogers, Y. (2004). New theoretical approaches for Human-Computer Interaction. *Annual Review of Information, Science and Technology*, 38, 87-143.

Rosch, E.H. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192-233.

Rosch, E.H., Mervis, C.B., Gray, W.D., Johnson, D.M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.

Rouse, W. B., Edwards, S. L. and Hammer, J. M. (1993). Modeling the dynamics of mental workload and human performance in complex systems. *IEEE Transactions on System, Man, and Cybernetics*, 23, 1662-1671.

Russo, M.B., Kendall, A.P., Johnson, D.E., Sing, H.C., Thorne, D.R., Escolas, S.M., Santiago, S., Holland, D.A., Hall, S.W. and Redmond, D.P. (2005). Visual perception, psychomotor performance, and complex motor performance during an overnight air refueling simulated flight. *Aviation, Space, and Environmental Medicine*, 76, C92-103.

Schneider, B. and Moraglia, G. (1994). Binocular vision enhances target detection by filtering the background. *Perception*, 23, 1267-1286.

Schultz, E.E., Nichols, D.A., and Curran, P.S. (1985). Decluttering methods for high density computer-generated graphic displays. In *Proceedings of the Human Factors and Ergonomics Society 29th Annual Meeting*, 300-303. Santa Monica, CA: Human Factors and Ergonomics Society.

Scribner, D.R., Wiley, P.H., Harper, W.H., and Kelley, T.D. (2007). *The Effects of Workload Presented via Visual and Auditory Displays on Soldier Shooting and Secondary Task Performance*. Aberdeen Proving Ground, MD: Army Research Laboratory. ARL-TR-4224

Shappell, S.A., and Wiegmann, D.A. (2000). *Human factors analysis and classification system-HFACS*. Department of Transportation, Federal Administration: Washington, DC. DOT/FAA/AM-00/7.

Shelden, S.G., Foyle, D.C., and McCann, R.S. (1997). Effects of scene-linked symbology on flight performance. *Proceedings of the 41st Annual Meeting of the Human Factors and Ergonomics Society* (pp. 294-298). Santa Monica, CA: Human Factors and Ergonomics Society.

Shepard, R.N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703.

Sheridan, T., and Parasuraman, R. (2006). Human-automation interaction. *Reviews of Human Factors and Ergonomics*, 1, 89-129

Simons, D.J., and Levin, D.T. (1997). Change blindness. *Trends in Cognitive Science*, 1, 261-267.

Smith, E.E., and Kosslyn, S.M. (2007). *Cognitive Psychology: Mind and Brain*. New York: Pearson Prentice Hall.

Smith, S.L., and Mosier, J.N. (1986). *Guidelines for designing user interface software*. ESD-TR-86-278, United States Air Force, Hanscom Air Force Base. Available on line at http://www.hcibib.org/sam/index.html

St. John, M., Smallman, H.S., and Manes, D.I. (in press). Interruption recovery tools for team collaboration. *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

St. John, M., Smallman, H.S., Manes, D.I., Feher, B.A., and Morrison, J.G. (2005). Heuristic automation for decluttering tactical displays. *Human Factors*, 47, 509-525.

Svensson, E., Angelborg-Thanderz, M., Sjöberg, L., and Olsson, S. (1997). Information complexity-mental workload and performance in combat aircraft. *Ergonomics*, 40, 362-380.

Tanaka, J.W., and Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23, 457-482.

Thomas, M., Sing, H., Belenky, G., Holcomb, H., Mayberg, H., Dannals, R., Wagner, Jr., H., Thorne, D., Popp, K., Rowland, L., Welsh, A., Balwinski, S., and Redmond, D. (2000). Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain. *Journal of Sleep Research*, 9, 4, 335-52.

Thomas, M., Sing, H., Belenky, G., Holcomb, H., Mayberg, H., Dannals, R., Wagner, Jr., H., Thorne, D., Popp, K., Rowland, L., Welsh, A., Balwinski, S., and Redmond, D. (2003). Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain. *Thalamus and Related Systems*, 2, 3, 199-229.

Treisman, A.M., and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.

Trick, L.M., and Pylyshyn, Z.W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 331-351.

Triesch, J., Ballard, D., Hayhoe, M., and Sullivan, B. (2003). What you see is what you need. *Journal of Vision*, 3, 86-94.

Tulving, E., and Thompson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 359-380.

Toms, M., and Williamson, J. (1998). *Aviation human-computer interface (AHCI) style guide, Version 2.2*. Joint Cockpit Office: U.S. Army Aviation Research Development and Engineering Center.

Verona, R.W., Rash, C.E., Holt, W.E., and Crosley, J.K. (1986). Head movements during contour flight. Fort Rucker, AL: .U.S. Army Aeromedical Research Laboratory. USAARL Report No. 87-1.

Walker, M.P., and Stickgold, R. (2005). Memory consolidation and reconsolidation: What is the role of sleep? *Trends in Neuroscience*, 28, 408-415.

Walker, M.P., Stickgold, R., Alsop, D., Gaab, N., and Schlaug, G. (2005). Sleep-dependent motor memory plasticity in the human brain. *Neuroscience*, 133, 911-917.

Watkins, W.R. (1997). Enhanced depth perception using hyperstereo vision. *Proceedings of SPIE, Targets and backgrounds: Characterization and Representation II.*, 3062, 117-125.

Welch, R.B. (1986). Adaptation to space perception. In: Boff, K.R., and Thomas, J.P. (Eds.), *Handbook of Perception and Human Performance*, 24.1, 24-45. New York: John Wiley and Sons.

Wertheimer, M. (1923). Untersuchen zur Lehre von der Gestalt, II. *Psychologishe Forshung*, 4, 301-305.

Wickens, C.D. (1980). The structure of attentional resources. In R.S. Nickerson (Ed.). *Attention and performance, VIII*. Hillsdale, NJ: Erlbaum.

Wickens, C.D. (1984). Processing resources in attention. In: Parasuraman, R., and Davies, D.R. (Eds.). *Varieties of Attention*. (pp. 63-102). London: Academic Press.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance*. New York: Harper-Collins Publishers.

Wickens, C.D., Gordon, S.E., and Liu, Y. (1988). *An Introduction to Human Factors Engineering*. New York: Longman

Wickens, C.D., and Long, J. (1995). Object versus space-based models of visual attention: Implications for the design of head-up displays. *Journal of Experimental Psychology: Applied*. 1, 179-193.

Wickens, C.D., McCarley, J.S., Alexander, A.L., Thomas, L.C., Ambinder, M. and Yheng, S. (2005). *Attention-Situation Awareness (A-SA) Model of Pilot Error*. Technical Report AHFD-04-15/NASA-04-5, University of Illinois.

Wildzunas, R.M. (1997a). They shut down the *wrong* engine! *Flightfax*, 25(9), 1-3.

Wildzunas, R.M. (1997b). Unpublished research protocol, "Ground hyperstereopsis viewing device: Training and adaptation effects." Fort Rucker, AL: .U.S. Army Aeromedical Research Laboratory.

Willingham, D.T. (2007). *Cognition: The Thinking Animal 3rd Edition*. New York: Prentice Hall Inc.

Wolfe, J.M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, 1, 202-238.

Wolfe, J.M. (1998). Visual search. In: Pashler, H. (Ed.), *Attention*. London UK: University College London Press.

Wolfe, J.M., Cave, K.R., and Franzel, S.L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 419-433.

Wolfe, J.M., Kluender, K.R., Levi, D.M., Bartoshuk, L.M., Herz, R.S., Klatzky, R.L., Lederman, S.J. (2006). *Sensation and Perception*. Sunderland, MA: Sinauer Associates.

Yeh, M., Merlo, J.L., Wickens, C.D., and Brandenburg, D.L. (2003). Head up versus head down: the costs of imprecision, unreliability, and visual clutter on cue effectiveness for display signaling. *Human Factors*, 45, 390-407.

Yeh, M., and Wickens, C.D. (1998). Performance Issues in Helmet Mounted Displays. University of Illinois Aviation Research Labortory, Technical Report ARL-97-09/ARMY-FED-LAB-97-1.

Yeh, M., Wickens, C.D., and Seagull, F.J. (1998). Effects of frame of reference and viewing condition on attentional issues with helmet mounted displays. University of Illinois Aviation Research Laboratory, Technical Report ARL-98-1/ARMY-FED-LAB-98-1.

Zangemeister, W.H., and Stark, L. (1981). Active head rotations and eye-head coordination, *Annals of New York Academy of Sciences*, 374:541-59.

16 PERCEPTUAL AND COGNITIVE EFFECTS DUE TO OPERATIONAL FACTORS

Clarence E. Rash
Keith L. Hiatt
Robert M. Wildzunas
J. Lynn Caldwell
John A. Caldwell
Melvyn E. Kalich
Gregory T. Lang
Ronald P. King
Robert Noback

"Non sentient viri fortes in acie vulnera – In the stress of battle brave men do not feel their wounds" – Cicero

Introduction to Stress and Stressors

Modern combat is violent, unpredictable, and cognitively challenging, and accordingly, few would argue against the premise that battlefields are highly stressful. They involve highly mobile operations, destructive weaponry, violent combat, continuous maneuvers, and decentralized command and control. Long hours, acceleration, noise and vibration, altitude effects on the body, and potential mechanical malfunctions are just a few examples of stressors inherent in operating complex military systems. This environment is also where multiple complex decisions must be made, some of which may be life-threatening. Combat is without question a potent, multifaceted stressor that every day involves Warfighters in multiple stressful situations, although these stressors are often accepted by the Warfighter as a standard part of the operational environment with potentially little or no relief in sight. Indeed, modern and future warfare will have a degree of intensity, fluidity, and lethality previously unknown. Yet, despite the advances of technological superiority currently enjoyed by the U.S. and her allied partners, there remains an undeniable human component to combat. The promise of bloodless victories with reliance on high-tech weapons has not replaced the flesh and blood Warfighter on the battlefield. The individual Warfighter remains the characteristic enduring center-point of war. Daily, Warfighters must face hostile combat scenarios involving extreme stressors and perform successfully to survive. Even during relatively calm periods between engagements, Warfighters face stress resulting from sleep deprivation due to sustained or continuous operations, information-overload due to operating complex equipment, emotional strain from exposure to extensive destruction and dead bodies resulting from combat, and anxiety for the welfare of their fellow Warfighters and for family members left back home.

This complex myriad of job stressors puts Warfighters at risk for psychological trauma and medical difficulties ranging from problems of memory and cognition, burnout, substance abuse, and decreased task performance, to severe depression, suicidal tendencies, and Post Traumatic Stress Disorder (PTSD). Unfortunately, Warfighter mental performance directly translates to system performance, combat unit effectiveness, and operational success. Consequently, in military operations, approximately 70% to 85% of all catastrophic mishaps are caused by human error (Wiegmann and Shappell, 2003). Since the research literature and common experience tells us that stressors can affect decision-making and performance on the battlefield, it is imperative that Warfighters, and those who design equipment for their use, be aware of and address these problems. This chapter addresses the human component of the human-machine interface and the effects of operational stressors on the warfighting system operator. It also strives to link operational stress factors to perception, cognition, and human performance errors and their implications for the design of combat systems – including helmet-mounted displays (HMDs). It is incumbent on the research community to address these stresses and strains of combat and to design systems that take degraded operator performance into account.

Notwithstanding, while addressing operational factors, this chapter also recommends countermeasures, leader actions, and design issues for controlling the negative effects of operational stressors. When reading, it is important that you consider the information presented not only from an individual Warfighter's perspective, but also from the perspective of a senior leader employing his units and soldiers as a combat system across the breadth and depth of the battlespace. Designers must understand how their users cope (or fail to cope) with these stressors and how the stress information presented could relate to problems encountered by operators when using their designs.

Psychological stress

Around 1926 an Austrian endocrinologist, Hans Selye, identified what he believed was a consistent pattern of mind-body reactions that he called "the nonspecific response of the body to any demand" (Gabriel, 2006). He later referred to this pattern as the "rate of wear and tear on the body." Selye's definition of stress is necessarily broad, as stress is a broad concept. However, it incorporates two very important points: (1) that stress is a physical or "body" phenomenon and (2) that stress involves some "demand" placed upon an individual. Today we still define stress as the nonspecific physical, psychological, and physiological responses of the body to any demand placed upon it. In popular usage the term *stress* often refers to both the event (technically the "stressor") and to how we react to the event (technically the "stress reaction"). Indeed, stress is a normal reaction to any demand placed on an individual, either physically or mentally.

Operators need stress because it serves as a motivator and an indicator (e.g., increased heart rate, respiration, perspiration) that helps prepare you to respond. Inasmuch, the Yerkes-Dodson Law states that a certain amount of stress is necessary for optimum performance (Figure 16-1). If there is too little stress (astress), operators are under-aroused, bored, and inattentive. Boredom can result in increased risk-taking behaviors and declines in vigilance (a key aspect of attention). On the other hand, if there is too much stress (distress) ability to perform is limited and burnout or overload can be expected. Designers of systems must strive to design for optimal arousal (eustress) and performance such that operators remain engaged and attentive without being overly task saturated.

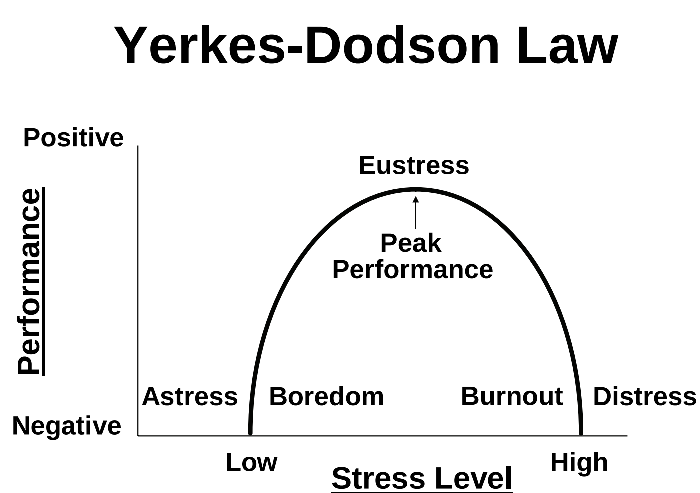


Figure 16-1. Yerkes-Dodson Law.

Additionally, odd as it may seem, some degree of stress response is also critical, in that failure to respond or adapt to stress is considered pathological. Another expression in usage for describing both stressors and stress responses is the General Adaptation Syndrome, a term used to describe the body's short-term and long-term reactions to stress and where the individual terms are described as (Selye, 1946; 1952):

General	– nonspecific response
Adaptation	– places a demand on body to adapt
Syndrome	– no adaptation = pathology

Fortunately humans have developed an inherent set of biological responses to address crisis situations. Generally, psychologists group these stress responses into a three stage process: alarm reaction, resistance, and exhaustion. In the alarm phase the bodily systems are mobilized in response to the event or demand. Therefore, when there is a perceived threat or challenge, a number of immediate, involuntary physiological changes occur – e.g., adrenaline is produced, heart rate and blood pressure increase; the pupils of the eyes dilate for better vision; the lungs take in more oxygen; the bloodstream brings extra oxygen and glucose into circulation for fuel; and digestion stops to allow the body to focus its energy on the muscles; and perspiration (required for evaporative cooling) increases. This adaptive alarm reaction is commonly called the *fight or flight response*, and it prepares the body to deal with (fight) or escape from (flight) the situation. During the resistance phase, the body maintains these efforts to cope with the threat, and eventually, if the threat is sustained, the body fatigues and fails to meet the threat challenges (exhaustion phase). If the stress response is activated too long or often, it can harm the body, causing damage to the immune system, brain, and heart.

Late 20th century psychologists found additional factors can contribute and moderate the stress experience and demonstrated that changes in the level of arousal have not been found to consistently correlate with stress (Lazarus, 1968). Specifically, the same physiological markers found to activate under stress also take place everyday in individuals who would not report being "stressed," but rather being angry, excited, etc. Inasmuch, Lazarus introduced the concept of psychological appraisal. He suggested that a primary appraisal of the event is conducted by the individual to determine the meaning of the event (positive, negative, or neutral). If appraised as negative, the individual then assesses the degree of harmfulness associated with the event. A secondary appraisal is then conducted to determine the availability of coping resources. Stress results when the perceived threat is greater than the perceived coping ability. Stress overload or prolonged stress can produce detrimental responses in individuals with poorly developed or weakened coping ability.

Responses to stress overload generally fall into one of four categories: physical responses, emotional responses, cognitive responses, and behavioral responses. As mentioned, the immediate physical response to a stressful situation involves overall heightened arousal of the body: increased heart rate, increased blood pressure, more rapid breathing, tensing of the muscles, and the release of sugars and fats into circulation to provide fuel for "fight or flight." Stress overload or prolonged stress and its continuous effects on the body may produce long-term physical symptoms such as muscle tension and pain, headaches, high blood pressure, gastrointestinal problems and decreased immunity to infectious diseases (Table 16-1).

Table 16-1.
Physical signs and symptoms of stress.

Immediate	Long-term
Sweaty palms	Sleep problems
Increasing heart rate	Backaches
Trembling	Increasing blood pressure
Shortness of breath	Immune system suppression
Gastrointestinal distress	Fatigue
Muscle tension	Anxiety disorders

Emotional responses to stress overload can range from a keyed-up sense of anxiety and irritability to social withdrawal, hostility, loss of self esteem and depression. *Anhedonia* is a symptom of depression involving an extreme loss of pleasure in activities that were once enjoyable. Persons suffering from stress overload may lose interest in hobbies and other leisure activities and find little happiness in life. If severe enough, depression could lead to suicide, but that topic is outside the scope of this discussion. The reader should only note that suicide can occur in the absence of a history of mental health problems. Extreme stress, like the loss of a loved one (or “dear John” letters in the combat zone), may cause previously healthy people to feel hopeless and consider harming themselves. In combat this could present as suicide by proxy (getting oneself killed intentionally) but may also be masked and appear to be a human error or system failure.

Prolonged stress may affect thinking (i.e., cognition) as well as emotions and behavior. This is a serious issue for Warfighters, as problems with judgment, attention, or concentration pose a great risk to personnel, the mission, and the warfighting system. For example, under high stress conditions, there is a tendency to oversimplify problem solving and ignore important relevant information, taking the “easy way out.” This is called the *simplification heuristic*.

Many individuals under high-stress conditions also tend to forget learned procedures and skills and revert to bad habits in a phenomenon called s*tress-related regression*. For example, a student aviator preparing for take-off may forget to turn on the fuel switch and then, realizing the problem and feeling stressed and embarrassed, turn the switch on and risk overheating the engine. This action is clearly contrary to training and represents a kind of regression or failure to utilize prior learning.

Yet another stress-related cognitive error is p*erceptual tunneling*. It is a phenomenon in which an individual or an entire crew under high stress becomes focused on one stimulus, like a warning signal, and neglects to attend to other important tasks or information, such as avoiding and defeating incoming fire. A similar situation may occur when Warfighters realize that they overlooked some aspect critical to mission success, such as missing a radio communication. They may then over-attend to rectifying this problem and become emotionally and mentally fixated on the error and forget other aspects of the operation, missing new information, and further compromising the mission. Beyond affecting memory, judgment and attention, stress can even decrease hand-eye coordination and muscle control.

The behavior responses to stress overload can also affect how we interact with others (e.g., at work, at home, and with friends). For example, explosiveness, social isolation, lateness to work, or a drop in work performance can be signs of stress overload. At times, stress may become so severe that alcohol is used to self-medicate anxiety or depression. Using alcohol as a coping strategy is particularly dangerous, since it impairs judgment and increases impulsivity (see *Smoking and alcohol*). Extreme stress, like the loss of a loved one, may cause previously healthy people to feel hopeless and consider harming themselves or others. This can lead to violence in the home or workplace, and like depression mentioned earlier, can result in suicide.

Psychosocial stressors are those that deal with relationships, career, and finances, as well as the factors that influence these three areas, such as your physical health. Psychosocial stress can be either positive (promotion at work, marriage, birth of a child) or negative life events (divorce or separation, death of a loved one, or illness/injury to self or family). A complete treatment of this subject is outside the scope of this work. The important thing here is to remember that psychological health enhances operator performance, and all aspects of stress have the capability to affect system effectiveness and should be considered in the initial design of a system. Given that consideration, designers must be aware that the typical user of a combat system will not be 100% capable and more likely will be operating in a diminished or degraded capacity.

Capacity to cope

Stresses decrease your capability to function in high stress environments where physical and mental capabilities must be optimal. Figure 16-2 illustrates the compounding effects of stresses, attention problems, environmental stresses, and system problems on one’s capacity to cope with normal stresses. Line *A* of the model represents the stresses most within your control. Self-imposed stresses like tobacco or alcohol use, poor sleep management, or self-medicating with over-the-counter (OTC) drugs diminish one's capacity to cope. This decrease in capacity to handle stress is depicted by a dip in line *A*. The more stresses one subjects themselves to, the more the decrease in the capacity to cope. Line *B* of the model depicts environmental and operational stresses beyond the Warfighter's control. For the most part, environmental stresses are based on mission profile, vehicle employed, and time of day. However, unpredictable stresses, like weather or mechanical problems, are also included in environmental stresses. The area between lines *A* and *B* represent the Warfighter's capacity to cope with any unknown stress the system or mission places on them.

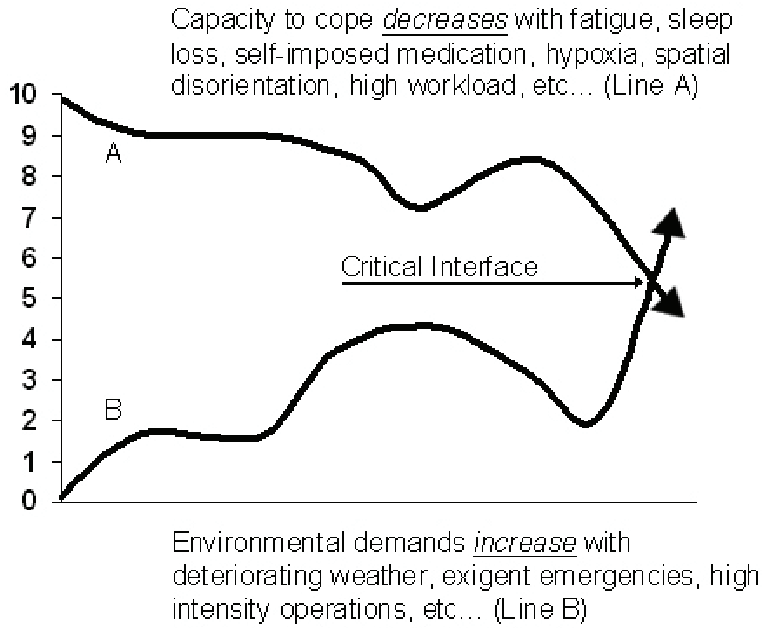


Figure 16-2. Critical interface concept for demands and capacity to cope.

If environmental stresses increase to the point where the capacity to cope and environmental demands intersect (critical interface), the Warfighter loses the capability to effectively cope with the situation increasing the potential for a catastrophic mishap. In most mishaps, there is seldom just one major factor causing the mishap. Many of the operational factors discussed in this chapter decrease one's capacity to cope, but you must remember that few occur in isolation. Instead, mishaps usually occur when many small factors add up to disrupt the Warfighters' capacity to control the system.

Therefore, it is critical to understand the importance of exposure to self-imposed stress. The closer Warfighters are to being 100% capable, the greater their ability to cope with unexpected problems or stresses arising during combat. Although we will discuss each of these concepts in greater detail later in this chapter, it is important to understand that capability decrements from self-imposed stresses result from actions taken by the Warfighters themselves. They can include the use of OTC drugs, caffeine, alcohol or tobacco. Self-imposed stresses also include nutrition, physical condition and life-style. These stresses, as well as circadian rhythm problems, can contribute to fatigue. All of these stresses strain the ability or Warfighters to function at an optimum level. Self-imposed stresses generally decrease performance, impair judgment, and decrease tolerance to other external and operational stresses.

Individual reaction to and performance during stress vary according to four major factors. The first factor is the *degree of mental effort* required by the task to be performed. In general, stress effects performance much less if an individual is engaged in a relatively simple task that is either over-learned (e.g., driving a car) or one that does not involve complex mental skills (like filling sand bags or some other form of manual labor). The *characteristics of* *the situation* in which a task is performed make up the second factor that affects the stress/performance relationship. For example, a student will do much better on a written achievement test if he is working in a quiet, comfortable room as opposed to working in a hot, noisy room. The *biological make-up* of the individual also influences the stress/performance relationship. So, an individual prone to fatiguing easily will not make a good Warfighter, where long hours and night operations are common. The fourth and final factor affecting the stress/performance relationship is one's *personality and mental health*. Individuals prone to obsession, perfectionism, and rigid thinking are less likely to perform well under stress than those persons with more flexible, realistic problem solving and decision making skills. Commonly referred to as Type A and B personalities based on how a person responds to stress. Type A persons (most common in the military) tend to respond to stress with hostility, anger, and greater competitiveness, while Type B persons tend to be easy going and cope well with stress.

Combating stress

Warfighters are competitive by nature; and combat is the ultimate win or lose scenario. Ultimately then, each Warfighter is a successful competitor, as unsuccessful ones tend to select out rather quickly. Arguably, outside of combat such competition is usually a healthy environment in which to function, but it can also be a source of stress. Constantly trying to succeed in a pressure-prone environment, impressing your superiors and outperforming your peers serve as continual sources of stress. Yet, even in peacetime, Warfighters live in a success-oriented and competitive society and are often placed in stressful environments. For example in aviation, it is obvious that an in-flight emergency evokes stress, but a checkride[1] is also a form of stress. Both elicit the same physiological response of fight or flight. Stress is useful if controlled; but a common expression remains true: "If you don't control it, it will control you."

As discussed, stress is necessary and can be both positive and negative. To be an effective Warfighter, each must learn to manage the stresses that are part of the demands of everyday life. However, to begin managing stress, it is necessary to first understand what it is and what it does – how it affects individuals both physically and mentally. Ultimately, each individual must learn to employ effective methods for coping – controlling or relieving stress. Coping techniques can be thought of as falling into one of four categories: avoiding stressors, changing your thinking, learning to relax, and ventilating.

Avoiding stressors is the most powerful technique for managing stress, since it actually prevents one from ever experiencing the full effect of a stressor. Avoiding does not mean running away from stress, however. Foresight and good planning go a long way in helping to avoid unnecessary stress. Prioritizing one's work load effectively will also help to avoid last minute crises. Planning and time management are especially important tasks for leaders, as subordinates will often model their work behavior after the examples set by their chain of command.

Realistic, mission-focused training and an effective physical training (PT) program also help prevent stress overload by providing Warfighters with the knowledge, skills, and physical endurance to perform under stressful conditions such as continuous operations, night operations, and sustained combat. If a Warfighter's comfort in performing mission essential tasks derives solely from garrison training with little realistic combat training, combat conditions and stressors will be new and unanticipated when encountered, and potentially fatal. Finally, paying close attention to communication and team coordination will also help avoid unnecessary stress and prevent mishaps. The stress of military operations can degrade communication, affecting the sound, rate, and content of speech, as well as the operator's ability to comprehend communications. Soldiers under high stress may be less precise in their messages, talk faster, and misinterpret messages more easily. The ability of a squad to work together as a cohesive team is also essential, as a number of accidents have resulted from individual members' feeling that they could not talk openly or disagree with an excessively authoritative leader.

---

[1] The *checkride* is a practical test to measure the skills developed throughout flight training. Pass/fail is based on performance against published test standards.

As discussed earlier, *how one thinks about stress* partly defines one's reaction to it, often creating a self-fulfilling prophecy. Pessimism and negativity will produce self-defeating behavior and negative results. Practicing positive self-talk is an important step toward accomplishing one's goals. Keeping a focus on what's going on *right now – today* – also helps prevent stress overload. We can't change the past, and we can only plan for the future; we cannot control the future. Spending time obsessing about past mistakes or worrying about future potential problems is distracting and creates a potential for failing at the task at hand. This is a serious danger for the Warfighter. While in combat, the Soldier's mind must be on warfighting and not on family, career concerns, or other issues past or future. Recognizing the choices one makes in life is also an important strategy for avoiding stress overload. Blaming failures and disappointments on others actually surrenders personal control and makes one's experience of life akin to being strapped down, blindfolded in the back seat of troop transport driven by someone else. It is important to make decisions, take appropriate risks, and accept responsibility for those decisions and risks. It is also important to recognize that sometimes one's decisions and actions will not be successful. When this happens, it is necessary to have the flexibility to accept setbacks and drive on, as opposed to engaging in self-pity or obsessing about repercussions.

Relaxation is an essential, albeit widely underutilized, coping technique. It is impossible to be relaxed and stressed at the same time. Find a relaxation technique that works, and use it regularly. Some examples include: meditation, yoga, self-hypnosis, reading, or pleasurable hobbies (like assembling models or listening to relaxing music).

Ventilation is the fourth and final category of coping techniques. It involves "letting off steam" either interpersonally by talking to someone or physically through exercise. Verbally expressing emotions helps resolve traumas and reduce stress and can be accomplished with a friend, family member, chaplain, or mental health professional. Exercise has long been recognized as a valuable way to "let off steam." Be careful not to overdo it, however, as this may result in injuries and thus more stress.

Fatigue

Fatigue is defined as a state of diminished mental and physical efficiency, i.e., decreased working capacity. Consequently, two types of fatigue often are discussed: mental and physical. Mental fatigue can be caused by continual mental effort and attention on a particular task, as well as by high levels of stress or emotion and exposure to environmental factors such as noise and thermal stress. The level of mental fatigue increases with time on task or exposure. As a result, tasks become more complicated to perform, concentration is reduced, and error rate is increased. Physical fatigue can result from loss of sleep, physical overexertion, medication side effects, thermal stress, and certain health problems.

While both types of fatigue may be, and often are, experienced separately, they more frequently are experienced simultaneously in the battlespace. In fact, in the warfighting environment, it is often difficult to separate the two types with respect to causes and effects. Therefore, the discussion of fatigue will be first discussed here under P*sychological stress* and then continued in the *Physiological stress* section.

We usually think of the negative effects of stress as those effects due to stress overload. Consequently, overload has deservedly received much attention as an important stressor. In overload the demands are such as to exceed the individual's ability to meet them. An example of overload is role conflict. This can be viewed as a situation in which a person finds, in essence, opposing demands being made. An individual often may be asked to work on one assignment when already having some other assignment. That person may have to stop what they are doing at that time to attend to the new task. When the issue concerns merely the sum total of work that must be done irrespective of its difficulty, we talk about quantitative overload. The person has more work than can be done in a given period of time. That person may be fully competent in the work, but time restrictions are what elicit the stress reaction. Quantitative overload could involve working for long hours without appropriate rest periods. When the work is overloading because it requires skills, abilities and knowledge beyond what the person has, then we talk about qualitative overload. The work may demand continuous concentration, innovation and meaningful decision. An important factor contributing to qualitative overload is job complexity. In general, the higher the inherent difficulty of the work, the more stressful the job. In some job situations there is a combination of both quantitative and qualitative overload; this is frequently encountered in Warfighters, particularly new trainees. Remember also, the Yerkes-Dodson Law teaches us that underload can create stress that is problematic.[2] This is a rare occurrence in intense warfighting operations (although, much of day-to-day warfighting consists of sitting around being bored). However, boring or monotonous tasks may result in astress conditions. A job may fail to provide meaningful stimulation or adequate reinforcement. Thus, jobs which involve dehumanizing monotony, no opportunity to use acquired skills and expertise, an absence of any intellectual involvement and repetitive performance provide instances of underload. In these situations boredom results from too high a degree of specialization (being overly qualified or too highly trained for the task). As with many other psychological and physiological problems, Warfighters may not be aware of task underload until they make serious errors. Sleep deprivation, disrupted circadian cycles, or life event stress may all play an additive role in producing even greater fatigue and concomitant performance decrements.

Inasmuch, it is evident that high stress overload or low stress underload both can compromise the Warfighters ability to safely accomplish the mission. Both high and low levels of stress do this by increasing fatigue in the Warfighter. Fatigue has been defined previously as a state of diminished mental and physical efficiency. Fatigue is normally caused by the common day-to-day activities a Warfighter performs. Stress can result in either *acute* or *chronic* fatigue. Acute fatigue is short-term fatigue caused by the normal daily activities of the Warfighter. It is usually remedied with a good night's sleep and rest. Unfortunately, if the Warfighter fails to remedy acute fatigue, then he begins to suffer from chronic fatigue. Chronic fatigue frequently develops gradually over time as is seen during combat deployments. However, problems can also arise when Warfighters fail to gain adequate rest in any situation where short-term fatigue evolves into long-term fatigue. For instance, when he fails to get adequate rest and sleep for several days, he becomes chronically fatigued. Other major causes of chronic fatigue include interrupted or poor sleep patterns, circadian rhythm shifts, illness, successive long missions with minimal recuperation time, and succumbing to self imposed stresses.

Chronic fatigue can lead to motivational exhaustion, commonly referred to as "burnout", and usually results from excessive unmanaged stress. Restorative measures for chronic fatigue are only temporary if stress continues. Signs and symptoms of stress related fatigue in an individual include: concentration and attention are difficult, feelings appear dull and sluggish, general attempts to conserve energy, feel or appear careless, uncoordinated, confused, or irritable. Cognitive effects include: "all or nothing" thinking, failure to focus on the here and now, and too many "musts" and "shoulds." Unfortunately, fatigue is an insidious stressor because Warfighters usually become mentally fatigued before they become physically fatigued. In fact, usually the cognitive deficits are seen by others before the physical signs and symptoms are felt by those affected.

Fatigue has a number of negative effects in the operation of complex systems. One possible result is a change in reaction time. Increases in reaction time occur because of the general decrease in motivation and sluggishness that often accompany fatigue. Decreases in accuracy also may occur, however, when individuals become impulsive and react too quickly and poorly. Fatigue also reduces attention. Fatigued Warfighters may exhibit a tendency to overlook or misplace sequential task elements, like leaving out items on a checklist. They may also become preoccupied with single tasks or elements, like paying too much attention to objects outside the system while on night vision devices, to the exclusion of checking systems and instruments inside their vehicle. Fatigue also impairs memory. Although long-term memory is reasonably well preserved during fatigue, short-term memory and processing capacity are greatly affected. Warfighters may have difficulty recalling operational events, like the location of the objective rally point, and may neglect peripheral tasks, like forgetting to check and ensure proper radio frequencies. Communication is also impaired by fatigue, as Warfighters may become more withdrawn or irritable, less clear in their speech, and more prone to misunderstanding messages. In general,

---

[2] The condition known as *underload syndrome* is defined as a lack of stimulation (such as a boring job) can result in depression and health problems, e.g., headache, fatigue and recurrent infection.

fatigued Warfighters have little awareness of their impaired performance and may feel physically okay. It is therefore important that other team members monitor each other closely in operations where fatigue is likely. Extreme fatigue can actually lead to hallucinations and problems thinking, causing the individual to appear as if they have a thought disorder or psychosis.

Combat stress, in the past commonly known as *shell shock* or *battle fatigue,* should not be confused with simple fatigue. Remember fatigue is the state of feeling tired, weary, or sleepy that results from prolonged mental or physical work, extended periods of anxiety, exposure to harsh environments, or loss of sleep. Combat stress is a specific military term used to categorize a range of behaviors resulting from the stress of battle that decrease the combatant's fighting efficiency. The most common symptoms are fatigue, slower reaction times, indecision, disconnection from one's surroundings, and inability to prioritize. Combat-stress reaction is generally short-term and can produce a wide range of behaviors, some of which are positive, such as heightened alertness, strength, and endurance, acts of courage and self-sacrifice, and strong personal bonding between soldiers. Combat stress reactions may manifest a broad range of symptoms from normal, common signs experienced by many soldiers such as hyperalertness, irritability, and loss of confidence to less frequently observed warning signs that require immediate attention such as impaired speech or muteness, impaired vision, touch, or hearing, paralysis, or hallucinations. (See Lee et al., 1997; and Solomon, 1993.)

Summary: Psychological stress

A high percentage of military mishaps are caused by user error. Some of the major contributors to human error are self-imposed stresses that decrease the Warfighter's capacity to cope with unforeseen environmental or mission-related stresses. Total prevention of stress and fatigue is impossible, but their effects can be significantly moderated. The major self-imposed stresses are self-medication with OTC drugs, alcohol and tobacco use, hypoglycemia and dehydration. Each of these contributes to fatigue, which is the crucible from which increased susceptibility to stresses such as spatial disorientation, visual illusions, and G-induced loss-of-consciousness arise. Fatigue also is a result of sleep cycle disruption and circadian rhythm shifts caused by transmeridian travel[3] (Meir, 2002). Financial, family, professional and social responsibilities are a few of the peacetime stresses which may confront the Warfighter. External stressors (noise, vibration, cold, heat, etc.) also may lead to negative behaviors associated with self-imposed stress.

Warfighters can strive to minimize self-imposed stresses; however, system designers also must both be aware of these effects and strive to design systems which are tolerant of expected degradations in operator performance. Failure to consider such effect in new designs can result in a less than optimal system at the least and loss of life at the worst. Fortunately, most unnecessary stress can be controlled or avoided with observance of the duty day, rest regulations, adequate recreation, good living quarters, and attention to morale factors. However, generally, the demands of combat are in conflict with the strain caused by internal and external stresses. Recognition, treatment, or better yet, avoidance of stress is essential for maintaining situational awareness, combat effectiveness, and ultimately safety. Resolution of the problems prior to combat is the only way to prevent them from adversely affecting system effectiveness and mission success. If efforts to resolve these stresses are unsuccessful, fault tolerant systems must be developed.

There are many ways to compensate for the effects of stress and fatigue. First, always start a task when well rested, especially when scheduled for extended missions or sustained operations. Warfighters should minimize the use of tobacco and alcohol, and ensure they are well hydrated. Proper diet can also reduce the effects of stress and fatigue. They should avoid high fat, high carbohydrate meals to reduce drowsiness instead placing emphasis on a meal moderately high in protein with moderate carbohydrates. Additionally, they should avoid eating large, filling meals prior to an operation to reduce the chance of drowsiness. It is more beneficial, in this instance, to eat several smaller meals (snacks) rather than a full meal at a single seating.

---

[3] Transmeridian travel refers to crossing a number of time zones.

The effects of fatigue and stress can be reduced through regular exercise and by staying active. During long convoys or flights, if possible, stop to get up and move around periodically. Internal vehicle lights can be turned up or down to optimize vision. Warfighters should never miss the advantage to take a nap if it can be done safely. When fatigued, it is important for Warfighters to increase their awareness of and cross-check fellow team member activities during critical tasks and functions to offset the potential for increased errors due to fatigue.

Although Warfighters are taught to push themselves to the limits of their abilities, to be tough and effective, it is wrong to think that denial of the effects of stress and fatigue will help accomplish these goals. To a certain extent Warfighters can increase their capacity to cope by eliminating or minimizing exposure to self-imposed stresses. However, stressors such as fatigue are not always controllable in the operational environment. Failing to identify and control for the effects of stress and fatigue in system design will weaken individual soldiers, units, and threaten safety and mission completion. Therefore, awareness of the causes and effects of such stresses is the key to decreasing the negative manifestations, and systems must be designed with the degraded operator in mind. Table 16-2 summarizes necessary Warfighting abilities and the potential effects of various stressors on these abilities.

The following recommendations should be considered in any Warfighter stress-reduction/endurance plan:

- *Place demands into perspective* – Doing well in military training, living comfortably, and being a good parent are all worthwhile aspirations, but they are not life threatening situations. Warfighters cannot control the reflexive physiological process that activates in a crisis situation, but they can control what they perceive as a crisis situation – the key is for them not to overreact.
- *Maintain a healthy diversity* – Entertainment and hobbies provide a healthy balance to life. A healthy balance will make the energy expanded on job and family more effective or meaningful. The military environment, particularly combat, is a demanding one, constantly changing and requiring a total mental and physical commitment from the Warfighter. The Warfighter must maintain focus and not become distracted. Any factor or condition that bothers someone enough to distract from their work is important and must be given adequate attention – but then compartmentalize! Put it away until it can be dealt with properly
- *Eliminate self-imposed stress* – Smoking, excessive drinking of alcohol, self-medicating, poor nutrition and lack of exercise are stressful and make it more difficult to deal with other stresses. Avoiding these behaviors eliminates their effect on the crewmember, minimizing self-imposed stress.
- *Maintain good physical fitness* – Regular, strenuous exercise will help resist the effects of fatigue.
- *Get plenty of natural sleep* – This is the most essential action to take for treating stress and fatigue once it has occurred. Although alcohol is the most widely used sleep aid in the U.S., alcohol-use as such is not appropriate, since it is disruptive to the quality of sleep. Alcohol will put you to sleep quickly, but later in the night you will not sleep as soundly.

Physiological Stress

Psychological stress is neruogenic (originating in the nervous system), is emotional in nature, and requires no physical interaction with the stressor(s). Physiological stress is homeostatic,[4] physical in origin, and similarly to psychological stress manifests itself in autonomic and anatomical changes (e.g., changes in blood pressure and heart rate). The physiological effects of stress occur as a result of certain biological function adjustments that occur in the body which are designed for the body to handle stress efficiently. If this response to the physiological effects of stress is present, then the individual would succumb to the hostility of the situation. The extent of

---

[4] *Homeostasis* is the ability or tendency of an organism or cell to maintain internal equilibrium by adjusting its physiological processes.

Table 16-2.
Fatigue and stress impact on critical Warfighter abilities.

Necessary (Top Level) Warfighting Abilities	
o Psychomotor coordination o Attention and vigilance o Memory	o Judgment and decision making o Prioritization of tasks o Effective communication
Stress Effects on Performance	**Fatigue Effects on Performance**
Psychomotor • Decreased tracking abilities	Psychomotor • Tracking not as smooth • Slow and irregular motor inputs

Attention
• Perceptual tunneling (decreased peripheral field of attention)
• Cognitive tunneling (narrowing salience – e.g., missed radio calls)
• Tunneling can be found with both cognitive and emotional stress
• Task shedding – entire tasks abandoned
	Attention
	• Perceptual tunneling (reduced audio-visual scan)
	• Reaction time increases
	• Errors in timing and accuracy
	• Vigilance is reduced
	• Concentration difficult
	• Lapse or “microsleeps”
	• Need enhanced stimuli salience
	  ▪ Increased volume
	  ▪ Increased contrast
	  ▪ Increased brightness

Memory
• Memory capacity declines (Short-term memory)
• Memory strategies compromised
  ▪ simplification heuristic
  ▪ speed/accuracy tradeoff
• New learning declines
• Stress related regression
	Memory
	• Diminished memory
	• Recall declines
	• Learning declines

Affect
• Group Think more common
  ▪ More confident in opinions when shared by others
  ▪ Less confidence in perceptions that contradict the majority
  ▪ Individual’s errors more difficult to identify
  ▪ Avoids personal responsibility and accountability
	Affect
	• Feel or appear dull and sluggish
	• General attempt to conserve energy
	• Feel or appear careless, uncoordinated, confused, or irritable
	• Cognitive deficits are seen before the physical effects are felt

Table 16-2. (Cont.)
Fatigue and stress impact on critical Warfighter abilities.

Stress Effects on Performance	Fatigue Effects on Performance
Combat Stress • Hyperalertness • Fear, anxiety • Loss of confidence • Impaired senses • Weakness/paralysis • Hallucinations or delusions	Communication • Impaired communication, cooperation, and crew coordination • More fragmented conversations • Misinterpretations

body's physiological response to stress varies across individuals. The physiological stress response can affect a number of organs, including brain, lungs, and heart. Apart from affecting the organs, stress also impacts the functioning of the metabolic system, immune system and cognitive function. Physiological stressors include fatigue, poor physical condition, hunger, disease. For the modern Warfighter, poor physical condition, hunger and disease usually are not of critical concern. However, fatigue is and, in this section, will be further explored via sleep loss and disruption in circadian rhythmicity. Interventions for physical fatigue also will be discussed.

Sleep deprivation

Fatigue is a significant concern for many civilian and military occupations; interest in fatigue and its potentially fatal effects within the truck driving and aviation communities has increased public awareness over the past decade. The fatigue from sleep loss and circadian factors is associated with degradations in response accuracy and speed, the unconscious acceptance of lower standards of performance, impairments in the capacity to integrate information, and narrowing of attention (Perry, 1974). Fatigued pilots tend to decrease their physical activity, withdraw from social interactions, and lose the ability to effectively divide mental resources among different tasks. In general, as sleepiness levels increase, performance becomes less consistent and vigilance deteriorates, cognition slows, short-term memory fails, frontal lobe functioning is impaired, and rapid and involuntary sleep onsets become marked (Bonnet, 1994; Dinges, 1992; Dinges and Kribbs, 1991; Horne, 1988, 1993; Koslowsky and Babkoff, 1992; Naitoh, 1975; Thomas, Sing and Belenky, 1993). Simply remaining awake for 18.5 to 21 hours can produce performance changes similar to those seen with blood alcohol concentrations of 0.05% to 0.08% (Dawson and Reid, 1997). Needless to say, the effects of lengthy duty periods are often compounded by the requirement to work and remain alert at night despite the fact that night duty is associated with a greater overall accident risk than day work (Dinges, 1995; Moore-Ede, 1993). Furthermore, is has been established that extended work shifts (i.e., those longer than 8 hours) are known to reduce the small margin-for-error that already exists in safety-sensitive jobs (Rosa, 1995).

Not only is alertness compromised by on-the-job fatigue, but several studies have found evidence of uncontrollable electroencephalogram (EEG) microsleeps[5] in pilots performing for long durations (Cabon, Coblentz, Mollard, and Fouillot, 1993; Samel, Wegmann, Vejvoda, Drescher, Gundel, Manzeu, and Wenzel, 1997; Rosekind et al., 1994; Wright and McGown, 2001). The presence of these events demonstrates that aircrew members flying today's missions often are suffering from significant cognitive difficulties while on duty

[5] *Microsleeps* are brief, unintended episodes of loss of attention associated with events such as blank stare, head snapping, and prolonged eye closure that may occur when a person is fatigued but trying to stay awake. Microsleep episodes can last from a few seconds to several minutes and often occur when a person's eyes are open.

(Belyavin and Wright, 1987; Ogilvie, Wilkinson, and Allison, 1989; Ogilvie, Simons, Kuderian, MacDonald, and Rustenburg, 1991), and such problems are at the heart of operational safety concerns. Goode (2003) concluded that flights longer than 13 continuous hours were six times more likely than shorter flights to result in fatigue-related mishaps, and a National Transportation Safety Board (NTSB) study of major accidents in domestic air carriers from 1978 through 1990 in part concluded that "…Crews comprising captains and first officers whose time since awakening was above the median for their crew position made more errors overall, and significantly more procedural and tactical decision errors" (NTSB, 1994).

Thus, long flights pose substantial risks for pilots and crews; however, it should be noted that *the duration of the flight itself* is of less importance than *the overall duration of continuous wakefulness*. Evidence of this comes both from the commercial world and from military settings. One study found that 80% of the regional airline pilots admitted to having inadvertently fallen asleep in the cockpit despite the fact that their flights were relatively short (Co et al., 1999). These pilots blamed their fatigue on scheduling issues which led to insufficient sleep (4.6 hours during "stand-up overnight" periods) and lengthy duty cycles (11.2 hours) rather than on prolonged flight durations. Given the nature of most rotary-wing operations, Army pilots may be in similar circumstances. In fact, surveys done on U.S. Army aviators (who typically make relatively short flights) and U.S. Air Force fixed-wing pilots (who sometimes engage in flights as long as 33 hours) reveal that both groups voice the same basic fatigue concerns as a consequence of insufficient sleep even though their routine flight durations differ substantially (Caldwell and Gilreath, 2002; O'Toole, 2004).

The increased sleep pressure from extended duty and the impaired arousal associated with night duty are exacerbated by sleep loss from circadian disruptions (Akerstedt, 1995a). All three factors are common in today's aviation operations. Thus, it is not surprising that the National Aeronautics and Space Administration's (NASA's) Aviation Safety Reporting System (ASRS) routinely receives reports from pilots blaming fatigue, sleep loss, and sleepiness in the cockpit for operational errors such as altitude and course deviations, fuel miscalculations, landings without proper clearances, and landings on incorrect runways (Rosekind et al., 1994). Such mistakes contribute substantially to the estimated 4% to 7% of civil aviation mishaps that are chalked up to fatigue (Kirsch, 1996), the 4% of Army aviation accidents that are considered fatigue related (Caldwell and Gilreath, 2002), and the 8% of Air Force class A mishaps that have been at least in part attributed to aircrew fatigue over the past decades (Caldwell, 2005). However, as disconcerting as these numbers are, a recent consensus report from a panel of experts suggests that the true extent of fatigue-related difficulties may be even greater. Akerstedt (2000) has asserted that fatigue is likely a causative or contributory factor in 10% to 15% of transportation mishaps; that existing statistics underestimate the real size of the problem; and that fatigue represents a greater safety hazard than drug or alcohol intoxication.

Circadian disruptions in alertness and performance

Jet lag

Rapid travel across time zones is not an uncommon occurrence in the modern 24-hour society. Air travel has become the primary mode of transportation to various locations for many people and for many reasons – academics (Takahashi et al., 2002), business, recreation, and participation in athletic and sports activities (Jehue, Street and Zuizenga, 1993; Recht, Lew and Schwartz, 1995; Wright et al., 1983). Many of these trips involve flights across multiple time zones which result in an acute condition known as "jet lag." The symptoms of jet lag have been reported to be caused by the transient internal desynchronization (i.e. mismatch or misalignment) between the internal clock that controls the sleep/wake cycle and the external geophysical clock set by the pervasive light/dark (L/D) cycle. The dissociation between the internal clock and the environmental and work or social obligations ultimately culminates in the impairment of health and productivity (Wisor, 2002). This state of temporal disarray after a change in time zone, before all rhythms return to their original internal phase-angle relationships, has been termed *transient internal desynchronization* (Moore-Ede, Kaas and Herd, 1977) or *transmeridian flight dysrhythmia* or *jet lag* (Dawson and Armstrong, 1996). Almost all individuals who travel these distances are subject to the physiological and psychological symptoms, at least for a few days. Normally, the symptoms of jet lag remit after a few days following the flight, but it may take over a week for some individuals to overcome the symptoms. Age, individual differences, number of time zones crossed, and direction of travel all contribute to the severity of jet lag symptoms (Leger, Badet and de la Giclais, 1993; Klein et al., 1970; Recht et al., 1995).

Some of the physiological symptoms of jet lag include insomnia, daytime somnolence, fatigue, stress, anorexia, nocturia, gastrointestinal discomforts, muscle aches, and head aches, (Haimov and Arendt, 1999; Cho et al., 2000; Cho, 2001). In addition, there are psychological disturbances which include moodiness/depressed mood, apathy, difficulty in concentrating, irritability, malaise, and decrements in both mental and physical performance (Bourgeois-Bougrine et al., 2003; Cabon et al., 2003; Haimov and Arendt, 1999; Minors and Waterhouse, 1988; Petrie, Power and Broadbent, 2004; Recht et al., 1995; Waterhouse et al., 2003). Women also experience delays in ovulation and menstrual dysregulation (Iglesias, Terres and Chavarria, 1980). Chronic or repeated jet lag exposure leads to cognitive decline and temporal atrophy (Cho, 2001).

Jet lag symptoms affect travelers in different ways and to different extents (Minors and Waterhouse, 1998; Waterhouse, Reilly and Atkinson, 1997). While these symptoms may be a minor inconvenience at the start and end of trips for holiday travelers, they may profoundly impair the decision making power of business executives, politicians, pilots, and aircrews. Both the severity and duration of jet lag symptoms are affected by the total numbers of time zones traveled by transmeridian flights as well as the direction of air travel (east or west). Generally, the fewer the number of time zones traveled, the lesser the side effects and discomforts (Aschoff and Wever, 1981). Eastward travelers (who were subjected to a phase advance) require more time to adjust than westward jet travelers (subjected to phase delay). The reason for this difference in adjustment is usually attributed to the ability of the internal body clock to adapt to a longer day than to a shorter day.

To successfully treat the symptoms of jet lag, the circadian rhythms must be retrained to the new time zone and result in minimal or no side effects. Numerous research groups have explored various methods to accomplish this goal, and various scientific and anecdotal observations have been proposed. One proposal states that the adjustment to a new time zone can be effectively accomplished by behavioral interventions; meals and exercise can be altered prior to or during the course of the flight schedule (Minors and Waterhouse, 1988; Winfree, 1987; Woodruff, 1988). However, no consensus has been reached regarding the exact manner to implement these behavioral changes.

Shift lag

Shift work in industrialized countries is very common, with over 27 million people in the U.S. working a shift outside the normal day shift (U.S. Bureau of Labor Statistics, 2004). The incidence of shift work in the military reflects that of the general public. According to a survey of U.S. Army aviation units, approximately 96% of the people surveyed indicated they worked a night shift at some point in their career (Caldwell and Gilreath, 2001). Personnel commonly are rotated to a night shift so that the 24-hour period will be manned at all times. Working night shift (or reverse cycle) presents problems to personnel who must be alert in order to carry out their duties. The initial period of adjustment from days to nights is particularly a problem since work still must be accomplished, but the human body is not capable of changing its internal sleep/wake rhythms quickly. Aviators are responsible for planning missions, flying aircraft, managing flight personnel, and performing a host of other duties while on reverse cycle and are faced with completing the mission even during this adjustment time. Sleepiness and fatigue can lead to dangerous consequences for all concerned. Research indicates that the problems associated with shift work, particularly night shift, include disturbed daytime sleep, and fatigue and sleepiness on the job (Akerstedt and Gillberg, 1982; Akerstedt, 1988; Penn and Bootzin, 1990; Harma, 1995). The reasons for these problems arise from the fact that the human body is programmed to be active during the day and to sleep at night (diurnal). Difficulties occur when one attempts to change these internal rhythms.

The main reason difficulties occur when working at night is due to the body's rhythms of sleep and alertness. Trying to sleep when the body's physiological arousal levels are rising is the main problem associated with daytime sleep. Most research indicates that daytime sleep is approximately 1 to 2 hours shorter than nighttime sleep (Tilley et al., 1982). While a person coming home from the night shift may have no problems initiating sleep, maintaining this sleep as long as desired is difficult at best. Early awakenings, paired with the feeling of non-refreshing sleep, are very common with day sleepers (Akerstedt and Gillberg, 1982). This shortened sleep accumulates during the course of the night shift period, increasing performance problems at night. Studies indicate that after 1 week on night duty, the night worker is functioning at the equivalent of a day worker with 1 night of sleep loss (Tilley et al., 1982).

Trying to stay alert during the time when the body's physiological signals are readied for sleep is a second problem associated with night work. The physiological tendency to sleep at night and to be awake during the day is powerful, with most research indicating that at least a week is needed for the majority of people to change their internal rhythms (Monk, 1990). Some research indicates that even permanent night workers do not adjust completely to night shift (Czeisler et al., 1990). The circadian rhythm is dictated mainly by the light/dark cycle and includes such physiological parameters as temperature, hormone secretions, and heart rate (Minors and Waterhouse, 1990). For example, high body temperature, heart rate, and blood pressure are associated with increased alertness and performance. Decreases in temperature, blood pressure, and cortisol occur in the evening, with a rise in the morning before awakening (Minors and Waterhouse, 1990). These fluctuations in various body rhythms generally occur whether we are asleep or awake. When the body's signals indicate the need for sleep, as occurs at night, the increase in sleepiness leads to decreases in performance.

These performance decrements which occur during night work are due not only to the physiological tendency to sleep during this time in the 24-hour cycle, but also from the accumulated sleep debt which occurs over the course of nights worked. Research indicates that as the number of nights accumulate for consecutive night duties, accidents increase and productivity decreases (Knauth, 1995). A study by Vidacek and associates (1986) found an increase in performance from the first to the third night of the shift, which they interpreted as circadian adjustment, but a decrease in performance occurred by the fifth night, attributable to the accumulated sleep dept. Among strategies which are used to help alleviate some of these problems is improving daytime sleep. Many techniques are suggested which may lead to better daytime sleep (Stone and Turner, 1997).

Sleep loss countermeasures

Unfortunately the scheduling demands posed by today's Warfighter missions (ground or air) are often incompatible with basic human physiological makeup, and this is at the heart of fatigue-related problems in military operations. In aviation, the multiple flight legs, long duty hours, limited time off, early report times, less-than-optimal sleeping conditions, rotating and non-standard work shifts, and jet lag that have become so common throughout modern aviation pose significant challenges for the basic biological capabilities of pilots and crews. Humans simply were not designed to operate effectively on the pressured 24/7 schedules that often define today's flight operations, whether these consist of short-haul commercial flights, long-range transoceanic operations, or around-the-clock military missions.

In order to manage fatigue that stems from acute sleep loss/sleep debt, sustained periods of wakefulness, and circadian factors, a well-planned, science-based, fatigue-management strategy is crucial (Rosekind et al., 1996). Strategies including education, behavioral countermeasures, and pharmacological interventions all have a place in preserving the safety of flight operations in both the fixed-wing and rotary-wing environments.

Education

Education about the dangers of fatigue, the causes of sleepiness while at a designated duty station, and the importance of sleep and proper sleep hygiene is one of the keys to addressing fatigue in operational military contexts. Ultimately, the Warfighters themselves and those scheduling routes and missions must be convinced that sleep and circadian rhythms are important, and that quality off-duty sleep is the best possible protection against on-the-job fatigue. Recent studies have made it clear that as little as 1 to 2 hours of sleep restriction almost immediately degrade vigilance and performance in subsequent duty periods (VanDongen et al., 2003; Belenky et al., 2003). Thus, educational programs must continue to educate leaders and Warfighters on the following points: (1) fatigue is a physiological problem that *cannot* be overcome by motivation, training, or willpower; (2) people *cannot* reliably judge their own level of fatigue-related impairment; (3) there are wide individual differences in fatigue susceptibility that *cannot* be reliably predicted; and (4) there is *no* one-size-fits-all "magic bullet" (other than adequate sleep) that can counter fatigue for every person in every situation. Warfighters and mission schedulers should be advised that it is important to: (1) make adequate off-duty sleep a priority; (2) gain 8 hours of sleep per day either in a consolidated block, or in a series of naps, whenever possible; and (3) adhere to "good sleep habits" to optimize sleep quantity and quality (Caldwell and Caldwell, 2003).

Warfighter-rest strategies

Warfighters can employ a number of strategies to reduce sleep loss. Most of these strategies have been developed within the aviation community but can be applied analogously to ground or vehicular-mounted Warfighters.

On-board sleep

For transport or cargo fixed-wing aviation, one technique for minimizing the impact of sleep loss and continuous duty is the implementation of short out-of-cockpit sleep opportunities (known as "bunk sleeps"). These sleep periods are extremely helpful for sustaining the alertness and performance of long-haul crews. In some fixed-wing military operations, an out-of-cockpit sleep strategy can be implemented in multi-crew aircraft. For B-2 bomber missions, which sometimes last for over 30 continuous hours, one of the pilots may sleep in a cot located behind the seats during low-workload flight phases while the other pilot maintains control of the aircraft. Such on-board sleep should be considered an important aviation fatigue countermeasure for any type of long-range flight operation where an adequate crew compliment is available.

Cockpit naps

Another fixed-wing counter-fatigue strategy related to out-of-cockpit bunk sleep is the cockpit nap. When cockpit naps are implemented, one pilot actually sleeps in his/her cockpit seat (rather than moving to another part of the aircraft) while the other pilot flies the aircraft. Many international airlines now utilize cockpit napping on long flights, and cockpit napping is often authorized for U.S. military flight operations as well. A 1994 NASA study has shown that naps of up to 40 minutes in duration are both safe and effective for long-haul pilots (Rosekind, et al., 1994). However, cockpit napping obviously is not feasible in dual-pilot rotary-wing aircraft, and it is worth noting that cockpit naps are not yet approved for U.S commercial operations.

Controlled rest breaks

Tasks requiring sustained attention, such as monitoring aircraft systems and flight progress, can pose significant problems for already-fatigued personnel (Dinges and Powell, 1988). This is in part why pilots often implement some type of work break strategy (chatting, standing up, walking around, or even simply swapping flight tasks – i.e., flying versus navigating) to help sustain alertness during lengthy duty periods. There is evidence from non-aviation studies that frequent rest breaks can improve physical comfort and reduce eyestrain during prolonged, repetitious tasks (Galinsky et al., 2000). More importantly, Neri et al., (2002) found that simply offering pilots a 10-minute hourly break during a 6-hour simulated night flight significantly reduced pilot sleepiness. Although positive benefits were transient (15 to 20 minutes), they were noteworthy and particularly evident near the time of the circadian trough.

Optimum crew work-rest scheduling

Since scheduling factors are often cited as the number one contributor to Warfighter fatigue, the development and implementation of more "human centered" work routines should be considered paramount for promoting on-the-job alertness. Unfortunately, crew scheduling practices in aviation have yet to incorporate the advanced knowledge of fatigue, sleep, and circadian rhythms that has been gained over the past 20 years. Concerted efforts must be made to develop schedules that recognize 1) sleep as being essential for optimum functioning, 2) breaks as being important for preserving sustained attention, and 3) recovery periods during each work cycle as being necessary to ensure full recovery from fatiguing work conditions (Dinges et al., 1996). In addition, crew schedules should include weekly rather than monthly recovery days to ensure recuperation from cumulative fatigue/sleep debt. Furthermore, scheduling practices must take into account the facts that: (1) circadian factors influence both sleep and performance; (2) homeostatic factors (continuous wakefulness) are similarly important; and (3) under certain conditions these two factors can interact to create sudden and dangerous lapses in vigilance. Also, it must be recognized that training, professionalism, motivation, and increased monetary incentives will have little impact on the basic physiological nature of circadian and homeostatic determinants of operator alertness. Finally, it is important to note that flight crews are made up of *individuals* who are differentially affected by sleep disruptions, long duty periods, circadian rhythms, and other potentially problematic factors. Thus, "one-size-fits-all" scheduling practices are almost certainly inadequate.

Sleep-promoting compounds

Sleep is often difficult to obtain, whether due to physical location (too noisy, hot, or uncomfortable), time of day (shift lag or jet lag), or physiological factors (too much excitement, apprehension, or anxiety). When the opportunity for sleep is available but is prevented by various circumstances, the limited use of sleep aids may be an appropriate solution. The U.S. military allows the use of *temazepam*, *zolpidem*, or *zaleplon* to help sleep under some situations. These hypnotics (sleep-inducing medications) can optimize the quality of crew rest in circumstances where there is an opportunity for sleep, but the situation creates difficulty in obtaining restful sleep. The choice of which compound is best for each circumstance must take several factors into account, including time of day, half-life of the compound, length of the sleep period, and the probability of an earlier-than-expected awakening, which may risk more sleep inertia effects. In addition to possible prescription hypnotics as countermeasures, "natural" substances such as melatonin are also available to personnel. While not approved for pilots, it will be included in the overview of potential countermeasures to insomnia.

Temazepam

Temazepam (Restoril®) (15 to 30 milligram [mg]) has been recommended in military aviation populations in Great Britain since the 1980's (Nicholson et al., 1986; Nicholson, Roth, and Stone, 1985; Nicholson and Stone, 1982). Most studies are mixed in whether next-day performance is affected by nighttime administration of temazepam. Roth and associates (1979) found that 30 mg of temazepam did not affect next-day alertness or performance. These findings were supported by other research (Mattila et al., 1984; Wesnes and Warburton, 1984; 1986). Wesnes and Warburton (1986) found that daytime administration of 10 mg and 20 mg of the soft capsule temazepam did not affect nighttime performance. Porcu and associates (1997) supported these findings.

However, given the long half-life of this medication, temazepam may best be used for optimizing 8-hour sleep periods that are out-of-phase with the body's circadian cycle. Under these circumstances, sleep is often easy to initiate, but difficult to maintain due to the circadian rise in alertness. The longer half-life of temazepam is desirable because the sleep *maintenance* and not sleep *initiation* is usually the problem. In addition, the pharmacokinetic disposition of temazepam is affected by the time of administration, with a faster absorption and shorter half-life and distribution after daytime administration as compared to nighttime administration (Muller et al., 1987). Research shows that temazepam facilitates daytime sleep and in studies involving simulated night operations, has been shown to improve nighttime performance by optimizing daytime sleep (Caldwell et al., 2003; Nicholson, Stone, and Pascoe, 1980; Porcú et al., 1997).

Thus, temazepam appears to be a good choice for maximizing the restorative value of daytime sleep opportunities. However, caution should be exercised prior to using temazepam in certain operational settings since the compound does have a relatively long half-life. Although residual effects were not reported in a military study in which personnel were able to gain suitable sleep before reporting for duty (Bricknell, 1991), nor in some other situations in which 30-40 mg of temazepam were given prior to a full sleep opportunity (Roth et al., 1979; Wesnes and Warburton, 1984), residual post-dose drowsiness has been reported elsewhere. Paul et al. (2004) observed that drowsiness was noticeable within 1.25 and 4.25 hours of a midmorning 15-mg dose. They also noted that psychomotor performance was impaired within 2.25 hours post dose (plasma levels were still elevated at 7 hours post dose). These data emphasize that there is certainly a possibility of sleep inertia hangover effects from temazepam's long half-life; however, the potential for this drawback must be weighed against the potential for impairment from sleep truncation in the event that temazepam therapy is withheld. Along these lines, it should be noted that Roehrs and associates (2003) found that just 2 hours of sleep loss produces the same level of sedative effect as the consumption of 0.54 grams/kilogram (g/kg) of ethanol (the equivalent of two to three 12-ounce bottles of beer), whereas the effects of 4 hours of sleep loss are similar to those of 1.0 g/kg of ethanol (five to six 12-ounce beers).

The same qualities that make temazepam desirable for maintaining the daytime sleep of shift workers make it a good choice for temporarily augmenting the nighttime sleep of personnel who are deployed westward across as many as nine time zones (Nicholson, 1990; Stone and Turner, 1997). Upon arrival at their destination, these travelers are essentially facing the same sleep/wake problems as the night worker. Namely, they are able to fall asleep quickly since their local bedtime in the new time zone is much later than the one established by their circadian clock (from the origination time zone); however, they generally are unable to sleep through the night. Based on a readjustment rate of 1.5 days per time zone crossed (Klein et al., 1970), it could take up to a week for adjustment to the new time zone to occur. Until this adjustment is accomplished, temazepam can support adequate sleep maintenance despite conflicting circadian signals, and the obvious benefit will be less performance-degrading sleep restriction. While the problem with daytime alertness due to circadian disruptions will not be alleviated, the daytime drowsiness associated with increased homeostatic sleep pressure (from sleep restriction) will be attenuated.

Thus, temazepam is a good choice when a prolonged hypnotic effect is desired as long as there is relative certainty that the hypnotic-induced sleep period will not be unexpectedly truncated. This compound is especially useful for promoting optimal sleep in personnel suffering from premature awakenings due to shift lag or jet lag since the hypnotic effect helps to overcome circadian factors that can disrupt sleep immediately following a time zone or schedule change. However, temazepam should not be used longer than is necessary to facilitate adjustment to the new schedule. Depending on the circumstances, temazepam therapy probably should be discontinued after three to seven days either to prevent problems associated with tolerance or dependence (in the case of night workers) or because adaptation to the new time zone should be nearly complete (in the case of travelers or deployed personnel) (Nicholson, 1990). When discontinuing temazepam after several continuous days of therapy, it is recommended that the dosage be gradually reduced for two to three days prior to complete discontinuation in order to minimize the possibility of rebound insomnia (Roth and Roehrs, 1991; U.S. National Library of Medicine and the National Institutes of Health, 2004).

Zolpidem

Zolpidem (Ambien®) (5 mg to 10 mg) may be the optimal choice for sleep periods less than 8 hours. This compound is especially useful for promoting short- to moderate-length sleep durations (of 4 to 7 hours) when these shorter sleep opportunities occur at times that are not naturally conducive to sleep. Just like daytime sleep in general, daytime naps are typically difficult to maintain (Costa, 1997; Lavie, 1986; Tilley et al., 1982), especially in non-sleep-deprived individuals. Furthermore, unless the naps are placed early in the morning or shortly after noon, they can be extremely difficult to initiate without some type of pharmacological assistance (Gillberg, 1984). Zolpidem is a good choice for facilitating such naps because its relatively short half-life of 2.5 hours provides short-term sleep promotion while minimizing the possibility of post-nap hangovers. Thus, it is feasible to take advantage of a nap without significantly lengthening the post-nap time needed to ensure that any drug effects have dissipated (Caldwell and Caldwell, 1998). However, as with temazepam, there should be a reasonable degree of certainty that there will not be an early interruption of the sleep period followed by an immediate demand for performance.

The efficacy of zolpidem as a nighttime sleep promoter has been clearly demonstrated in clinical trials (with up to one year of administration) in normal, elderly, and psychiatric patient populations with insomnia (Blois, Gaillard, Attali, and Coqueline, 1993). Rebound insomnia, tolerance (treatment over 6 to 12 months), withdrawal symptoms, and drug interactions are absent, and the dependence/abuse potential is low (Bartholini, 1988). Overall, zolpidem is a clinically safe and useful hypnotic drug (Palminteri and Narbonne, 1988; Sanger et al., 1987).

Zolpidem may also be helpful for promoting the sleep of personnel who have traveled eastward across three to nine time zones (Suhner et al., 2001). Unlike westward travelers who experience sleep *maintenance* difficulties, eastward-bound personnel suffer from sleep *initiation* problems. For example, a 6-hour time zone change in the eastward direction creates difficulty with initiating sleep because a local bedtime of 2300 translates to a body clock time of only 1700, and it has been well established that such early sleep initiation is problematic (Nicholson et al., 1986; Stone and Turner, 1997; Waterhouse et al., 1997). Thus, eastward travelers need something that will facilitate early sleep onset and suitable sleep maintenance until the normal circadian-driven sleep phase takes over; however, they do not need a compound with a long half life. This is because, in this example, any residual drug effect would only exacerbate the difficulty associated with awakening at a local time of 0700 that corresponds to an origination time (or body-clock time) of only 0100 in the morning. As stated above, sleep difficulties are only part of the jet-lag syndrome, but alleviating sleep restriction or sleep disruption will help to attenuate the alertness and performance problems associated with jet lag.

Thus, zolpidem is a good compound for facilitating naps of moderate durations (4 to 7 hours), even when these naps occur under less-than-optimal circumstance or at the “wrong” circadian time. Zolpidem is also appropriate for treating sleep-onset difficulties in eastward travelers. However, as is the case with any hypnotic, this medication normally should be used only when necessary, i.e., prior to circadian adaptation to a new work or sleep schedule. More chronic zolpidem administration may be essential for promoting naps that occur under uncomfortable conditions or naps that are “out of phase” since, by definition, these generally are difficult to initiate and maintain, but zolpidem probably should not be used for more than seven days to counter insomnia from jet lag. After this time, most of the adjustment to the new time zone should be accomplished (Stone and Turner, 1997; Waterhouse et al., 1997).

Zaleplon

Zaleplon (Sonata®) (5 mg to 10 mg) may be the best choice for initiating very short naps (1 to 2 hours) during a period of otherwise sustained wakefulness. Clinical trials of the hypnotic efficacy of zaleplon have shown improvement in sleep initiation, particularly with the 20-mg dose (Chagan and Cicero, 1999; Elie, Ruther, Farr, Emilien, and Salinas, 1999; Fry et al., 2000). In people diagnosed with primary insomnia, the latency to sleep onset decreased significantly compared to placebo (Chagan and Cicero, 1999). After zaleplon exerts its initial effects, the drug is subsequently (and quickly) eliminated in time for more natural physiological mechanisms to take over and maintain the remainder of the sleep period. There is evidence that there are no hangover problems as early as 6 to 7 hours later (Chagan and Cicero, 1999). Studies indicate that next-day performance is not affected by administration of zaleplon as soon as 4 hours before awakening (Mitler, 2000). Another study concluded that 10 mg of zaleplon may be taken up to 5 hours before driving with little risk of serious impairment (Vermeeren, Danjou and O'Hanlon, 1998). Another study found that 10 mg of zaleplon administered 1.25 and 8.25 hours before testing produced no serious impairment of behavioral performance, memory, or psychomotor performance at either time period (Troy et al., 2000). Other studies have not shown any residual effects in doses as high as 20 mg. Daytime mood and anxiety were not affected when zaleplon was given at bedtime, and residual sedation was not found 5 to 6.5 hours after administration of 10 mg (Forbes and Berkahn, 1999). Paul et al. (2004) found that 10 mg zaleplon increased drowsiness for 2 to 5 hours after dosing, with plasma drug levels equal to placebo by 5 hours post-dose. These authors recommend zaleplon for times when an individual may have to awaken no earlier than 3 hours after drug ingestion. Overall, 10 mg of zaleplon does not affect performance if given at least 5 hours prior to testing.

Thus, zaleplon (10 mg) is a good hypnotic for promoting short naps (2 to 4 hours) which would otherwise be difficult to initiate and maintain. In addition, as was the case with zolpidem, zaleplon can be considered useful for the treatment of sleep-onset insomnia in eastward travelers who are experiencing mild cases of jet lag. For instance, those who have transitioned eastward only 3 to 4 time zones can use this short-acting drug to initiate and maintain what the body believes to be an early sleep period. As with any hypnotic, the course of treatment should be kept as short as is reasonably possible to minimize drug tolerance and drug dependence (Nicholson, 1990). However, a study comparing 10mg zaleplon to 10mg zolpidem found that insomniac patients preferred zolpidem over zaleplon based on sleep initiation and sleep quality (Allain et al., 2003), an important point for physician's who are trying to determine which of these compounds to use.

Melatonin

Melatonin (N-acetyl-5-methoxytryptamine) is considered a chronobiotic in humans (Armstrong, 1999; Claustrat, Kayumov and Pandi-Perumal, 2002; Dawson and Armstrong, 1996; Lewy et al., 1992; Lewy et al., 1998; Pevet et al., 2002; Sack et al., 1996; Short and Armstrong, 1984; Simpson, 1980). The term *chronobiotics* refers to a chemical substance that is capable of therapeutically re-entraining short-term dissociated or long-term desynchronized circadian rhythms, or prophylactically preventing their disruption following an environmental insult (Armstrong, 2000). Melatonin is a potent synchronizer of the locomotor activity rhythms in non-human animals (Armstrong et al., 1988) as well as in humans (Kunz and Bes, 2001). Melatonin has a direct action on the central circadian pacemaker, the suprachiasmatic nucleus (SCN), to modulate its activity and influence circadian rhythms (Reppert et al., 1988; Weaver and Reppert, 1996).

The effects of exogenous melatonin in humans are generally attributed to the ability of this neurohormone to re-entrain the underlying circadian pacemaker (Pandi-Perumal et al., 2002). Properly timed melatonin administration shifts circadian rhythms, facilitates re-entrainment to a novel light/dark (L/D) cycle and alters the metabolic activity of the SCN (Reiter, 2003). Melatonin phase shifts the endogenous rhythm of core body temperature (cBT) and its own endogenous rhythms, as well as the sleep/wake cycle (Arendt et al., 1997). The beneficial effects of melatonin in alleviating the symptoms of jet lag have been extensively explored by various investigators (Arendt, 1999; Arendt and Marks, 1982; Arendt et al., 1995; Atkinson et al., 2003; Cardinali et al., 2002; Lino et al., 1993; Oxenkrug and Requintina, 2003; Parry, 2002; Petrie et al., 1989; Skene et al., 1988; for review, Herxheimer and Petrie, 2002). While the hypnotic properties of melatonin also have been demonstrated in some studies (Cajochen, Krauchi and Wirz-Justice, 2003; Stone et al., 2000), the current school of thought on the mechanism of melatonin is not as a direct hypnotic, but as a soporific agent (Reiter, 2003). It has been postulated that melatonin induces sleepiness by opening the sleep gate and exerts a slight reduction in body temperature that promotes sleep

(Gilbert, Van Den Heuvel and Dawson, 1999; Kennaway and Wright, 2002). Thus, the therapeutic benefit of melatonin for jet lag is a consequence of increasing sleep propensity by inducing an acute suppression of cBT and through a synchronizing effect on the circadian clock mechanisms (chronobiotic effect) (Arendt et al., 1995; 1997; Cardinali et al., 2002). The direction in which melatonin phase-shifts the circadian clock depends on its time of administration (Lewy et al., 1992; Reiter, 2003).

Since melatonin is considered a dietary supplement (nutraceutical), it is not regulated by the FDA as a drug in the U.S. However, the consumption of melatonin is high and caution should be observed in the uncontrolled use of melatonin. Its effects during pregnancy, potential interactions with other pharmaceuticals, long-term usage, purity of the chemical preparation, toxicity and many other considerations remain to be addressed (Arendt and Deacon, 1997; Guardiola-Lemaitre, 1997).

General precautions for hypnotic therapy

Sleep promoting compounds can be useful for promoting sleep in operational contexts where there are problems with sleep initiation or sleep maintenance. However, it should be noted that, like all medications, there are both benefits and risks associated with the use of these compounds. These should be considered by the prescribing flight surgeon, the aviation safety officer, and the individual pilot before the decision to utilize hypnotic therapy is finalized (U.S. military pilots are never *required* to use hypnotics of any type). A hypnotic of any type should not be used if a person is on-call and may be awakened for immediate duty at any time. Although temazepam, zolpidem, and zaleplon are widely recognized as being both safe and effective, personnel should be cautioned about potential side effects and instructed to bring these to the attention of the unit flight surgeon. Potential problems may include morning hangover which may cause detrimental effects on performance, dizziness and amnesia that may be associated with awakenings that are forced before the drug has been eliminated, and various idiosyncratic effects (Balter and Uhlenhuth, 1992; Menkes, 2000; Nicholson, 1990; Roth and Roehrs, 1991). If any difficulties occur, it may be necessary to discontinue the specific compound or to abandon hypnotic therapy altogether. However, it is likely that significant side effects can be reduced or eliminated by using an alternate compound or by modifying dosages or dose intervals (Nicholson, 1990). For these reasons, military personnel are required to experience a test dose of the hypnotic of interest under medical supervision before using the medication during operational situations. Even after the test dose yields favorable results and it is clear that operationally-important side effects are absent, hypnotics should be used with particular caution when the aim is to aid in advancing or delaying circadian rhythms in response to time-zone shifts (Nicholson, 1990; Stone and Turner, 1997; Waterhouse et al., 1997). Reviews by Waterhouse and associates (1997), Nicholson (1990), and Stone and Turner (1997) offer detailed information on this rather complex issue. While melatonin is available over the counter, it is not authorized for use in military pilots.

Alertness-enhancing compounds

For those situations in which, despite everyone's best intentions, adequate sleep opportunities are simply nonexistent, stimulants (or alertness-enhancing drugs) will help to stave off the deleterious effects of fatigue (prescription stimulants are an option only for military pilots). Unavoidable manpower constraints, hostile environmental circumstances, extremely high workloads, and unexpected enemy attacks all may require a postponement of sleep until a break in the operational tempo permits rest and recuperation. Although stimulants should not be viewed as a substitute for proper staffing or adequate work/rest cycles, they can be life saving in circumstances in which sleep deprivation is unavoidable (Cornum, Caldwell and Cornum, 1997). Stimulants are effective and easy to use, and because their feasibility is not dependent upon environmental manipulations or scheduling modifications, their usefulness, especially for short-term applications, can be significant (Caldwell and Caldwell, 2005). Caffeine, modafinil, and dextroamphetamine are approved for certain aviation operations by the

U.S. Air Force, and both caffeine and dextroamphetamine are approved for limited use by the U.S. Army and Navy.[6]

Caffeine

Caffeine is a good choice for situations where medical oversight is limited or not available. This is because caffeine is not a controlled substance, and prescriptions are not required. Also, since caffeine is already in widespread use and is generally viewed as quite safe, there is little concern that there will be adverse physiological consequences associated with its ingestion. Caffeine is available in a number of forms (i.e., tablets, candy, gum, food, and beverages). An 8-ounce cup of drip-brewed coffee contains an average of 135 mg of caffeine, an 8-ounce cup of brewed tea contains approximately 50 mg of caffeine, and a 12-ounce cola drink contains an average of 44 mg caffeine, ranging from 23 to 58 mg, depending on the drink. An 8-ounce cup of Starbucks™ contains 250 mg of caffeine (Center for Science in the Public Interest, 1996).

Although caffeine can produce some minor side effects (Committee on Military Nutrition Research, 2002; Serafin, 1996), in general, these are inconsequential in comparison to the improvements in reaction time and cognitive performance, the improved mood, and the reduction of sleepiness in fatigued subjects (Bonnet et al., 1995; Committee on Military Nutrition Research, 2002; Lieberman et al., 1987; Wyatt et al., 2004). Militarily-focused studies at the Walter Reed Army Institute of Research (WRAIR) (Silver Spring, MD) have shown that 600 mg single-dose caffeine is beneficial for sustaining the performance and alertness of sleep-deprived personnel kept awake for over 50 continuous hours (Wesensten et al., 2002). Other researchers have found that 150 mg to 300 mg bolus doses of caffeine are sufficient to increase performance over placebo when the sleep deprivation period is short, for example less than 24 hours (Penetar et al., 1993).

Despite these and other positive findings, wholesale dependence on caffeine to mitigate the effects of sleep deprivation in the military operational aviation environment is controversial since the effects of tolerance have not been adequately studied (Wyatt et al., 2004). Over 80% of adults in the U.S. daily consume behaviorally active doses of caffeine (Griffiths and Mumford, 1995), and it is possible that acute caffeine administration in operational contexts may not effectively alert severely fatigued individuals. Nonetheless, caffeine should be considered a "first line" approach to pharmacologically-based alertness enhancement because caffeine has been shown to exert a number of positive effects. No medical oversight of caffeine use is required as long as the caffeine comes in the form of coffee, soft drinks, chocolate, or other standard foods and beverages.

Modafinil

Although modafinil (Provigil®) (100-200 mg) is a relatively new alertness-enhancing substance, there is substantial evidence that it is useful for sustaining performance during continuous or sustained military operations (Lagarde and Batejat, 1995). These authors found that the drug reduced episodes of microsleeps and attenuated decrements in reaction time, math, memory-search, spatial-processing, grammatical-reasoning, letter-memory, and tracking tasks. Wesensten et al. (2004; 2002) found 200 mg to 400 mg modafinil to be effective for restoring the performance and alertness of sleep-deprived non-pilots in a typical research setting, however, it was concluded that neither modafinil nor dextroamphetamine (20 mg) offered greater efficacy than a 600-mg dose of caffeine.[7] In active-duty military pilots, Caldwell et al., (2000a) found that 200 mg of modafinil every 4 hours

[6] Note that such approvals are generally "Service-wide" rather than location specific. For instance, U.S. Air Force policy authorizes the use of modafinil for dual-seat bomber missions longer than 12 hours in duration, and authorizes dextroamphetamine on a wider basis for similar circumstances. Although individual units or bases can choose not to utilize these compounds, they are not permitted to authorize the use of medications that have not been officially sanctioned by the U.S. Air Force, Army, or Navy without obtaining a waiver from higher headquarters.

[7] Note that high-dose caffeine should be used judiciously in pilots because adverse reactions such as nausea and vomiting sometimes occur.

maintained flight performance and basic cognition at near-well-rested levels despite 40 hours of continuous wakefulness. However, there were reports of nausea and vertigo that were attributed to the large cumulative dose (600 mg within a 24-hour period). A more recent study with U.S. Air Force F-117 pilots indicated that 100-mg doses of modafinil administered every 5 hours sustained flight control accuracy to within 27% of baseline levels, whereas performance under the no-treatment condition degraded by over 82% during the latter part of a 37-hour period of continuous wakefulness (Caldwell et al., 2004). Similar beneficial effects were seen on measures of alertness and cognitive performance. Furthermore, the lower dose (in comparison to those used in the Caldwell et al., 2000 study) produced these positive effects without causing the side effects noted in the earlier study.

The frequency of adverse side effects with modafinil is low, drug tolerance seems nonexistent even after weeks of continuous use, and the abuse liability is limited (Cephalon, 1998). As a result, modafinil is easier to dispense compared to dextroamphetamine which will be discussed shortly. Another advantage of modafinil is that it appears to have a relatively small adverse effect on recovery sleep even when given fairly close to the time of sleep onset (Buguet et al., 1995). Thus, modafinil may be an optimal choice for use in sustained military operations in which there is a moderate possibility that a short break in the operational tempo could provide an unexpected sleep opportunity. Initial concerns that modafinil caused overconfidence in sleep-deprived people (Baranski and Pigeau, 1997) have not been substantiated by more recent research (Baranski et al., 2002). Nonetheless, modafinil has not been as widely assessed as caffeine and amphetamine in normal, sleep-deprived people engaged in real-world tasks (Akerstedt and Ficca, 1997); work with clinical populations suggests that modafinil is less effective than amphetamine (Mitler and Aldrich, 2000); and some believe that there is insufficient information available concerning modafinil's long-term safety and efficacy (Banerjee, Bitiello and Grustein, 2004). However, for short-term fatigue management, modafinil should be considered a possible option because of its alertness-enhancing capacity and its favorable side-effect profile. Future military policies may make modafinil more widely available in the aviation setting, but pilots will first need to "pass" a ground test for adverse effects and sign an informed consent agreement for using modafinil for an "off label" indication (i.e., none of the prescription alertness-enhancers have been explicitly approved for keeping sleep-deprived but otherwise normal people awake).

<u>Amphetamine</u>

The effects of dextroamphetamine (Dexedrine®, 5 mg to 20 mg) have been well-researched. In comparison to caffeine, amphetamine appears to offer a more consistent and prolonged alerting effect (Mitler and Aldrich, 2000; Weiss and Laties, 1962), and in comparison to modafinil, some reports suggest it is more efficacious (Lagarde and Betejat, 1995; Mitler and Aldrich, 2000). However, there is some disagreement on this point as three other reports have suggested that dextroamphetamine is equivalent to modafinil for sustaining the performance of sleep-deprived normal individuals in sleep-deprivation periods of up to 40 hours (Caldwell, 2001; Pigeau et al., 1995; Wesensten et al., 2004). Real-world operational comparisons of dextroamphetamine to caffeine or modafinil are currently nonexistent due to the difficulties of conducting such studies under warfare conditions. Consequently, simulator-based studies continue to be the best alternative and are still ongoing (e.g., Estrada et al., 2008)

Although dextroamphetamine can produce side effects such as palpitations, tachycardia, elevated blood pressure, restlessness, euphoria, and dryness of mouth (Physician's Desk Reference, 2009), the properly-controlled administration of this compound remains a viable strategy for the sustainment of combat performance in select military aviation operations where sleep is difficult or impossible to obtain. The U.S. Navy's guide for Flight Surgeon's and the U.S. Army's guide for leaders both discuss policy-based guidance for the use of dextroamphetamine in sustained and continuous flight operations (U.S. Army Aeromedical Research Laboratory, 1996; U.S. Navy Aerospace Medical Research Laboratory, 2001), and the U.S. Air Force has authorized the use of dextroamphetamine in certain types of lengthy (i.e., 12 or more hours) bomber and fighter flight missions.

In a study conducted at the Walter Reed Army Institute of Research (Newhouse et al., 1989) a 20-mg dose of dextramphetamine produced marked improvements in mathematical ability, a gradual improvement in logical-reasoning, better performance on choice reaction-time, and an increase in alertness in non-pilots who were sleep deprived for 48 hours. A 10-mg dose produced similar effects, but they were fewer and shorter in duration. In two studies conducted on pilots at the Army's Aeromedical Research Laboratory (Caldwell, Caldwell and Crowley, 1997; Caldwell et al., 1995), repeated 10-mg doses of dextroamphetamine maintained flight performance, an array of cognitive skills, and alertness indicators close to well-rested levels despite 40 hours of continuous wakefulness. These results were later confirmed in an in-flight study with 40 hours of sleep loss (Caldwell and Caldwell, 1997) and in a follow-on simulator study in which the sleep-deprivation period was extended to 64 hours (Caldwell et al., 2000b). Data from actual field environments have further established amphetamine's capacity for reducing the impact of fatigue (McKenzie and Elliot, 1965; Tyler, 1947; Winfield, 1941), and there are reports of beneficial amphetamine effects in combat situations such as Viet Nam (Cornum, Caldwell and Cornum, 1997), the 1986 Air Force strike on Libya (Senechal, 1988), Operation Desert Shield/Storm (Cornum, Cornum and Storm, 1995; Emonson and Vanderbeek, 1995), and Operation Iraqi Freedom (Kenagy et al., 2004). To date, no major side effects or other problems have been reported from the medical use of dextroamphetamine in several military settings (referenced above), and concerns about "judgment impairments" are to some extent negated by reports that amphetamine decreasing risk-taking behavior and the sleep-loss-induced liberal response bias often seen on cognitive tests in sleep-deprived subjects (Newhouse, et al., 1989; Shappel, Neri and DeJohn, 1992) without impairing the ability of such subjects to self-evaluate their own performance (Baranski and Pigeau, 1997).

Thus, dextroamphetamine is a viable counter-fatigue medication useful for military aviation missions in which significant fatigue is a risk factor; however, amphetamine should only be used under proper medical supervision since this medication possesses significant abuse potential. As with modafinil, the use of dextroamphetamine to counter the effects of fatigue in healthy individuals requires an informed-consent agreement for off-label use as well as a suitable ground test to rule out idiosyncratic reactions.

Summary: Physiological stress

Fatigue (the most uncontrollable physiological stressor) is a known risk factor in the operational environment, and it warrants treatment with scientifically-validated fatigue countermeasures. Since a large percentage of operator fatigue stems from insufficient sleep, the best countermeasure would be to avoid sleep deprivation by: (1) ensuring adequate manpower levels to properly staff all work periods; (2) consider scheduling of naps or taking advantage of opportunities for naps; and (3) establishing work/rest schedules that enable personnel to gain sufficient restorative sleep in their off-duty hours. However, if real-world demands disrupt or prevent sleep, and behavioral or administrative counter-fatigue strategies are found to be insufficient or impractical, pharmacological adjuncts can help to safely sustain alertness.

In the event that sleep opportunities are available but compromised due to operational factors that prevent the onset and maintenance of restful sleep, the hypnotics temazepam, zolpidem, and zaleplon should be considered. Temazepam is best for maintaining sleep for relatively long periods during the night or for optimizing daytime sleep, while zolpidem and zaleplon are better for promoting an earlier-than-usual sleep onset or for inducing and maintaining short naps. Also, as discussed earlier, these compounds can help to minimize sleep disruptions associated with circadian factors (jet lag and shift lag). In this regard, the choice of compound depends on when the new sleep opportunity is offered and the probability that the sleep period will be unexpectedly truncated. An effort should be made to balance the need to improve sleep with the need to avoid residual effects, taking into account the effects of sleep restriction versus any residual effects which may occur from medication-induced sleep.

The duration of prescription sleep medication therapy should be kept as short as possible, usually for only a few days, to help with jet lag symptoms, or intermittently to help with shift lag symptoms. While the modern hypnotics are much safer and shorter acting than the hypnotics of years past, caution is still needed with prolonged use of any hypnotic. Continued use of hypnotics for several weeks or months may lead to tolerance or dependence, but the extent of these problems remains an issue of debate (Menkes, 2000; Roth and Roehrs, 1991). In addition, sudden withdrawal after several weeks of therapy may lead to rebound insomnia (Menkes, 2000; Nicholson, 1990).

When considering the use of medications for aid in operational contexts, the following points should be kept in mind: (1) drugs are not a substitute for good work/rest scheduling; (2) sleep-promoting and alertness-enhancing compounds should not be administered to personnel indiscriminately or in the absence of proper medical oversight; and (3) with regard to situations devoid of sleep opportunities, there has not been a drug of any description that has been found capable of indefinitely postponing the basic physiological need for 8 hours of restful daily sleep. However, clearly there are circumstances that warrant the operational use of pharmacological fatigue countermeasures, and in these situations, properly-administered, appropriately-supervised medication therapies can enhance both the safety and effectiveness of military aviation personnel.

It is well known that sleep deprivation affects performance, whether the deprivation is due to long work hours or to shortened sleep length due to changes in work schedule or time zone. A pilot's task in flying the aircraft requires divided attention and vigilance, both of which are affected by long work hours and lack of adequate rest. When the pilot requires the aid of an HMD while flying the aircraft, additional complexity is added to the task, thereby potentially lowering performance even further (Brown, 2004). Therefore, risk of in flight performance errors increase with the combination of sleep deprivation and wearing HMDs, and the pilot and crew should be aware of this increased risk. Countermeasures to decrease the impact of sleepiness on performance will be useful when HMDs or any other complicating factors are added to the equation, however, there is no research to indicate which countermeasures will address the added risk of flying with HMDs specifically. Thus, a pilot's best strategy will be to recognize the potential for fatigue-related dangers and take general steps to ensure optimal alertness given the circumstances of the mission.

Self-Imposed (Internal) Stressors

Use of approved over-the-counter and prescription medications

Warfighters, as a subsection of the general population, are overall in better physical condition than the general public. This is a consequence of fairly stringent medical selection criteria that all prospective Soldiers are required to meet prior to induction in an all volunteer force as well as mandatory physical training and a strongly encouraged regimen of extramural exercise and recreational sports. However, even when battlespace-related injuries are discounted, Warfighters, like all civilians, face disease and other physical maladies. Consequently, Warfighters can be expected to need both prescription and OTC medications. In addition, as with their civilian counterparts, Warfighters will use other legal substances believed to be health or performance enhancers.

Approved medications

The effects and concerns of medication use (both prescription and OTC) on human operational performance are important for the Warfighter. This is also true in the HMD environment. Although the entire Warfighter community currently has access to HMD technology and often employs it, the aviation community certainly has the most experience. Consequently, the Aerospace Medicine community (i.e., Flight Surgeons) has a greater depth of knowledge and experience on the effects that medications may have in both the general flight environment and the HMD flight environment than do the corresponding ground-based Surgeons. All branches of the military services have published specific guidance for Flight Surgeons regarding the use of medications in the flight environment. However, and of note, our ground-based counterparts do not have any in-depth nor published guidance documents that indicate when a Warfighter placed on specific medications should be restricted from various occupational activities – to include the use of HMDs – but instead mainly rely on past experience and the art of medicine. It should be noted that in the current Middle Eastern theater of operations, the Army has stated that approximately 12% of the combat forces in Iraq and 17% in Afghanistan are taking prescription antidepressants and or “sleeping pills” in order to cope with harsh operational demands – both of which would adversely affect not only physical performance but also to a greater extent the individual Warfighter’s performance with HMDs (Thompson, 2008). Although the vast majority of the discussion that follows will be Aviation Medicine-centric, most is applicable to ground HMD use for all of the Warfighter force.

The use of medications by the Warfighter in any situation is of concern because these products contain a number of chemical compounds that may have negative physical performance effects and neurological (cognitive, perceptual and sensory) effects on the human body. Additionally, an underlying condition also may significantly affect these parameters. For the HMD user (both aviators and ground forces), this fact must be further emphasized due to the very unique perceptual environment that an HMD presents to the user and the often complex cognitive processes that our Warfighters must use to interpret what is being presented to them on the HMD as compared to what their actual four-dimensional (4-D) environment is. The ability to know where you are in an operationally harsh and complex battlespace is paramount to individual Warfighter survival and mission accomplishment.

Whenever an aircrew member presents with a medical condition requiring an OTC or a prescription medication, a Flight Surgeon must evaluate the condition and the proposed course of drug therapy to determine if that aircrew member can continue to fly. In the U.S. Department of Defense (DOD), all three services have different regulations and guidance regarding the treatment of aircrews, what medications they can take, which ones are waiverable for flight, and what processes they must go through to return to the crew member to aviation duty. In general, any medication “grounds” an aircrew member, even if it is a waiverable medication. An aircrew member on a waiverable drug is usually returned to flight status only after an observation period. Drugs for aircrew members must be prescribed by, or with the knowledge of, a Flight Surgeon. Almost any medication can impair a person’s ability to fly an aircraft but, more importantly, the condition being treated is often more of a factor in “grounding” a pilot or aircrew member than the drug itself. For example, amoxicillin is a relatively benign drug used commonly for otitis media (middle ear infection). The drug is quite safe, but the middle ear infection may impair the ability to fly. The pilot can fly when the condition resolves, even though he or she may have several more days to complete the course of antibiotics. Conversely, many conditions are fairly benign, but the medications required for treatment can significantly impair cognition, judgment or the sensorium such that safe flight or the optimal use of HMDs would be severely and negatively affected.

As to each individual service, the Army has AR 40-501 (Department of the Army, 2008) that determines medical standards of fitness and AR 40-8, *Temporary Flying Restrictions Due to Exogenous Factors Affecting Aircrew Efficiency* (Department of the Army, 2007). They also have published numerous Aeromedical Policy Letters (APLs) that address medication waivers (Department of the Army, 2006). The medication policy letters break medications down into 5 classes: (1) OTC medications; (2) no waiver action required or information only; (3) chronic use; (4) chronic use requiring waiver; and (5) mandatory disqualifying medications. All of these APLS are web accessible (Department of the Army, 2006).

The Air Force has Air Force Instruction 48-123 which covers use of medications in Air Force aircrew members (Department of the Air Force, 2006). This extensive instruction governs medications, medical conditions, medical standards, etc. affecting aircrew members as well as special duty operators, missile crews, ground controllers, and so forth. Medications may be: (1) approved for use without medical consultation; (2) approved for use by a flight surgeon without removal from flying duty; (3) require a waiver (specifies level of the command structure that waiver must come from), or 4) not waiverable. Medications listed as not waiverable may be approved or granted a waiver after physiological testing at the U.S. Air Force School of Aerospace Medicine, Brooks AFB, Ohio. They also have a waiver guide similar to the Army APLs and again these are internet accessible (U.S. Air Force School of Aerospace Medicine, 2008).

The U.S. Navy has Navy Instruction 3710 (NATOPS General Flight and Operating Instructions) (Department of the Navy, 2001), which includes information governing Navy aircrew members that is similar to AR 40-8 and AR 40-501. Also available on the web is the Naval Operational Medicine Institute (NOMI) guidance regarding medical conditions and medications in Navy aviators (Naval Aerospace Medical Institute, 2007a; 2007b). The guidance and list of waiverable medications is a bit longer than the Army's list. As a final consideration for our civilian aviation HMD users, the Federal Aviation Administration (FAA) also provides some guidance on medical conditions and medication use which is again web available (FAA, 2008).

In treating an aviator, the Flight Surgeon's view regarding medication use is that any aircrew member should be evaluated for restriction from flying duties when initiating any medication and also be advised of potential side effects. Additionally when we place an individual on a medication, we consider: (1) Is the medication compatible with flight duty and, more importantly, is the underlying medical condition compatible with flight duty; (2) is the medication effective and essential to treatment; and (3) is the individual free of aeromedically significant side effects after a reasonable observation period. Since medication side effects are very hard to predict and they occur with irregularity and often differently in any given individual, Flight Surgeons are quite cautious in prescribing patterns. They are especially cautious prescribing medications whose side effects relate to central nervous, cardiac, ophthalmologic, and labyrinthine systems. Finally they also consider the unique environmental considerations present in the aviation environment, i.e., G-forces, hypoxia, pressure changes, noise, heat, cold, acute and chronic fatigue; and how these effect the medication or the underlying medical condition. Since flight surgeons determine suitability for flight duty, they are inherently determining suitability for HMD use (Batchelor, 2006; Orford and Silberman, 2008).

For example, in prescribing practices, U.S. Army flight surgeons rely on the guidance provided by the Commander, U.S. Army Aeromedical Center, Fort Rucker, AL, who has reviewed and classified a wide range of medications for use in the aviation environment. Medications are designated Class 1, 2A, 2B, 3 and 4. Medications not discussed in the APL are currently incompatible with the aviation environment or little information of its safe use in the aviation environment exists. New medications are reviewed constantly and waiver requests are considered on a case-by-case basis but often take a great deal of time to process (Department of the Army, 2006). The following is a brief discussion of the Army model of medication classification for aviation; however the other services do a similar system of classification.

- **Class 1 Medications:** These are OTC medications that may be used without a waiver. Occasional and infrequent use of these OTCs does not pose a risk to aviation safety nor does it violate the intent of AR 40-8. Generally OTCs are approved for acute non-disqualifying conditions and do not require a waiver. They may be used in accordance with standard prescribing practices. Note however, that OTC medications are frequently combination medications, with one or more components contra-indicated for safety of flight. Many OTC medications do not provide a listing of ingredients on the package and often give quite sketchy information on side effects. Also of note is that aircrew members require constant alertness requiring full use of all senses and reasoning powers. Many OTC medications as well as most prescribed medications cause sedation, blurred vision, disruptions of vestibular function, etc. Often the condition for which the medication is used is mild; however, it can produce very subtle effects which may also be detrimental in both the flight and the HMD environment. Just as with the subtle deterioration of cognitive ability that occurs with hypoxia and alcohol intoxication, medication effects may not be appreciated by the individual taking the medicine. These effects may have disastrous results in situations requiring full alertness and rapid reflexes. Of a final note is that all OTCs should only be used infrequently and for short periods of time. The list of approved army OTCs is found in Table 16-3 (Since medications are constantly being reviewed, the reader is directed to refer to the APLs for all other classes of medications on the web) (Department of the Army, 2006).

Table 16-3.
Class 1 Medications
(Department of the Army, 2006)

Type	Comments
Antacids	Tums, Rolaids, Mylanta, Maalox, Gaviscon, etc.®
Antihistamines	Loratidine - Short term use by individual aircrew is authorized but the aircrew member must report use of this medication to the Flight Surgeon as soon as possible. The Flight Surgeon must also be concerned not only with the use of this medication but also the underlying problem that the individual is self- treating (e.g. allergic rhinitis) and the aeromedical implications of the diagnosis.
Artificial Tears	Saline or other lubricating solution only. Visine or other vasoconstrictor agents are prohibited for aviation duty.
Aspirin/Acetaminophen	When used infrequently or in low dosage.
Cough Syrup/ Cough Lozenges:	Many OTC cough syrups contain sedating alcohol, antihistamines or Dextromethorphan (DM) and are prohibited for aviation duty.
Decongestant	Pseudoephedrine - When used for mild nasal congestion in the presence of normal ventilation of the sinuses, and middle ears (normal valsalva).
Pepto Bismol	If used for minor diarrhea conditions and free of side effects for 24 hours.
Multiple Vitamins	When used in normal supplemental doses. Mega-dose prescriptions or individual vitamin preparations are prohibited.
Nasal Sprays	Saline nasal sprays are acceptable without restriction. Phenylephrine HCL may be used for a maximum of 3 days. Long-acting nasal sprays (oxymetazoline) are restricted to no more than 3 days. Recurrent need for nasal sprays must be evaluated by the flight surgeon. Use requires the aircrew member to be free of side effects.
Psyllium Mucilloid	When used to treat occasional constipation or as a fiber source for dietary reasons. Long term use (over 1 week) must be coordinated with the flight surgeon due to possible side effects such as esophageal/bowel obstructions.
Throat Lozenges	Acceptable provided the lozenge contains no prohibited medication. Benzocaine (or similar analgesic) containing throat spray or lozenge is acceptable. Long term use (more than 3 days) must be approved by the local flight surgeon.

- **Class 2A medications:** These are medications which are available by prescription only, have proven to be quite safe in the aviation environment. These medications, when dispensed and their usage monitored by Flight Surgeons, have been quite effective in returning aviators more rapidly to their respective flying positions. While generally safe, one still must take into consideration the underlying medical condition and the ever present possibility of side effects. Note that occasionally the underlying health condition dictating the need for the medication may require a waiver; and if the medication is required on a frequent or maintenance basis, a waiver may also be needed (Department of the Army, 2006).
- **Class 2B medications:** This classification of drugs requires a prescription and must be used under the supervision of the flight surgeon. Unlike Class 2A, they are often employed for chronic long term use and more likely to be used for underlying medical conditions which require a waiver. They also have greater potential for side effects, so all must have a period of observation of at least 24 hours (Department of the Army, 2006).
- **Class 3 medications:** These medications are generally given for treatment of underlying conditions which require a waiver, may have significant side effects, or require significant evaluations as follow-up for safe use. Specific requirements are given under each drug or drug category listed below. Other

requirements as dictated by the underlying medical condition also may be added at the discretion of the Consultant, Aeromedical Activity (Department of the Army, 2006).

- **Class 4 medications:** These medications are strictly contraindicated in the aviation environment due to significant side effects. The underlying cause or need for use of these medications may result in a permanent disqualification or require a waiver for return to flying duty. Generally, a period of continuous grounding is mandatory from the initiation of therapy through cessation of these drugs plus a specified time period to rid the drug completely from the body (usually at least three half lives). Continuous use of these medications is incompatible with continuation of aviation status (Department of the Army, 2006).

In conclusion, the use of medications by the Warfighter is of great concern, first because the underlying condition may significantly affect cognitive, sensory and physical performance and second because of the multiple influences that medications have on these performance parameters. Again, it cannot be overemphasized that for the HMD user who has a unique view of his environment and the 4-D battlespace surrounding him, any decrement in his ability to interpret where he is in an operationally harsh combat scenario is critical to his survival and mission accomplishment.

Dietary supplements

In both Western and Eastern cultures, before the advent of modern pharmacology, individuals relied on naturally occurring substances (e.g., plants, minerals and animal parts) as healing agents. While some of these substances are the basis of many modern drugs, sometimes the cure was worse than the ailment.

2000 B.C.	“Here, eat this root.”
A.D. 1000	“That root is heathen. Here, say this prayer.”
A.D. 1850	“That prayer is superstition. Here, Drink this potion.”
A.D. 1940	“That potion is snake oil. Here, swallow this pill.”
A.D. 1985	“That pill is ineffective. Here, take this antibiotic.”
A.D. 2000	“That antibiotic doesn't work. Here, eat this root.”

-- *Anonymous*

In between are a host of substances that have developed a wide following along side modern drugs as remedies and as dietary *supplements* that are believed to promote health or enhance performance.

As with the use of prescription and OTC medications, the effects and concerns of the use of dietary supplements on Warfighter physical, cognitive and perceptual performance in the HMD environment is another important issue. Dietary supplements include vitamins, minerals, proteins, botanicals/herbs, amino acids, metabolites (including ergogenics) and extracts. In a recent survey, it was noted that the annual sales of dietary supplements in the United States was approaching $16 billion. Additionally, on average, 1,000 new products are developed each year. Although manufacturers are restricted from claiming that using their products leads to therapeutic benefits, surveys show that many people take supplements for purposes such as treating colds or alleviating depression. Surprisingly, the majority of consumers don’t believe these products are definitely safe nor work as promised, but still continue to use them (Institute of Medicine, 2004).

Unlike prescription medications, which are highly regulated by the FDA, dietary supplements are regulated under the auspices of the Dietary Supplement Health and Education Act (DSHEA). This act was passed in 1994 and states that dietary supplements are to be regulated like foods instead of drugs, meaning that they are considered safe unless proven otherwise and are not required to be clinically tested before they reach the market. It is up to the U.S. Food and Drug Administration (FDA) to determine whether a particular substance on the market is harmful, based upon information available in the public domain. Thus, it is fairly obvious that the use of dietary supplements is largely unregulated in both the U.S. military and civilian populations (De Smet, 2002; U.S. Congress, 1994). Many dietary supplements are ineffective, and some have been found to be dangerous (Gardner et al., 2007; Lonn, 2003; Noonan and Noonan, 2006; Solomon et al., 2003; U.S. Federal Register, 2004). To illustrate, Table 16-4 shows a few of the more commonly consumed dietary supplements and their purported benefits contrasted with many of their reported problems. In addition, Table 16-5 provides a list of supplements classed with regard to their well-documented side/toxic effects. Both Tables were compiled from information available in the Physician's Desk Reference (PDR) for Herbal Medicines (PDRHealth, 2007). Finally, due consideration must be given to the fact that some dietary supplements interact with prescribed medications (Scott and Elmer, 2002; Wilson et al., 2006).

As a large organization with a focus on health protection and readiness, the DOD is dedicated to maintaining the health and well-being of the armed forces; these responsibilities include policy development and education regarding dietary supplements. Additionally, the military has the responsibility to train and maintain its members at an optimal readiness posture as well as a mission performance standard. As such, DOD has the responsibility of guiding its service members to make appropriate decisions that best enhances their health, including nutrition.

Unfortunately, as with all sectors of the U.S. population, the use of dietary supplements to promote health has become increasingly popular among members of the military. The prevalence of use among service members has been well documented in a number of reports. For example, in one study, a dietary supplement survey was administered to 2,215 males (mean age, 25 years; range, 18 to 47 years) entering U.S. Army Special Forces and Ranger training schools. Eighty-five percent of the men reported past or present use of a supplement, 64% reported current use, and 35% reported daily use (Arsenault and Kennedy, 1999). In another study, a U.S. Army Special Forces unit was studied to determine characteristics of supplement users and found that most Warfighters (87%) reported current supplement use (Bovill, Tharion and Lieberman, 2003).

Supplements available to service members range from those that might impart beneficial effects to health and performance with negligible side effects to others that have uncertain benefits and might be potentially harmful to health and performance. Furthermore, the military, cognizant of the potential benefits of dietary supplements, is conducting research on some promising supplements. However, there are no service wide military policies (e.g., education or regulations) to guide commanders in management practices for safe use of dietary supplements. With this in mind, the Committee on Military Nutrition Research (CMNR) convened an add hoc working group – the Committee on Dietary Supplement Use by Military Personnel to assist in the assessment of the effects that dietary supplements, whether beneficial or detrimental, might have on different military service members and for some subpopulations facing heightened risks (e.g., Special Forces, Rangers, aviators). They were also asked to review the patterns of dietary supplement use among military personnel (Tables 16-6 and 16-7) (Lieberman, 2008), to recommend a framework to identify the need for active management of dietary supplement use by military personnel, and to develop a systematic approach to monitor adverse health effects. The committee was further tasked with selecting a subset of dietary supplements and, by examining published reviews of the scientific evidence, identifying those that are beneficial or warrant concern. This group has recently published an extensive guide regarding their initial findings of the use of supplements by the military along with the requirements for continued monitoring and research (Institute of Medicine, *in press*).

As with our earlier discussion of medications, the use of dietary supplements by the Warfighter in any situation is of concern because these products contain substances that may have a variety of effects that are not adequately documented. With the HMD user (both aviators and ground forces), this fact must be emphasized due to the very unique perceptual environment that an HMD presents to the user and the complex cognitive processes that Warfighters must use to interpret what is being presented on the HMD as compared to what actual 4-D environment. Undoubtedly, some dietary supplements have clear benefits, some have uncertain benefits, and

Table 16-4.
Common Supplements: Issues and Problems
(PDRHealth, 2007)

Supplement	Purported Benefits	Reported Problems/Issues
Echinacea	Purported benefit is stimulation of cellular immune system Commonly used for fevers, colds, bronchitis and "tendency towards infection"	Long term use not recommended due to unknown effect on immune system with chronic use Not for use in immune system/autoimmune diseases (Multiple Sclerosis, RA, Lupus, etc) or in those with documented allergies to plants in the *Asteraceae/Compositae* family (ragweed, chrysanthemums, marigolds, and daises).
Saw Palmetto	Used primarily as treatment for Benign Prostatic Hypertrophy (BPH). Fairly good evidence that it relieves symptoms in mild BPH (decreased nocturia, hesitancy, post-void dribbling and improved stream).	No change in objective parameters such as prostate size or PSA levels. May cause Gastro- intestinal upset similar to the side effects of radiation therapy.
Creatine	Used as body building supple-ment. Research is conflicting, and there are mixed results in literature. May have mild benefit for less conditioned weight lifters.	No documented benefit in endurance activities. Weight gain of 1to 3 Kg. Questionable true muscle growth – since discontinuation results in loss of weight and muscle size. Heavy use may lead to cramps, nausea, diarrhea, dehydration. Risks of long term use unknown. A decrease in endogenous creatine production has been noted. Case reports of heat casualties with use.
Ephedra	Stimulant found in many body building/weight loss supplements that are advertised to improve endurance.	Reported deaths/disabilities in healthy, young individuals due to use. Possibility exists for sudden incapacitation due to stroke, and heart attack. Banned by the FDA.
DHEA and Androstenedione	Precursor of androgens (testosterone) and estrogen. The so called "Fountain of Youth". Believed effects are anabolic secondary to steroid conversion and possible osteoblast stimulation as well as promotion of protein anabolism.	Banned by NCAA, NFL, IOC. Side effects like anabolic steroids. Many effects are reversible after discontinuation. However, irreversible virilization and gynecomastia has been noted. May potentially increase the risk of hepatic, uterine and prostate CA. Possible positive effect on HDL and total cholesterol.

others are unsafe, especially if taken in combination with medication or in certain work environments. The short term effects of some of these preparations are dangerous and use can result in incapacitation. The long term effects of many of these unregulated preparations are unclear and have not been studied to any degree in the HMD environment. The bottom line is that many of the supplements contain a number of chemicals that can have negative overall health effects, physical performance effects, and neurological (cognitive, perceptual and sensory) effects on the human body, and this can greatly impact an HMD Warfighter's ability to know where he is in an operationally harsh and complex battlespace, which is vital to his survival and mission accomplishment. Again, flight surgeons, under the auspices of various regulations and published guidance (e.g., AR 40-8 and the APLS for

Table 16-5.
Supplements classes by side/toxic effects.
(PDRHealth, 2007)

Supplement Class	Examples
Herbals causing increased bleeding time	Ginseng , Gingko, Garlic, Feverfew
Plants with sedative properties	Hops, Valerian Root, St. John's Wort, Hemlock, Opium Poppy, Passion Flower, Skullcap Mushroom, Wild Lettuce, Wolf's Bane
Hallucinogenic plants	Peyote, California Poppy, Kava-kava, Mandrake, Nutmeg Periwinkle, Thorn apple, Yohimbe Bark
Cardiac active plants	Ma Huang (Ephedra) , Foxglove (Digitalis) both Yellow and Purple, Squill/White Squill, Broom, Lilly of the Valley Pheasant's Eye
Liver toxic plants	Germander, Comfrey, Chapparal, Life Root

the US Army) (Department of the Army, 2006; 2007) can attempt to strictly regulate what the aircrew HMD user's are allowed to consume, but ground based surgeons generally do not have that same guidance when dealing with their HMD Warfighters.

Nutrition

Self-imposed stresses such as fatigue and hypoglycemia are reduced by taking proper care of your body. Certain life-style factors that contribute directly to health and well-being also result in decreased stress effects and optimal performance. Two tools that can be used effectively to increase combat performance and increase resistance to fatigue are a proper healthy diet incorporated with a well-rounded exercise program that includes both aerobic and anaerobic exercise.

In order for the human body to function, it must have fuel to burn, specifically the sugar glucose. Glucose liberated during the digestion process enters the blood stream and is transported to the organs and tissues needing it. If there is apparent excess to the body's needs, it is stored as glycogen in the liver itself. The nervous system in general, i.e. the brain, nerves and especially the retina in the back of the eye, are all highly -dependent on blood sugar levels to function. When glucose levels in the blood fall below levels adequate to supply these tissues, the liver converts glycogen to glucose and releases it into the blood stream. Hypoglycemia results when the glycogen stores in the liver are depleted and there is not enough glucose in the blood stream. Hypoglycemia means "low blood sugar" and has a variety of causes. The most common cause is skipping meals or eating foods that are predominantly simple sugars. Other causes of hypoglycemia are high protein/low carbohydrate diets and diets where a Warfighter does not eat for extended periods of time (fasts or starvation diets).

Short-term symptoms of hypoglycemia are shakiness, decreased mental ability, physical weakness, irritability, fatigue and sleepiness. These symptoms arise within 4 to 6 hours after the last meal. However, if the meal consisted primarily of complex carbohydrates, like pasta, potatoes, or whole wheat breads, hypoglycemia does not occur as quickly. If the last meal consisted of simple carbohydrates, like those found in candy and soft drinks, then hypoglycemia occurs much more quickly because of the rapid digestion and rapid metabolism of the simple sugars. Complex carbohydrates, proteins and fat require more time for digestion and utilization. Their glucose is slowly released into the blood and stored in the liver over a period of time, avoiding erratic shifts in metabolism. Simple carbohydrates are absorbed into the blood quickly, causing the blood sugar level to rise dramatically. As the blood sugar rises, the brain senses there is too much glucose in the blood and signals the pancreas to release insulin into the blood stream which acts to remove glucose from the blood and take it to the liver. Unfortunately,

Table 16-6.
Supplement use in the Army-wide survey exercise frequency.
(Institute of Medicine, in press)

	Total N	At Least Once/ Week	1-2 Supplements/ Week	3-4 Supplements /Week	5 + Supplements /Week	Multi-vitamin Use	Sport Drinks Use	Protein Amino Acid Mix Use	Money Spent per Month
Sex									
Male	553	58%	31%	12%	15%	32%	19%	14%	$58
Female	63	71%	37%	20%	14%	37%	20%	10%	$58
Age									
≤20	70	57%	28%	9%	20%	24%	17%	9%	$56
21-29	283	59%	28%	10%	15%	26%	21%	13%	$47
30-39	143	69%	35%	17%	15%	44%	19%	18%	$67
≥40	133	63%	40%	19%	10%	43%	16%	13%	$67
Education									
High School or GED	194	52%	27%	8%	14%	21%	14%	12%	$46
Some College/Associate Degree	289	64%	32%	12%	16%	33%	20%	14%	$71
Bachelor Degree	107	78%	40%	24%	13%	52%	25%	17%	$53
Graduate Degree	39	72%	40%	39%	8%	41%	23%	15%	$25
Occupation									
Combat Arms[a]	219	57%	33%	10%	15%	29%	19%	16%	$64
Combat Support[b]	214	61%	31%	16%	15%	36%	21%	16%	$51
Combat Service Support[c]	192	60%	32%	15%	13%	33%	19%	10%	$59
Rank									
E1-E4	298	53%	30%	9%	15%	24%	21%	12%	$59
E5-E9	195	61%	31%	16%	15%	39%	14%	16%	$67
WO1-WO5	100	67%	36%	18%	13%	43%	15%	13%	$48
O1-O3	31	88%	47%	22%	19%	50%	38%	22%	$32
O4-O6	5	60%	20%	40%	0%	40%	40%	0%	$40
Exercise – Aerobic									
None	28	52%	28%	3%	21%	35%	14%	7%	$31
≤4 days a week	234	66%	38%	17%	12%	40%	21%	16%	$48
≥5 days a week	361	56%	29%	12%	15%	28%	19%	13%	$66
Exercise – Strength									
None	176	45%	28%	9%	8%	19%	14%	4%	$73
≤2 days a week	141	62%	34%	16%	13%	36%	28%	10%	$46
≥3 days a week	306	67%	34%	14%	19%	40%	19%	22%	$55

[a] Infantry, armor, field artillery, air defense, special forces, aviation
[b] Engineer, chemical, military intelligence, military police, signal, civil affairs
[c] Ordinance, quartermaster, transportation, legal, medical, finance, chaplain

Table 16-7.
Supplement use and exercise frequency among survey populations.
(Institute of Medicine, in press)

	Army Wide		Army War College		Ranger	Special Forces
	Male	Female	Male	Female	Male	Male
Use Supplements at least once a week	58%	71%	72%	82%	82%	66%
Use 1-2 different supplements per week	31%	37%	29%	36%	45%	41%
Use 3-4 different supplements per week	12%	20%	14%	6%	19%	14%
Use 5+ different supplements per week	15%	14%	12%	23%	15%	7%
Multivitamin Use	32%	37%	39%	52%	23%	32%
Sport Drinks Use	19%	20%	10%	0%	41%	36%
Protein/Amino Acid Mixture Use	14%	10%	3%	0%	18%	17%
Creatine Use	5.2%	0%	2%	0%	19%	16%
Exercise Frequency						
Aerobic Exercise >3 times / week	91%	90%	75%	N/A	98%	96%
Aerobic Exercise >5 times / week	60%	50%	N/A	N/A	N/A	65%
Strength Training >3 times / week	50%	31%	34%	N/A	45%	36%

if the blood sugar levels are high, insulin removes most of the sugar, leaving a blood sugar level that is lower than before the candy was eaten.

Long-term symptoms of hypoglycemia can include convulsions and fainting, usually occurring as a result of large swings in blood sugar levels. One of the major effects of hypoglycemia is a lapse in mental processes. When the brain cannot get the glucose it needs from the blood, it begins to slow down. For the Warfighter, common symptoms could include math errors, checklist errors, and decreased attention span which cause missed communication errors and perception errors.

To prevent hypoglycemia Warfighters must eat regularly. When meals are missed, snacks of complex carbohydrates are more beneficial than candy and soft drinks. Some snacks designed to keep the amount of sugar in the blood at a constant level include bagels, pretzels, fig or fruit bars, granola bars, yogurt, milk, fresh fruits and vegetables. The bottom line on nutrition and combat is to eat sensible meals containing complex carbohydrates low in fat, at regular intervals. If accustomed to eating three meals a day, then try not to skip a meal since the glycogen stores in the liver may become depleted. Avoid fad diets or high protein/low carbohydrate diets designed to build bulk. Furthermore, protein is an inefficient source of energy and is primarily used to build muscle and bone. Carbohydrates, however, are efficient sources of energy and are easily converted to glucose.

Diet pills should not be relied upon to maintain weight. They often contain the same medications found in decongestants (discussed in the medications section of this chapter). They are stimulants with unwanted side effects including nervousness, tremors, increased blood pressure and heart rate, dehydration due to increased sweating, and sleep disturbances. There is a significant synergistic effect when diet pills are used in conjunction with caffeine. This effect includes a marked increase in blood pressure and increased dehydration. Weight loss can be accomplished without diet pills; a sensible diet and a regular exercise program is a much healthier and safer alternative for losing weight.

Dehydration

Dehydration, like hypoglycemia, is a major contributor to fatigue. There are varying degrees of dehydration, with different symptoms. Unfortunately, most people are constantly in a slightly dehydrated condition. When dehydration is combined with the combat environment, fatigue onset is quicker. Also, in the aviation environment, dehydrated aviations are at a higher risk of experiencing decompression sickness, spatial disorientation, visual illusions, airsickness, and loss of situational awareness.

The first common indication of dehydration is a sensation of thirst. At this point, the Warfighter is about 2% dehydrated or about 1.5 quarts (1.6 liters) low on water. If combined with the diuretic effects of caffeinated drinks (coffee, colas) Warfighters can quickly become 3% or more dehydrated. At a dehydration level of 3%, they may experience sleepiness, nausea, mental impairment, and mental and physical fatigue. After a night of drinking alcoholic beverages, the 3% dehydration level is reached more quickly because of the diuretic effects of alcohol. In addition to mental impairment, dehydration decreases your ability to do high intensity physical work. The best method to prevent the problems of dehydration, obviously, is to drink plenty of water before, during and after each operation. If water is unappealing or unpalatable, drinks that are low in sugar, nonalcoholic, and decaffeinated can be substituted. Many Soldiers prefer “sports drinks” like Gatorade®. These drinks are marginally helpful, but some contain higher amounts of salt than the body normally needs. In addition, some of the drinks are heavily sugared. Usually, Warfighters won’t lose enough salts or electrolytes during normal activity to warrant the use of these types of drinks. However, if they prefer sports drinks to water, then its recommended they drink whatever they like best providing it is not alcoholic, caffeinated or heavily sugared. Staying hydrated before, during and after exertion has a pronounced positive effect on how well you perform combat related duties.

Smoking and alcohol

There are two very commonly used drugs not discussed in the preceding section, *Smoking and alcohol*. The acts of imbibing of these drugs, smoking and drinking, are very prevalent in both the civilian and military communities. Tobacco products are primarily used as stimulants; alcohol is a central nervous system (CNS) suppressant. For historical and social reasons, the use of these drugs are not prohibited or severely limited, although many occupations and especially the aviation community does place some time-related restrictions on the use of alcohol prior to the associated vocational activity. In the discussions to follow, it will be shown that these drugs do have a significant influence on Warfighter performance, especially on visual and cognitive performance. Long-term health effects also have been associated with their use.

Tobacco

First, the effects and concerns of the use of tobacco products on Warfighter physical, cognitive and perceptual performance and specifically their impact in the HMD environment is discussed. Tobacco comes from the plant Nicotinia Tabacum that has in it the drug *nicotine*. Nicotine is a poisonous alkaloid contained in the leaves, roots and seeds of tobacco plants. It is used as an insecticide as well as in some medications, primarily and ironically in smoking cessation medications.

Historically, the military has had a reputation as an environment in which tobacco use is accepted and common. As with the civilian community, military personnel use all forms of tobacco, to include cigarettes, cigars, pipes and smokeless tobacco. Overall in the U.S. DOD population, the prevalence of tobacco use has been reported as 51% in 1980, 53% in 1982, and 47% in 1987 (Edwards, Sanders and Price, 1988). However, when Edwards, Sanders and Price (1988) investigated the impact of smoking on U.S. Army aviation initial-entry rotary-wing (IERW) training flight school performance, they reported only 15% as smokers. In recent years, the DOD has increased efforts to lower tobacco use by members of the Armed Forces, and the rate has declined. Nevertheless, in a recent 2005 survey it again was found that tobacco use remained moderately high among military personnel (Figure 16-3) (Department of Defense, 2005).

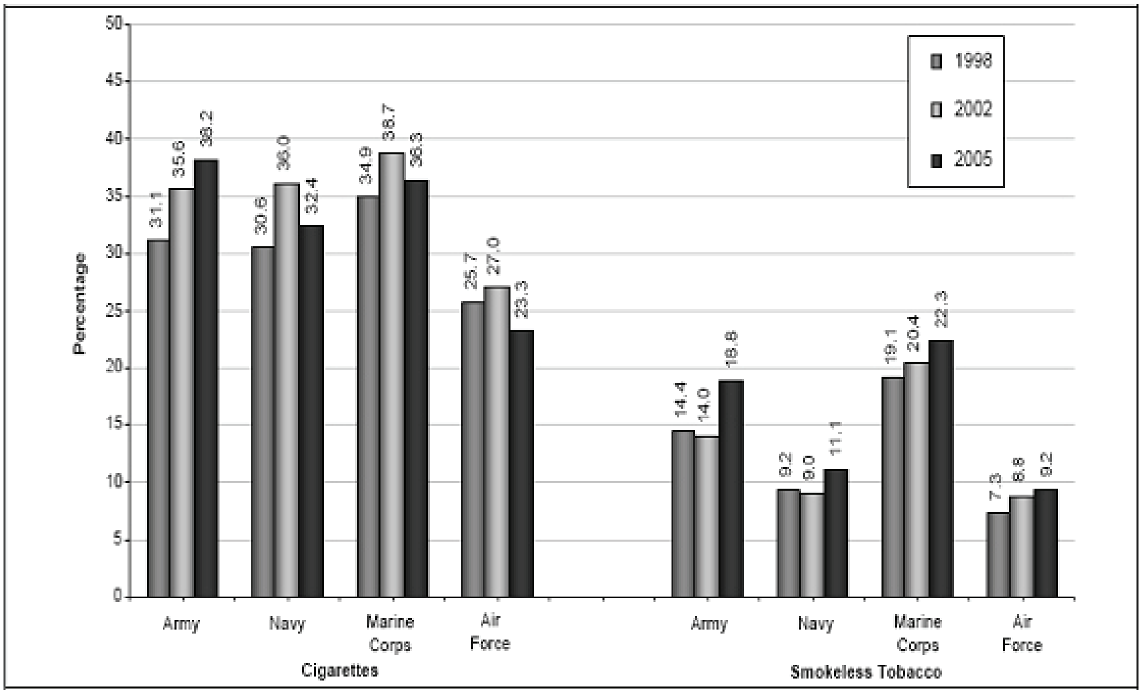


Figure 16-3. Service comparisons in the prevalence of any cigarette use and smokeless tobacco use, past 30 days, 1998-2006 (Department of Defense, 2005).

This high rate of tobacco use is of concern to the DOD for the following reasons:

- Smoking-related illnesses take a toll on the physical readiness of the Armed Forces. Thousands of studies have demonstrated an association between the use of tobacco and negative health outcomes, such as cardiovascular diseases, various cancers, and pulmonary disease (Haddock et al., 1998).
- The use of tobacco also has been associated with negative performance outcomes, such as higher absenteeism, diminished motor and perceptual skills, and poorer endurance (Chisick, Poindexter and York, 1998).
- There is a financial concern. Each year, the DOD spends an estimated $875 million on smoking-related health care and productivity loss (Conway, 1998).
- There is a concern that most of the individuals currently serving in the Armed Forces will eventually return to civilian life, and the DOD has an obligation to return veterans to the civilian sector in the healthiest condition possible (Chisick, Poindexter and York, 1998).

The use of tobacco products by the Warfighter in any situation is of concern because these products contain nicotine and a number of other chemicals that have negative overall health effects, physical performance effects, and neurological (cognitive, perceptual and sensory) effects on the human body.

With the HMD user (both aviators and ground forces), this fact must be further emphasized due to the very unique perceptual environment that an HMD presents to the user and the often complex cognitive processes that Warfighters must use to interpret what is being presented to them on the HMD as compared to their actual 4-D environment.

Overall health effects

From the DOD standpoint, having the healthiest Warfighter population is of utmost importance and holding on to these highly-trained individuals is paramount. Cigarette smoking has been declared as hazardous to health by numerous world health organizations, due to its contributions to hypertension and chronic lung disorders such as bronchitis and emphysema. Tobacco contains at least 28 known carcinogens (cancer-causing agents). The most harmful carcinogens in tobacco are the tobacco-specific nitrosamines. They are formed during the growing, curing, fermenting, and aging of tobacco. Other cancer-causing substances in smokeless tobacco include N-nitrosamino acids, volatile N-nitrosamines, benzo(a)pyrene, volatile aldehydes, formaldehyde, acetaldehyde, crotonaldehyde, hydrazine, arsenic, nickel, cadmium, benzopyrene, and polonium-210. All tobacco, including smokeless tobacco, contains nicotine, which is an addictive substance (National Cancer Institute, 2008).

Tobacco is one of the strongest cancer-causing agents. Tobacco use is associated with a number of different cancers, including lung cancer, as well as with chronic lung diseases and cardiovascular diseases. Lung cancer is the leading cause of cancer death among both men and women in the United States, with 90% of lung cancer deaths among men and approximately 80% of lung cancer deaths among women attributed to smoking. Cigarette smoking remains the leading preventable cause of death in the United States, causing an estimated 438,000 deaths – or about one out of every five deaths each year (National Cancer Institute, 2008).

Tobacco users also increase their risk for cancer of the oral cavity. Oral cancer can include cancer of the lip, tongue, cheeks, gums, and the floor and roof of the mouth. People who use oral snuff for a long time have a much greater risk for cancer of the cheek and gum than people who do not use smokeless tobacco.

The possible increased risk for other types of cancer from smokeless tobacco is being studied. Possible increased risks for heart disease, diabetes, and reproductive problems are being studied (Centers for Disease Control and Prevention, 2004; National Cancer Institute, 2008).

Physical performance effects

Warfighters need to be in top physical condition to be able to survive the harsh operational environments that they operate. Several studies have shown that smoking is associated with impaired cardiovascular fitness and reduced heart rate response to exercise. Chronic smoking is found to affect young male smokers' cardiovascular fitness, impairing the efficiency and decreasing the capacity of their circulatory system. It is not known whether these associations are present in adolescence or whether they change over time. But moderate to heavy smoking (≥10 grams of tobacco per day)[8] has been shown to reduce cardiovascular fitness and heart rate response to exercise in young otherwise healthy smokers (Bernaards et al., 2003; Papathanasiou et al., 2007).

Sensory, perceptual and cognitive effects

From an HMD Warfighter perspective, the ability to think clearly, see well, and react quickly and appropriately are the key requirements to survival and the successful execution of the mission. As a stimulant, nicotine has been found to improve performance on attention and memory tasks. Clinical studies using nicotine skin patches have demonstrated the efficacy of nicotine in treating cognitive impairments associated with Alzheimer's disease, schizophrenia, and attention-deficit/hyperactivity disorder (ADHD) (Levin et al., 2006; Levin and Rezvani, 2002; Rezvani and Levin, 2001). Experimental animal studies have demonstrated the persistence of nicotine-induced working memory improvement with chronic exposure, in addition to the efficacy of a variety of nicotinic agonists. Nicotine has also been shown in a variety of studies in humans and experimental animals to improve cognitive function. Nicotinic treatments are being developed as therapeutic treatments for cognitive dysfunction. Several studies have found that transdermal nicotine significantly improves attentional function in people with

[8] Approximately 14 to 20 cigarettes.

Alzheimer's disease, schizophrenia or Attention Deficit Hyperactivity Disorder (ADHD) as well as normal nonsmoking adults.

Nicotine studies also have been conducted on smooth pursuit eye movements and have showed that nicotine administered by patch improved antisaccade performance and smooth pursuit eye movements. Nicotine also induces loss of anticipatory saccadic eye movements; provides for improved acceleration of eye movements during smooth pursuit initiation; and improves pursuit gain during the maintenance phase (steady-state velocity). However, nicotine does not appear to modify peak predictive pursuit. Thus, Nicotine appears to improve visual attention (Kumari, 2003).

Nicotine is known to improve performance on tests involving sustained attention and recent research suggests that nicotine may also improve performance on tests involving the strategic allocation of attention and working memory. Nicotine improves visual search performance by speeding up search time and enabling a better focus of attention on task-relevant items. This appears to reflect more efficient inhibition of eye movements towards task irrelevant stimuli, and better active maintenance of task goals. When the task is novel, and therefore more difficult, nicotine lessens the need to refixate previously seen letters, suggesting an improvement in working memory (Zingler, 2007).

A few studies have shown that nicotine may improve the ability of humans to focus on auditory information and filter out background noise (Baldeweg et al., 2006; Harkrider et al., 2001). In one study, nonsmokers received nicotine transdermally and their auditory processing was measured. These measurements indicated that nicotine in these nonsmokers appeared to improve the transmission of information in the midbrain and cortex. These areas are believed to involve processing of auditory information related to alertness to changes in the environment and also to the screening of sensory input (Harkrider and Champlin, 2000). On the other hand, clinical studies also have suggested that cigarette smoking is associated with hearing loss, a common condition affecting older adults. One study showed smokers were 1.69 times as likely to have a hearing loss as nonsmokers (Cruickshanks et al., 1998). Two other studies showed that smoking was associated with increased odds of having high frequency hearing loss in a dose-response manner (Mizoue et al., 2003; Nakanishi et al., 2000).

Nicotine has well-known, unpleasant side effects, e.g., transient dizziness, nausea, and nicotine-induced nystagmus (NIN). Motion stimulation increases nicotine-induced dizziness and nausea, but does not significantly influence NIN or postural imbalance. The view is that all measured adverse effects reflect dose-dependent nicotine-induced vestibular dysfunction. Additional motion stimulation aggravates dizziness and nausea, i.e., nicotine increases sensitivity to motion sickness (Zingler, 2007).

Of even greater concern are the effects that smoking tobacco has on night vision. Early studies showed a significant decrease in scotopic dark adaptation with smoking, which was attributed to the hypoxic effects of carbon monoxide (CO). Later studies found that smoking seemingly improved night visual performance on some psychophysical tests. This improvement was presumed to be a result of the stimulant effect of nicotine. More recent studies have reported that smokers have reduced mesopic vision when compared with nonsmokers (Miller and Tredici, 2002).

Although the literature is somewhat confusing,[9] smoking is discouraged in aviation for several reasons, which include:

- There is some evidence that it may degrade mesopic and night vision.
- Although many night flights are low level, the hypoxic effect of CO is additive with altitudinal hypoxia. Cigarette smoke contains a minute amount of carbon monoxide. Just three cigarettes smoked at sea level will raise the physiological altitude to between 5,000 and 8,000 feet (ft) (1500 and 2400 meters [m]). The effect of altitudinal hypoxia on night vision is primarily one of an elevation of the rod and cone threshold. Although decreased cone function is clearly demonstrated by the loss of color

[9] In that the comparison with pure nicotine drug administration via e.g. skin patch vs. cigarette smoking confounds the meta-analysis.

vision at hypoxic altitudes, the decrement in central visual acuity is usually insignificant. However, scotopic (night) vision at altitude can be significantly reduced. Scotopic vision has been reported to decrease by 5% at 3,500 ft (1,050 m), 20% at 10,000 ft (3,050 m), and 35% at 13,000 ft (4,000 m), if supplemental oxygen is not provided. Thus, the use of oxygen, even at low pressure altitudes, can be very important at night (Miller and Tredici, 2002).

- Smoke is a significant irritant for aircrew who wear contact lenses or for those with dry eyes.
- Smoke forms filmy deposits on windscreens, visors, and spectacles and HMDs that can degrade contrast at night.
- The effects of smoking withdrawal during long missions may be dangerous.
- The chronic long-term effects of smoking are hazardous to overall health.

When any Warfighter is required to fly or to rapidly ascend to elevations greater than 10,000 ft (3,050 m) (common elevations found in areas of current conflict such as the mountains along the Afghanistan and Pakistan border as well as those in northern Iraq), it has been noted that they will experience substantial impairment in cognitive performance. Because of their CO load, Warfighters who smoke are already at a physiologic altitude of between 5,000 and 8,000 ft (1500 to 2400 m) above sea level (ASL) thus only compounding the issue and placing them at even a higher physiologic altitude. For example studies have shown that activities requiring decisions, strategies, and memory retention are more vulnerable than automatically performed activities, complex tasks are affected more than simple tasks, and tasks that are not already well learned at sea level will be difficult to learn or perform, especially during initial exposure to altitude. Also, initial exposure to high altitude will likely also adversely affect mood, balance, reaction time, and manual dexterity of fine and complex motor tasks (Banderet and Burse, 1988; Banderet and Shukitt-Hale, 2002; Crowley et al., 1992). For individuals who are already at artificially high physiologic altitudes because of smoking, all these issues are compounded.

With acclimatization, acquired while living at the same altitude or via staging at moderate altitudes, the large cognitive impairments are typically eliminated within one to two days. This has been shown in a number of studies. For example, the large impairment in cognitive function (represented by a code substitution task) that occurs at least during the first few hours for unacclimatized sea-level residents who ascended to 14,000 ft (4,300 m) was eliminated in about 12 hours (Figure 16-4). Also note that there was no cognitive impairment for mountain-area residents who had lived for >21 months at 7,000 ft (2,100 m) prior to their ascent to 14,000 ft (Cymerman et al., 2005; 2006a; 2006b). Unfortunately all of these studies were done on non-smokers and little is known as to if the smoker would be able to acclimatize as quickly.

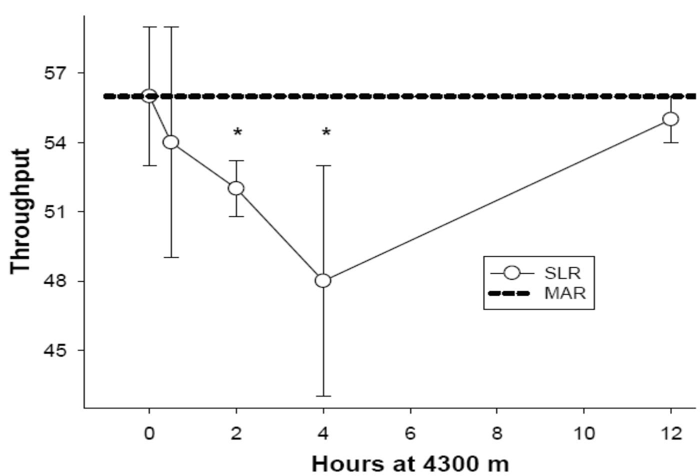


Figure 16-4.Cognitive impairment at altitude (Cymerman et al., 2007).

In 1988, Edwards, Sanders and Price (1988) conducted a study comparing flight school performance in groups of nonsmoking and smoking Army aviation students. The study's intent was to determine whether the effect of smoking enhances or decrements initial flight training performance. Academic and in-flight grades for five phases of IERW classes between January 1984 and November 1986 were extracted from U.S. Army Aviation Community of Excellence, Fort Rucker, AL, (formerly U.S. Army Aviation Center) records and compared to the student's responses to behavior activities on the auxiliary questionnaire portion of the Aviator Epidemiologic Data Register, a comprehensive database collected yearly on every Army aviator by the joint effort of the U.S. Army Aeromedical Research Laboratory (Fort Rucker, AL) and the U.S. Army Aeromedical Activity (Fort Rucker, AL). There were 2,025 student aviators with data sufficiently complete, with the average age of 24.5 years, and with a rank and sex distribution as follows: 96.3% males, 3.7% females; 53.2% commissioned officers, 46.7% warrant officers. Using strict criteria defining smokers and nonsmokers for this study, a 15:85 ratio of smokers to nonsmokers was found (recent quitters and those who smoke less than one pack/day were not included in the analysis). While recognizing that a number of controlled medical studies had determined that smoking is detrimental to overall health, no evidence of a statistically significant relationship was found between smoking behavior and flight school performance.

While not current, the results of a 1986 literature review conducted by the U. S. Army Medical Research and Development Command regarding research into smoking as it related to soldier performance is worth examining (Dyer, 1986)). Research on smoking and other nicotine effects was included in the review. The research reviewed was related to position disclosure in combat; the effects of smoking on physical work capacity and endurance; the effects of smoking on perceptual processes; the effects of smoking on arousal and ability to deal with stress, pain, and fear; smoking-induced hormonal changes; the effects of tobacco deprivation; smoking-disease relationships and their effects on productivity and absenteeism; smoking and abuse of other substances, delinquency, and accidents; and associations between smoking and other factors of potential relevance to soldier performance. Among the main findings, the review disclosed detrimental effects of smoking on physical performance of soldiers, particularly soldiers with several years of tobacco exposure. The review also identified nicotine-related improved performance on vigilance and rapid information processing tasks, including tasks that may be relevant to some soldier tasks. It also showed an abundance of negative behaviors that are correlated with smoking such as drug abuse, delinquency and driving accidents. Research in many areas critical to soldier performance, such as the effects of smoking on dark adaptation and the effects of smoking on testosterone production, showed contradictory results, which the authors argued required additional research for resolution.

In conclusion, the use of tobacco products by the Warfighter is highly discouraged due to detrimental affects on overall health, physical performance and to a greater extent because of the multiple influences that nicotine and CO have on visual perception. Again, it cannot be overemphasized that for the HMD user who has a unique view of his environment and the 4-D battlespace surrounding him, any decrement in his ability to interpret where he is in an operationally harsh combat scenario is critical to his survival and mission accomplishment.

Alcohol

> *Aircrew will not perform aviation duties for a minimum of 12 hours after the last drink consumed and until no residual effects remain* - Army Regulation 40-8 (Department of the Army, 2007*).*

The effects and concerns of the use of alcohol products on Warfighter physical, cognitive and perceptual performance in the HMD environment is an important issue. Alcohol use is fairly ubiquitous in American society, with an estimated 80% of adults imbibing in beer, wine or spirits with a per capita consumption of approximately 25 gallons per year (Orford and Silberman, 2008). It should not be surprising that Warfighters also consume alcohol. In a recent 2005 survey, it was found that alcohol use remains fairly high among military personnel (Table 16-8) (Department of Defense, 2005).

Table 16-8.
Estimates of average daily ounces of ethanol, among entire population and drinkers, by military service. (Department of Defense. 2005)

Population	Service				
	Army	Navy	Marine Corps	Air Force	Total DoD
Entire Population, per Year	1.9 (0.2)[a,b]	1.4 (0.2)[b,c,d]	1.9 (0.1)[a,b]	0.7 (0.1)[a,c,d]	1.4 (0.1)
Drinkers Only, per Year	2.4 (0.3)[a,b]	1.8 (0.2)[b,c,d]	2.3 (0.1)[a,b]	1.0 (0.1)[a,c,d]	1.8 (0.1)
Drinkers Only, per Drinking Day	8.7 (0.7)[a,b]	6.6 (0.3)[b,c,d]	8.9 (0.4)[a,b]	4.5 (0.4)[a,c,d]	7.0 (0.3)

Note: Table entries for average daily ounces of ethanol are average values among military personnel by Service. The standard error of each estimate is presented in parentheses. Pairwise significance tests were conducted between all possible Service combinations (e.g., Army vs. Navy, Navy vs. Marine Corps). Differences that were statistically significant are indicated.

[a]Estimate is significantly different from the Navy at the 95% confidence level.

[b]Estimate is significantly different from the Air Force at the 95% confidence level.

[c]Estimate is significantly different from the Army at the 95% confidence level.

[d]Estimate is significantly different from the Marine Corps at the 95% confidence level.

Source: DoD Survey of Health Related Behaviors Among Active Duty Military Personnel, 2005 (Average Daily Ounces of Ethanol, Q18-Q26 and Q32-Q34).

Twenty-one percent of service members admit to drinking heavily – a statistic the U.S. military hasn't managed to lower in 20 years. Additionally, young Warfighters between 18 and 25 tend to engage in heavy drinking more than their civilian peers. Binge drinking (now more commonly referred to in the scientific literature as heavy episodic drinking) is also at higher levels than for the civilian population (16.6%). The 2005 estimate of binge drinking, defined as five or more alcoholic drinks within a 2-hour period at least once in the past 30 days, is 44.5% for the military. This estimate is not significantly different from the 2002 estimate (41.8%). It should be noted, however, that the rate of binge drinking among college populations (44.8% in 2001) is very similar to the military rate (Wechsler et al., 2002).

Overall health effects

From the DOD standpoint, healthy Warfighters are tantamount to mission success. Retaining these highly trained individuals is a requisite for an all volunteer force. An extensive body of data shows associations between long-term, heavy alcohol intake and a variety of adverse health outcomes, including coronary heart disease, diabetes, cirrhosis, various cancers, hypertension, congestive heart failure, stroke, dementia, Raynaud's phenomenon, and all-causes mortality. Additionally, binge drinking, even among otherwise light drinkers, increases cardiovascular events and mortality. However, light to moderate alcohol consumption (up to 1 drink daily for women and 1 or 2 drinks daily for men) is associated with cardioprotective benefits, whereas increasingly excessive consumption results in proportional worsening of outcomes (O'Keefe, Bybee and Lavie, 2007). Additionally, moderate alcohol consumption, up to 2 drinks per day, has been shown to be significantly protective for ischemic stroke after adjustment for cardiac disease, hypertension, diabetes, current smoking, body mass index, and education (Sacco et al., 1999). Ethanol itself, rather than specific components of various alcoholic beverages, appears to be the major factor in conferring health benefits, providing that the individual is a moderate drinker (O'Keefe, Bybee and Lavie, 2007).

Physical performance effects

Warfighters need to be in top physical condition to be able to survive the harsh operational environments that they encounter. The effects of ethanol vary and can depend on the extent of its consumption and environmental context. Acutely, ethanol consumption is not consistent with operating machinery or firing weapon systems – this would only be exacerbated in an HMD environment where the Warfighter must use his full cognitive and perceptual capabilities to determine his actual position in the 4-D battlespace. This being the case, the DOD prohibits the consumption of alcohol while on duty and during any deployment. Additionally, the various services have strict guidelines for aircrew that restrict them from operating any aircraft for at least 12 hours after consuming ethanol (i.e., Army regulation 40-8 [Department of the Army, 2007]), which is informally known as the "12 hours bottle to throttle" rule. In actuality, the formal regulation states that aircrew will not operate aircraft for 12 hours after consumption and without after-effects, since the effects of a hangover can greatly affect performance. The latter has been well documented in a number of studies. One example is an alcohol study done on military pilots that documented that 14 hours after consuming the alcohol, pilot performance was worse in the hangover condition on virtually all measures (Yesavage and Leirer, 1986). In another example, a study demonstrated that alcohol use among athletes revealed that alcohol has a causative effect in sports-related injuries, with an injury incidence of 54.8% in drinkers, compared with 23.5% (less than half) of non-drinkers. Researchers believe that this is due to the hangover effect of alcohol consumption, which has been shown to reduce athletic performance by 11.4% (O'Brien and Lyons, 2000).

In addition to its acute effects, ethanol can impede physical performance when its consumption is of a chronically abusive nature, i.e., alcoholism. It has been known for some time that individuals diagnosed with alcohol dependence have displayed various degrees of muscle damage and weakness (Martin and Peters, 1985). However other studies have demonstrated that at low doses, the acute effects of ingestion of ethanol on the response to submaximal and maximal exercise resulted in heart rates at rest and during submaximal exercise were higher after ingestion of ethanol, but there was no effect on stroke volume and the circulatory response, oxygen uptake and pulmonary ventilation to maximal work was not affected by ethanol. These findings are in agreement with data from animal experiments suggesting that ethanol in blood concentrations below 200 mg/100 ml has no significant depressive effect on performance of the normal heart (Blomqvist et al., 1970). Table 16-9 provides a comprehensive listing of some of the more commonly documented acute effects on motor skills, strength and power, and aerobic performance (The University Health Center, 2008).

Sensory, perceptual and cognitive effects

From an HMD Warfighter perspective, the ability to think clearly, see well, and react quickly and appropriately are the key requirements to survival and the successful execution of the mission. From a cognitive standpoint, heavy alcohol drinking is acknowledged by a substantial percentage of young adults in the military population, despite the known cognitive demands associated with their endeavors and the cognitive impairments associated with alcohol usage. Researchers have assessed the acute effects of ethanol (0.6 g/kg) on the acquisition of both semantic and figural and noted that ethanol significantly impaired memory acquisition in both domains (Acheson and Swartzwelder, 1998).

Yet another study examined the effects of ethanol on several complex operant behaviors in rats as a human model. Tasks included: temporal response differentiation (TRD) to assess timing behavior; differential reinforcement of low response rates (DRL) to assess timing and response inhibition; incremental repeated acquisition (IRA) to assess learning; conditioned position responding (CPR) to assess auditory, visual, and position discrimination; and progressive ratio (PR) to assess motivation. Ethanol was found to reduce accuracy or percent task completed for the TRD, DRL, and CPR tasks. This experiment demonstrated that ethanol selectively impairs performance on cognitive-behavioral tasks and that these effects can occur at doses that do not affect the subjects' ability to respond (Popke, Allen and Paule, 2000).

Table 16-9.
Acute effects on motor skills, strength and power, and aerobic performance.
(The University Health Center, 2008)

Physical Performance	Ethanol Effects
Motor skills	Low amounts of alcohol (0.02-0.05 grams/deciliter) result in • decreased hand tremors • slowed reaction time • decreased hand-eye coordination Moderate amounts of alcohol (0.06-0.10 grams/deciliter) result in • further slowed reaction time • decreased hand-eye coordination • decreased accuracy and balance • impaired tracking, visual search, recognition and response skills
Strength, power, and short-term performances	Alcohol will not improve muscular work capacity and results in • a decrease in overall performance levels • slowed running and cycling times • weakening of the pumping force of the heart • impaired temperature regulation during exercise • decreased grip strength, decreased jump height, and decreased 200- and 400-m run performance • faster fatigue during high-intensity exercise
Aerobic performance	Adequate hydration is crucial to optimal aerobic performance. The diuretic property of alcohol can result in • dehydration and significantly reduced aerobic performance • impaired 800- and 1500-m run times • increased health risks during prolonged exercise in hot environments

As far as vision, ethanol has been repeatedly shown to cause significant visual perceptual issues and should be of great concern for all HMD Warfighters. For example:

- Consuming alcohol can have short-term negative affects on vision. For a low blood alcohol level, visual performance is less affected by the visual changes than by alteration in brain functions (Quintyn et al., 1999).
- At higher concentrations, such as when the legal blood-alcohol level is reached and surpassed, depth perception and night vision are affected. It becomes impossible to accurately judge how far away objects are when depth perception deteriorates. Vision becomes blurred or doubled since eye muscles lose their precision causing them to be unable to focus on the same object. Alcohol also affects night vision by keeping the pupils from adapting from darkness to light. Alcohol consumption also produces tunnel vision and can make night blindness worse (Department of Transport South Africa, 2007).
- Contrast sensitivity can be reduced, which can prevent an individual from detecting obstacles within the field-of-view (FOV) for some situations. A reduction in contrast sensitivity, when combined with changes in ocular-motor control and attention deficits, also degrades performance (Pearson and Timney, 1998).

- Studies have illustrated that motion parallax (the ability to recover depth from retinal motion generated by observer translation) is important for visual depth perception. Thresholds in a motion parallax task are significantly increased by acute ethanol intoxication (Nawrot, Nordenstrom and Olson, 2004).
- Also there is a higher incidence of blue-yellow color blindness (tritanopia) found when ethanol is consumed. Individuals showed poorer color discrimination in all spectra but with significantly more errors in the blue-yellow versus the red-green color range ($p < 0.005$, $p < 0.01$). Thus, ethanol appears to act as a toxin to inner retinal layers, which could account for the higher incidence of tritanopia found among alcoholics (Russell et al., 1980).

In regards to auditory perception, numerous studies have shown that the acute ingestion of ethanol can cause auditory distraction on visual forced choice reaction time. This suggests that the attention-capturing effects of the deviant sounds were suppressed by ethanol, thus demonstrating a detrimental effect of ethanol on involuntary attention (Teo and Ferguson, 1986).
Additionally, the effects of ethanol on the evoked response potentials evoked by auditory stimuli are to decrease stimulus attention, and stimulus categorization (Jaaskelainen et al., 1996). Finally ethanol has been noted to specifically blunt lower frequencies affecting the mostly 1000 Hertz (Hz), which is the most crucial frequency for speech discrimination (Upile et al., 2007).

In conclusion, the use of ethanol-containing products by the Warfighter is highly discouraged due to detrimental affects on overall health, physical performance and to a greater extent because of the multiple influences that ethanol has on cognition, vision and auditory perception. With the HMD Warfighter this fact must be further emphasized due to the very unique perceptual environment that an HMD presents to the user and the often complex cognitive, visual and auditory processes that our Warfighters must use to interpret what is being presented to them on the HMD as compared to what their actual 4-D environment is. The ability to know where you are in an operationally harsh and complex battlespace is paramount to individual Warfighter survival and mission accomplishment.

Environment (External) Stressors

> Key concept: Normal physiology in abnormal environments will cause HMD related performance impediments unless these environmental effects are identified, considered and mitigated in HMD design.

This section seeks to address the environmental factors that directly or indirectly affect human performance and will thus affect the human-machine interface associated with HMDs. Generally these factors are characteristics of the aviation environment that require unique countermeasure development versus being under the direct control of the Warfighter. Exceptions to this rule are usually related to lessening the impact of a particular environmental stressor as in the example of smoking and hypoxia noted in the text above. Thus, it becomes incumbent upon the HMD designers to be cognizant of these environmental stressors and understand how the Warfighter will perform when exposed to these conditions.

Thermal stress

Hot and cold environments have been shown to have adverse effects on human sensation, perception, and cognition. There is a wealth of scientific information and analysis of human performance measures with respect to physiological and psychological changes that occur as a result of exposure to heat or cold. However, some of the greatest challenges to human performance when operating in climates outside the body's thermoneutral zone are those that result from issues that at first would appear mundane. For example, the sweat that soaks through the helmet liner of a helicopter pilot flying in the Iraqi desert at 120°F (49°C) can make it extremely difficult to keep NVGs positioned correctly for more than about ten minutes at a time before the helmet begins to shift. Similarly, the wearing of a balaclava[10] to keep one's head warm in the mountains of Afghanistan will necessarily change the way that a combat helmet fits. As a result, displays that do not have a wide range of adjustment in multiple planes may not allow for a full FOV. Furthermore, changes in ambient temperature that might arise in going from a heated (or cooled) ready room to a chilled (or sun-baked) cockpit can lead to decreased resolution due to condensation. This section presents information on the ways in which humans respond to thermal stress. While the preponderance of the available scientific knowledge focuses on objective measures of physiological or psychological performance, the reader is encouraged to consider the practical design implications for HMDs that result from operation in both static and dynamic thermal environments.

Overview of human thermoregulation

Human beings are homeotherms – meaning that circadian and seasonal variation in core temperature is maintained within a relatively narrow range about 99°F (37°C) with normal fluctuations being less than 0.6°F (1^{o}C) (Stocks et al., 2004; Wright et al., 2002). In contrast, the temperature of the skin can vary significantly depending upon environmental conditions; this is especially true for the nose, the ears, and the extremities. Human thermoregulation is a complex process that occurs at multiple levels. The thermoregulatory system is comprised of four main components: (1) thermoreceptors located throughout the body; (2) neural pathways mediating information to and from the central nervous system (CNS); (3) the controlling system within the CNS; and (4) the thermoeffector system, which includes autonomic and behavioral responses (Pozos and Danzl, 2001). While humans can survive in a wide range of thermal conditions, the thermoneutral zone (TNZ) for a naked resting body, which is the range of ambient temperature in which thermoregulation is achieved without changes in metabolic heat production or evaporative heat loss, is relatively narrow and falls between 83°F to 86°F (28°C to 30^{o}C) (Faerevk et al., 2001). Within the TNZ, thermal balance is maintained primarily by regulation of skin blood flow (Wright et al., 2002). Once thermal regulatory action goes beyond minor postural or vasomotor control, thermal stress is experienced.

Thermoreception and thermal comfort

The body's core temperature must be maintained at a high level within a very narrow range for human survival, and both core and peripheral temperature sensing systems are required to maintain homeostasis (Stocks et al., 2004). Thermosensitive nerve endings, or thermoreceptors, are located in different areas of the skin and muscle, and throughout the deeper parts of the body to include arteries, internal organs, and the CNS. The peripheral sensors located in the skin and muscles provide the first line of physiological information. These thermoreceptors are either "warm" or "cold" types according to their responses to external stimuli. The determinants affecting the activation of thermoreceptors and the subsequent thermal sensation are: (1) the number of receptors in a specific region, (2) the intensity of the stimulus, (3) the individual's adaptation to temperature, (4) the rate of temperature change, and (5) the size of the area stimulated. Thermal sensation and comfort are related to the thermal state of the body. Skin temperature is a major determinant of thermal comfort; however, the influence of local sensation varies for different parts of the body (Simmons et al., 2008). For example, local cooling of the hands and feet may produce a whole-body sensation of cold that is not related to average skin temperature. It has been suggested that overall thermal sensation and comfort follow the warmest local sensation in a warm environment and the coldest in a cool environment. It must be emphasized, however, that skin temperature cannot be used as a surrogate for

---

[10] A balaclava is a form of headgear covering the whole head, exposing only the face and often only the eyes.

core body temperature due to the centrally mediated physiological responses to thermal stress (Pozos and Danzl, 2001).

Thermoregulation and the CNS

The CNS controls all physiological and behavioral responses to thermal stress. The extreme complexity of the thermoregulatory system necessitates that only a cursory overview will be presented. Incoming signals from the periphery and the deep sensors are processed at multiple levels within the CNS to include the spinal cord. The hypothalamus, an area within the brain, is considered to be the body's thermostat (Pozos and Danzl, 2001). At present, it is not clear which variables (i.e., core temperature, temperature change, body heat content, or rate of heat outflow) are regulated. Furthermore, the establishment of a "set-point" is not well understood; however, it is believed that this point may change temporarily due to factors such as acclimatization, hydration, or fever (Sawka and Pandolf, 2001). Changes detected by the hypothalamus can trigger efferent pathways of the thermoregulatory system through parallel processes of behavioral and physiological responses. Examples of thermally oriented behavioral responses can include the donning or doffing of clothing, seeking shelter, or modifying activity levels. Nearly all physiological systems respond in some way to thermal stress. The systems that are most immediately activated include the cardiovascular system, the musculoskeletal system, and the neuro-endocrine system (Pozos and Danzl, 2001).

Thermal balance

Thermal stress is the nonspecific response of a subject to temperatures that fall outside of the TNZ. The basis of all human thermal stress lies in an energy balance equation which satisfies the continuity requirement for energy exchanged between the body and its surroundings which can be summarized as follows (Parsons, 2003):

$$S = (M - W) - (C + R + E + K) \qquad \text{Equation 16-1}$$

where S = storage of body heat, M = metabolic energy transformation, W = work, C = convective heat transfer, R = radiant heat exchange, E = evaporative heat loss, and K = conductive heat transfer. The maintenance of core temperature requires the continuous elimination of metabolic heat in addition to the compensation for any environmental heat gain or loss. The environmental factors that affect the thermal balance equation are ambient temperature, radiant temperature, air (or water) movement, and humidity. Together with metabolic heat production and clothing, these variables can be used to define human thermal environments (Figure 16-5).

Heat production and loss

At increased activity levels, heat generated from metabolic energy transformation and utilization moves from the core to the skin via tissue conduction and circulatory convection. It must then be dissipated to the environment. Within the TNZ, the body makes minor adjustments via cutaneous vasomotor dilation (to dissipate heat) or constriction (to conserve heat) in order to maintain thermal homeostasis (Faerevik et al., 2001). The body experiences thermal stress when vasomotor control alone cannot maintain thermal balance. To compensate, the thermoregulatory control center in the hypothalamus initiates both physiological and behavioral changes. The primary physiological defense against heat stress is the secretion of sweat. Each liter of evaporated sweat removes 580 kilocalories of heat, and sweat rates may approach two liters per hour during strenuous work in hot environments (Finnoff, 2008). However, high ambient humidity, clothing, and other protective gear can impede the evaporation of sweat thereby negating the potential for heat loss while simultaneously exacerbating dehydration. Additionally, clothing can trap heat and hinder other methods of heat exchange (e.g., radiation, convection, conduction) between the body and the environment.

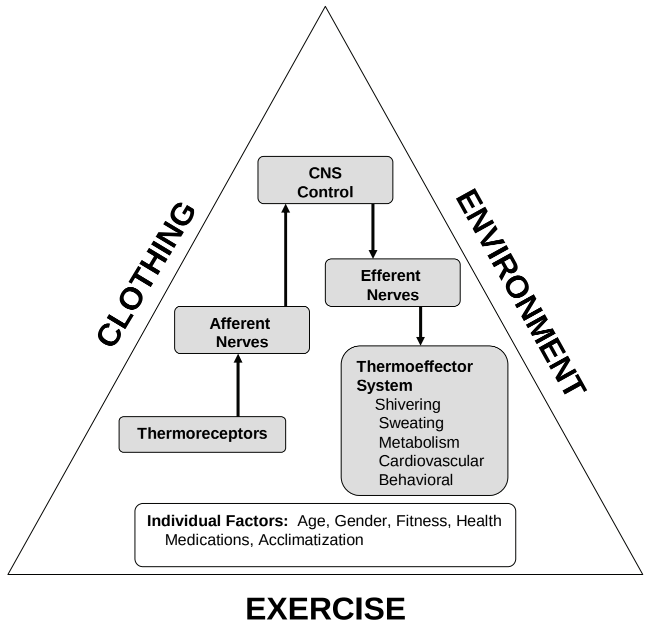


Figure 16-5. Factors affecting thermal balance (Parsons, 2003).

Responses to cold

Vasoconstriction of the superficial blood vessels is an efficient means of reducing heat loss to the environment (Enander and Hygge, 1990). The shell of cooled tissue which includes skin, inactive muscle, and subcutaneous fat provides a layer of insulation for the internal organs. Skin temperatures of the peripheral areas are reduced while the excess blood that is shunted to the inner parts of the body leads to compensatory changes in the cardiovascular system (Stocks et al., 2004). If cooling continues, the hypothalamus sends efferent signals that lead to the involuntary contractile activity of skeletal muscles, or shivering. This increases the metabolic production of heat two to four times above basal levels (Stocks et al., 2004). Simultaneously, behavioral responses to cold are initiated. As with hot environments, excessive or inappropriate clothing may trap heat and lead to sweating which can lead to decreased insulation and enhanced heat loss.

Thermal tolerance

Physiological tolerance to hot or cold environments is a function of both severity and duration of exposure. The core temperature provides the most reliable indicator to predict physical impairment in an environment outside the body's TNZ (Sawka and Pandolf, 2001). Core temperature will continue to rise if evaporative cooling is unable to compensate for the heat gained either from increased metabolic activity or from the environment itself. Humans are better suited to compensate for heat stress, and they are able to tolerate heat stress to a greater extent than cold stress before incapacitation ensues. Heat exhaustion is inevitable when the core temperature goes above 104°F (40°C) (Gonzalez-Alonso et al., 1999). In a cold, dry environment more than one half of heat loss occurs through radiation. Absent appropriate behavioral responses to cold, the minor increase in heat production derived from shivering is often inadequate, and is unsustainable for more than a few hours (Stocks et al., 2004). From a clinical perspective, hypothermia begins as the core temperature falls to 95°F (35°C) or below. There are many common factors that can contribute to the development of uncompensable heat or cold stress. In addition to the more obvious risk factors such as temperature, wind, and humidity, they can include lack of fitness, dehydration, fatigue, and a history of a previous heat or cold injury. One factor that can lead to increased thermal tolerance is acclimatization.

Acclimatization

Acclimatization occurs when prolonged or repeated exposures to an environmental condition lead to significant physiological changes. Heat acclimatization has been shown to greatly improve physical performance, and it leads to greater tolerance for heat exposure (Sawka and Pandolf, 2001). The adaptive changes include increased sweating and earlier onset of sweating, decreased loss of electrolytes through sweat, decreased heart rate, and lower core temperatures. It is theorized that acclimatization to heat changes the thermoregulatory "set point" within the hypothalamus. The majority of the improvement is experienced within the first week of exposure with complete acclimatization by two weeks (Sawka and Pandolf, 2001). The benefits of acclimatization can be partially nullified by fatigue and dehydration, and it is suspected that they are lost shortly after periodic exposure is ended. On the other hand, physiological adaptation to cold is difficult to prove in part due to the behavioral responses invoked by exposure to a cold environment, such as avoidance (Enander and Hygge, 1990). Any adaptation that may occur after repeated exposure to cold is suspected to be relatively minor as compared to those benefits afforded by heat acclimatization.

Clothing and microclimate systems

Protection from environmental extremes can be provided by specialized clothing and systems that can be worn to modify the environment immediately adjacent to the body. Unfortunately, many of the advances in fabrics capable of wicking moisture from the body, thus facilitating heat loss through evaporation, are incompatible with the work environment. More often than not, clothing limits the heat exchange with the surroundings by increasing insulation and inhibiting evaporative heat loss. This can lead to a hot, humid microclimate next to the skin. Even in a cold environment, additional layers of protective clothing may cause significant heat stress, especially when the individual is exposed to a wide range of ambient temperatures over the course of a single duty period. Faerevik et al. (2001) were able to show that standard issue protective clothing worn by aircrew actually shifted the TNZ from to 83°F to 88°F (28°C to 31°C) to a lower range of 50°F to 58°F (10°C to 14°C), and that the clothing hindered evaporative cooling at 65°F (18°C) yet was not sufficiently insulated to prevent shivering at 32°F (0°C). However, their study did not include the wear of protective armor which further inhibits the evaporation of sweat and increases the metabolic cost of physical work. Warfighters operating in uniforms and gear such as shown in Figure 16-6 are at increased risk of uncompensable heat stress especially when exposed to high ambient temperatures. In some cases, the only solution may be the use of an active thermal control system which heats or cools the microclimate within the clothing. Ventilated suits can distribute air over the skin to facilitate evaporation; whereas, liquid cooled garments consisting of interwoven tubing transfer body heat to an external sink through convection.

Immersion

Water is a potent heat sink with a cooling power that far exceeds that of air at the same temperature. Cold water immersion is capable of causing a substantial convective heat loss which can rapidly overcome the body's ability to maintain its core temperature and subsequently lead to uncompensable hypothermia. The rate of heat loss is a function of the water temperature, the water current, metabolic rate, and the body's subcutaneous fat content. Shivering offers much less protection in the water due to increased convective loss with movement (Stocks et al., 2004). In general, the greatest performance decrements can be expected to occur in individuals immersed in cold water or those who remain wet and are exposed to cold air (Hoffman, 2001). Therefore, special consideration must be given to the thermal protection utilized in underwater operations.

Figure 16-6. Typical uniform of a U.S. Army Soldier in Iraq circa 2008.

Psychological aspects of performance under thermal stress

Thermal stress is also capable of inducing changes in psychological performance measures. In fact, some researchers believe that changes in psychological measures will often precede critical changes in physiological status (Johnson and Kobrick, 2001). Thus, monitoring certain aspects of behavior can give an early warning of uncompensable thermal stress. Some of the measures used to assess psychological performance include sensory tasks such as vision or hearing, perceptual tasks which require interpretation of environmental changes such as target discrimination, and cognitive tasks that require reasoning or mathematical calculations. A possible explanation for observed decrements in performance under thermal stress is that changes in temperature somehow limit human attention leading to a narrowing of focus in sensory, perceptual, and cognitive abilities thereby forcing task prioritization of finite mental capabilities (Hancock, 1986).

Unfortunately, this field of research is replete with many conflicting reports of the effects of thermal stress on performance (Pilcher et al., 2002). This is likely a result of the diversity in experimental conditions used, the specific performance tasks measured, the severity of the thermal stress, and the duration of the exposure found between studies. Human behavior is influenced by several factors to include the environment, the person, the task, and the situation; and within these are many sub-variables as illustrated in Figure 16-7 (Johnson and Kobrick, 2001).

Thus variations in tasks, conditions, and performance measures can lead to dissimilar outcomes. For example, the simple concept of standardizing the quantification of ambient temperature can become complex quickly when variables such as air velocity, relative humidity, and radiant heat are considered. Furthermore, establishing a relevant measure of the thermal stress induced can be problematic (Enander and Hygge, 1990). Taking the results of multiple studies of the effects of thermal stress on psychological measures, one can conclude that performance is negatively affected by exposure to either heat or cold especially when there is dynamic change in the core body temperature (Hancock, 1986; Pilcher et al., 2002; Wright et al., 2002).

Psychological performance changes in the heat and cold

There are many more studies that have examined the effects of *increased* temperatures on human psychological performance in contrast to the fewer that have studied the effects of *decreased* temperatures. This is probably due to the increased likelihood of exposure to heat stress either in the workplace or through increased body temperature as a result of exercise. As previously mentioned, the results are often difficult to compare. Measures of sensation have found that tactile discrimination is greatest in moderate temperatures (Johnson and Kobrick, 2001), but sensitivity decreases as temperature decreases with measurable impairment at hand skin temperatures

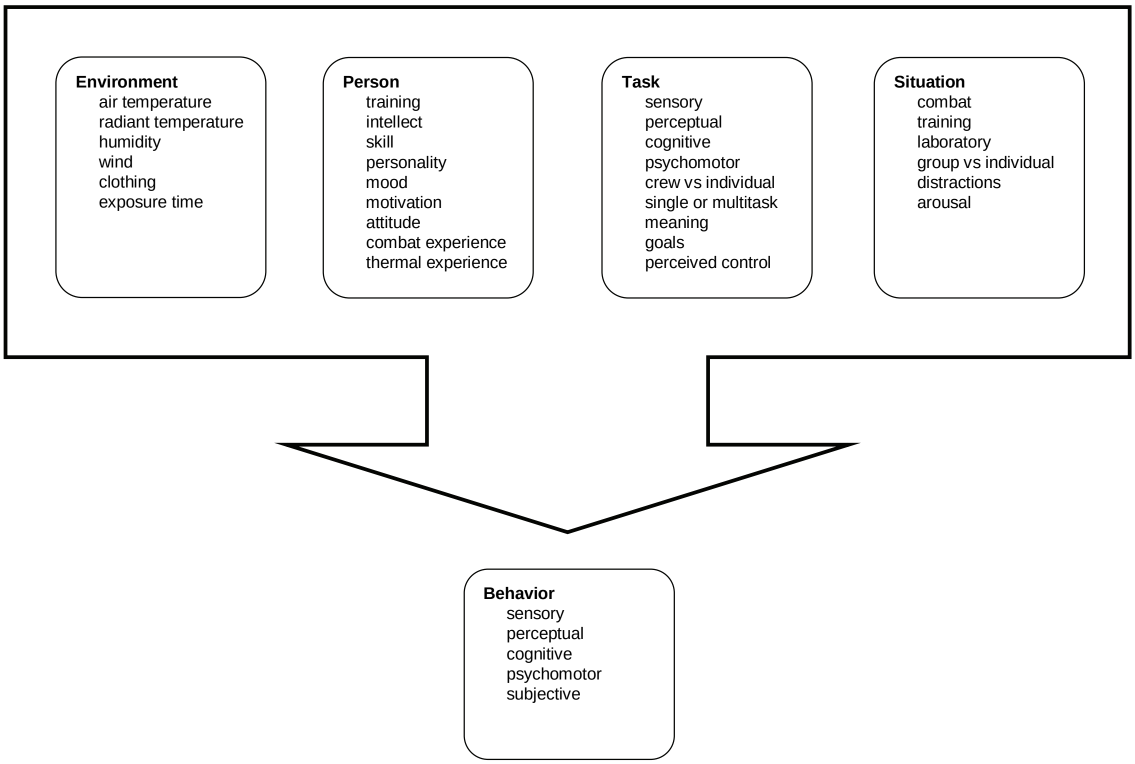


Figure 16-7. Basic psychological model in which behavior is a function of the environment, the person, the task, and the situation (Johnson and Kobrick, 2001).

below 68°F (20°C) (Enander and Hygge, 1990; Hoffman, 2001). The effects of heat on visual acuity and contrast sensitivity were indeterminate (Johnson and Kobrick, 2001); however, there is a valid concern that heat can indirectly interfere with vision through sweat dripping in the eyes and by shifting head gear when the hair, scalp, and helmet interface become wet.

Perception is often measured by subject response times to either visual or auditory stimuli. Higher core body temperatures tend to produce faster response times but lead to more mistakes (Simmons et al., 2008). Mild static hyperthermia improves performance in simple reaction time as long as the body is able to compensate for the thermal stress; whereas, complex reaction times become slower (Enander and Hygge, 1990; Grether, 1973; Hancock, 1986;). Similarly, cold exposure increases the error rate in complex reaction tasks (Enander and Hygge, 1990; Thomas et al., 1989).

There are a wide variety of measures of cognitive functioning which may include target tracking, vigilance, and memory tasks. Vigilance is a complex behavior that consists of attention, alertness, cognition, judgment, and decision making. Visual and auditory vigilance tasks are impaired at elevated ambient temperatures (Johnson and Kobrick, 2001). Psychomotor performance tasks such as tracking, and cognitive functions that require some type of judgment or reasoning are impaired above 85°F (30°C) wet bulb globe temperature (WBGT) (Grether, 1973; Johnson and Kobrick, 2001). Similarly, visual motor tracking is significantly impaired by exposure to temperatures below 10°F (12°C); however, the exposure time does not appear to have a significant effect until the core body temperature begins to drop (Giesbrecht et al., 1993; Hoffman, 2001). In studies of cold water immersion, the speed of complex mental tasks is reduced by one-half at core body temperature below 95°F (35°C), and memory registration is reduced by almost three quarters of what would have been retained under normal physiological conditions (Coleshaw et al., 1983).

Interestingly, physiological adaptation to a controlled environment, or acclimation, does not improve cognitive performance in the heat (Curely and Hawkins, 1983). Nor does acclimation improve the disruption to sleep patterns and sleep effectiveness that is seen during exposure to hot conditions (Johnson and Kobrick, 2001). While acclimation does not appear to be helpful, studies of differential body cooling indicate that head cooling[11] can modulate the detrimental effects of elevated skin and core body temperatures on comfort and alertness (Nunneley et al., 1982; Simmons et al., 2008). This suggests that psychological performance has some correlation with subjective assessments of comfort in both hot and cold environments (Hoffman, 2001; Nunneley et al., 1982).

Conclusions on thermal stress

Thermal stress will compromise cognition, but the level of deterioration is dependent upon the severity of the stress, the resultant core temperature, and the complexity of the task (Giesbrecht et al., 1993; Simmons, 2008; Tikuisis and Keefe, 2007). In hot environments, performance is degraded when thermal homeostasis is disturbed; that is, performance suffers when there is a dynamic change in core body temperature (Hancock, 1986). It is not solely the ambient temperature that affects performance, but the combination of ambient temperature and exposure time that is sufficient to change the core body temperature (Johnson and Kobrick, 2001). In hot and cold environments, cognitive performance decrements will occur when thermal stress becomes uncompensable (Giesbrecht et al., 1993; Simmons et al., 2008).

Many psychological performance measures follow an inverted U-shaped distribution with decreased performance at both higher and lower ambient temperatures (Hoffman, 2001; Pilcher et al., 2002). The nearer the ambient temperature is to the body's TNZ, the less effect it has on performance in both hot and cold environments (Figure 16-8). The range of optimal temperatures may vary by specific task as different types of brain function appear to have different zones of thermal sensitivity with respect to performance (Pilcher et al., 2002; Wright et al., 2002). In general, simple behavioral performance measures show some improvement when core body temperature is statically elevated within a compensable zone; however, the more complex the task the more likely it will deteriorate with exposure to heat or cold (Enander and Hygge, 1990; Wright, 2002).

Implications for HMD design

The head represents only 10% of body surface area, but its potential for heat transfer is amplified because of the extensive vasculature. HMDs that heat the head will tend to increase core body temperature; whereas, HMDs that incorporate some type of cooling device can reduce both thermal discomfort and core temperature thereby improving psychological performance (Nunneley et al., 1982; Simmons et al., 2008). The design of any equipment intended for use in even moderately hot or cold temperatures should take into account the expected

---

[11] Recent work also has been directed to heat extraction via the hands (e.g., Grahn, Cao and Heller, 2005).

performance decrement in many sensory, perceptual, and cognitive tasks with the understanding that complex tasks, to include vigilance, will suffer impairment to a greater extent.

Of a more practical nature, designers should be cognizant of the indirect effects of hot and cold environments upon human sensation, perception, and cognition. Military uniforms and equipment impose added heat load due to the decreased effectiveness of evaporative cooling as illustrated in Figure 16-6.

HMDs that restrict airflow to the head can exacerbate thermal stress and lead to increased unevaporated sweat that can either drip into the eyes and reduce vision or cause the helmet to shift out of position on the head. In cold environments, users of HMDs can be expected to wear additional clothing to include some type of thermal protection for the head and face as illustrated in Figure 16-9.

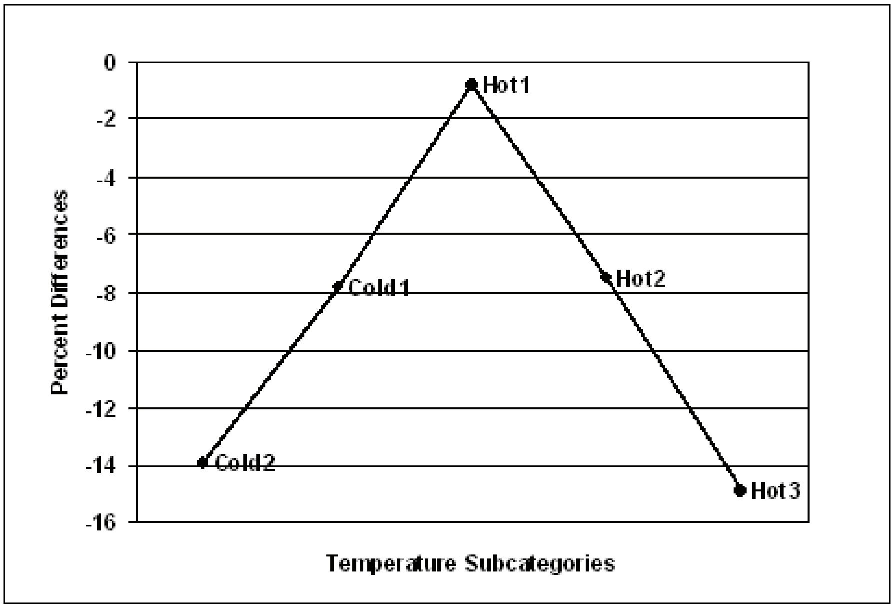


Figure 16-8. Mean percent difference in performance between the neutral temperature groups and the five temperature subcategories. Cold2 = <50°F (10°C); Cold1 = 50°F to 64.9°F (10°C to 18.3°C); Hot1 = 70°F to 79.9°F (21.1°C to 26.6°C) WBGT: Hot2 = 80°F to 89.9°F (26.7°C to 32.2°C) WBGT; Hot3 = 90°F (32.3°C) or greater WBGT (Pilcher et al., 2002).

Figure 16-9. Cold weather gear worn to protect the head and face.

These additional layers necessitate a wide range of image adjustment in order to maximize FOV. Furthermore, gloves worn in both cold and hot environments can restrict manual dexterity and limit tactile sensitivity. Thus, tasks that require fine motor control with the fingers may require additional lighting to allow for visual monitoring in dark or low light conditions. For all military purposes, HMD controls should be designed for operational use with gloves on, and metal surfaces should be coated with rubber to decrease heat conductivity (see later section on *User adjustments* in this chapter).

Thermal stress is a major environmental concern for military operations. Any additional equipment intended to be worn by the Warfighter must be designed and evaluated to minimize (or if possible to reduce) the thermal load.

Altitude threats and hypoxia

"The higher, the fewer…"
-unknown RAF Apprentice
Halton, Buckinghamshire, UK

Warfighter-interface designers must be cognizant of operating environments in order to help foresee and potentially mitigate performance decrements that may Warfighters may incur at high operational altitudes. In the past, when discussing altitude issues and human factors, platform specific categories were reasonable to consider. We still have those full time operators that can be divided into orbital, suborbital, high altitude reconnaissance, fast jet, transport and rotary wing; each with over water and over land caveats. Threats associated with changes in altitude are routinely encountered by pilots and aircrew but are also increasingly experienced by dismounted Warfighters in mountainous operations. In aviation, these dangers exist in both pressurized and unpressurized cabins – especially since a pressurized cabin can become unpressurized in an emergency situation.

However, as we discuss altitude as an operationally relevant factor, the reader should keep in mind that with increasing integration of ground, naval and air assets and the development of joint warfighting doctrine, a single combatant may find himself in multiple environments in quick succession as he executes a mission. Consider a hypothetical example of a 12-hour ingress flight at 50,000 ft and a high altitude parachuting to water at sea level near the objective. After reaching the coast, the Warfighter is required to make an overland trek to 14,000 ft (4,300 m) to reach the mission site. Recovery occurs via helicopter over a 20,000 ft (6,100 m) mountain range and via transport aircraft standing by at a friendly neighboring base. Can HMD devices be designed to be compatible with the wide range of altitude extremes that this Warfighter will experience?

In general, humans live in a gaseous envelope with a set mixture of nitrogen (78%), oxygen (21%), inert gases (1%), carbon dioxide (0.03%), and water vapor (varies) known colloquially as "air." The percentages of these components remain stable as one ascends through the troposphere,[12] but the barometric pressure decreases with distance above the Earth's surface, in an approximately exponential manner, meaning that the partial pressure of available oxygen decrease as well (Dalton's Gas Law).[13] This can lead to both hypoxia and decompression related problems like trapped gas disorders, barotraumas (Boyle's Gas Law)[14] and decompression illness (Henry's Gas Law).[15] Other physical properties that change predictably include temperature (about 2°C per 1000 ft) (Figure 16-10), decreasing humidity, increasing ionization and radiation exposure. HMD designers also should have a

---

[12] The *troposphere* is the lowest level of the Earth's atmosphere and is considered to extend from the surface of the Earth to an average height of 7 miles (11 kilometers).

[13] Dalton's law (also called Dalton's law of partial pressures) states that the total pressure exerted by a gaseous mixture is equal to the sum of the partial pressures of each individual component in the mixture.

[14] Boyle's law describes the inversely proportional relationship between the absolute pressure and volume of a gas, if the temperature is kept constant within a closed system.

[15] Henry's law states that at a constant temperature, the amount of a given gas dissolved in a given type and volume of liquid is directly proportional to the partial pressure of that gas in equilibrium with that liquid.

general understanding of altitude countermeasures so as to help minimize conflict and interaction with these life support devices.

Hypoxia

Hypoxia can be defined as the lack of adequate tissue oxygen available to support the body's normal metabolism. In healthy individuals, this is usually due to a lack of adequate inspired oxygen and can eventually lead to in-

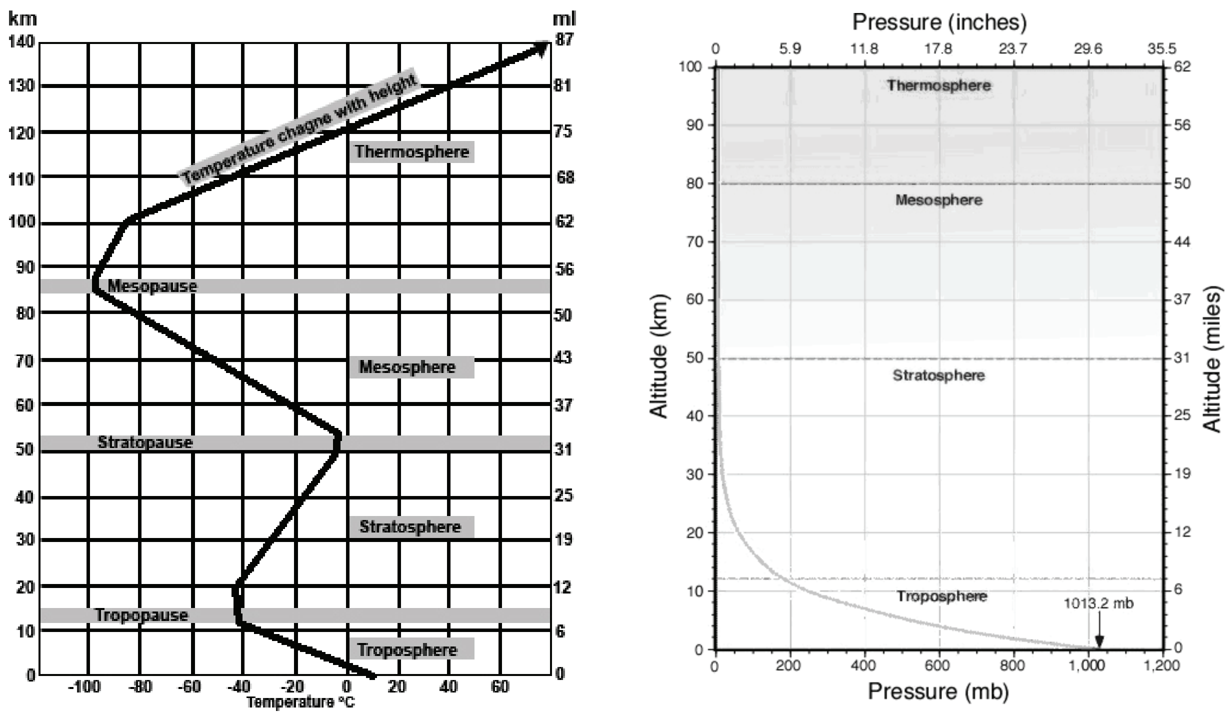


Figure 16-10. Atmospheric temperature (left) and pressure (right) changes as a function of altitude

sufficient energy production, cell dysfunction and, if left unchecked, cell death. It should be noted that the neurologic system is most sensitive to hypoxia, and that even though the brain comprises less than 5% of bodyweight it consumes almost 20% of the oxygen acquired by the circulatory system. This means that higher cognitive functions, as well as vision, are more acutely affected by lack of oxygen. Furthermore, because the brain is affected directly and the symptoms can develop insidiously, the untrained Warfighter is usually unable to detect that he is becoming hypoxic. To further complicate the early detection of hypoxia, individuals vary widely in their initial hypoxia symptom complexes making physiology training in altitude chambers very important for any Warfighters, especially pilots and aircrew, who routinely may operate at high altitudes. Traditionally, extended exposure to cabin altitudes above 10,000 to 12,500 ft (3,050 to 3,800 m) have required supplemental oxygen but recently subtle operationally and physiologically significant effects of hypoxia have been noted at lower altitudes as well (Smith, 2005). Finally, the overall physiologic state of the Warfighter influences the onset of hypoxic symptoms since other factors like alcohol, smoking, general health and life stressors can lower the individual resilience.

Physiologists recognize four types of hypoxia which are categorized based on the cause for the lack of oxygen available to cellular metabolism (Dehart and Davis, 2002). *Hypemic hypoxia* occurs when the body's ability to transport the available oxygen is impaired and may occur due to lack of adequate red blood cells (i.e., bleeding, genetic abnormalities), carbon monoxide poisoning or other chemical poisoning (i.e. sulfa drugs, nitrites). This is analogous to a delivery company not having enough trucks on the road due to fleet shortages or maintenance. In *stagnant hypoxia*, there is a reduction in either regional or whole body blood flow, thereby lessening the delivery of oxygen to tissue. Stagnant hypoxia occurs in heart failure, excessive G-forces, blood clots, tourniquets or strokes. In this case, the delivery company has adequate trucks on the road, but they are stuck in traffic jams or waiting on road construction. *Histotoxic hypoxia* refers to the tissue's inability to accept oxygen that is offered by the circulatory system. It can be caused by metabolic toxins like alcohol, cyanide and some narcotics. By analogy, the delivery company has brought the package to your house, but no one is there to sign for it or accept it, so it goes back on the truck to attempt redelivery the next day.

Hypoxic hypoxia is the most familiar to the aviation community and refers to a lack of available oxygen in the inspired air. As humans ascend into lower atmospheric pressure, the partial pressure of oxygen also decreases meaning that, on a per breath basis, less oxygen molecules are available for the lungs to transfer into the blood stream. For instance, the pressure at 18,000 ft (5,500 m) is only half the normal ground level 760 mm of Hg, so only about half as much oxygen is available to the lungs. Fortunately, due to the design of hemoglobin, the oxygen carrying proteins in red blood cells, there is actually only a 75% to 80% decrease in available oxygen in the blood stream. The non-linear relationship between oxygen saturation and ambient oxygen tension is illustrated in the oxygen dissociation curve (Figure 16-11). Other than altitude, other causes of *hypoxic hypoxia* include asthma, drowning and respiratory arrest. Unlike the previous examples, the delivery company finally has everything in order – enough trucks, clear roads and customers ready to accept packages – but the distribution center has gone on strike leaving partially or completely empty trucks to drive the routes.

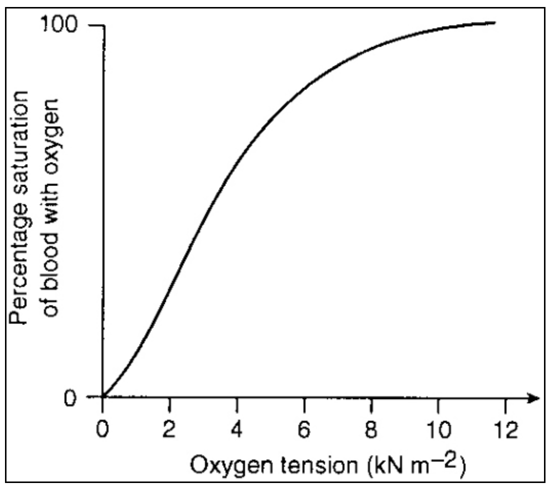


Figure 16-11. Oxygen dissociation curve.

What are the effects of hypoxia? As stated earlier, decrements in higher cognitive functions and visual effects are the most important and obviously observable consequences. Although the exact symptom complexes vary by individual and have an insidious onset, there are generally accepted decrements in functioning that have been divided into four stages: Indifferent, compensatory, disturbance and critical (Table 16-10). Time of useful consciousness (TUC) is defined as the amount of time an individual is able to perform efficiently in a hypoxic environment; after which, the individual is no longer capable of taking proper corrective and protective action. More importantly, it should not be viewed as the time to loss of consciousness.

In studies looking for cognitive deficits, participants exhibited disturbances of memory functions and delayed recall but graphic and semantic memory showed less frequent errors. Simple arithmetical errors, perseveration, impaired visual-motor coordination (jerkiness, illegible writing and poor reproduction of geometric figures) and thought blockage with an inability to complete written tasks were also described. Neuromuscular symptoms of tremor or twitching were noted and later subjects lapsed into a semi-conscious state, mentally switched off and became unresponsive, with eyes open and head upright. Also reported was a feeling of being unable to execute commands, or feelings of euphoria or carelessness, which cast doubt on the ability of these individuals to respond to an emergency (Ernsting, Nicholson and Rainford, 1999; Westerman, 2004).

Table 16-10.
Generalized Hypoxic Symptoms by Stage and Altitude
(Adapted from Reinhart, 1992)

Altitude (ft)	Indifferent Stage O_2 Sat: 90-98% TUC: unlimited	Compensatory Stage O_2 Sat: 80-90% TUC: >30 min	Disturbance Stage O_2 Sat: 70-80% TUC: 20-30 min	Critical Stage O_2 Sat: <70% TUC: 4-10 min
25000				Circulatory Failure Convulsions Death
20000			Impaired speech Impaired muscle control Impaired coordination Worsening flight control Impaired visual acuity	
15000		Drowsiness Poor judgment Impaired efficiency Decreased coordination Impaired color vision		
10000				
5000	Decreased night vision			

Example of actual cognitive degradation and euphoric indifference experienced by the author during physiologic training: “Once ascent in the hypobaric chamber had been completed, every other student was asked to remove their oxygen mask long enough to develop their distinct hypoxia symptoms. After several minutes of performing cognitive tasks on a clipboard, I stopped and stared straight ahead. When the instructor ordered me to replace my mask, I fully and cheerfully acknowledged the instruction but did not execute the activity. After three requests, the fully oxygenated student next to me was asked to replace my mask and set it to 100% oxygen allowing me to regain full control of my mental facilities.”

Of importance to HMD display design engineers is that decreases in night vision occur at relatively low altitudes during the indifferent stage and that mild hypoxia also impairs some color vision at lower light levels (Connolly et al., 2008). In fact, some research has shown that dark adaptation occurs more rapidly at ground level when supplemental oxygen is supplied to the subject suggesting that some of the tissue in our eyes may normally be somewhat hypoxic (Wangsa-Wirawan and Linsenmeier, 2003). Lower oxygen availability in the cabin air also directly affects the corneas since they get most of their oxygen via diffusion rather than via the blood supply. This becomes even more critical when the ergonomics of an HMD, as in the AH-64 Apache helicopter, requires the use of corrective contact lenses rather than spectacles. Later stages of hypoxia also can lead to visual convergence issues and diplopia, which may of import when considering HMD placement and focal ranges.

Potentially compounding weight and center-of-gravity issues, hypoxia can lead to early and potentially painful neck muscle fatigue over time. This becomes especially acute in higher G-environments and for aircrew that have higher metabolic demands for oxygen due to movement around the cabin or frequent head motion (Smith, 2006).

Consider medical personnel attending to patients en route, loadmasters aboard cargo helicopter or disembarking rescue/recovery personnel that will return to the aircraft after significant exertion.

As discussed earlier, other stressors also affect the overall effect of hypoxic hypoxia. For instance, carbon monoxide, a major component of tobacco smoke, has a 20 times greater affinity for blood than oxygen; given a choice between carbon monoxide and oxygen, the red blood cell will choose the carbon monoxide. This compounds the hypoxic hypoxia with hypemic hypoxia, accelerating symptom development. Some researchers estimate that a regular smoker has a physiologic altitude of 3,000 to 8,000 ft (900 to 2450 m)[16] while at sea level, and he/she will usually display a higher red blood cell count as a result of the chronic hypoxia.

The link between longer mission durations in modern military operations, with its associated extended relative immobility and potential hypoxia, and potentially fatal blood clot formation is the subject of continuing debate. According to some research, prolonged civilian air travel may increase blood levels of clotting factors, particularly among individuals with risk factors for blood clotting disorders. However, it remains unclear whether the reduced cabin pressure and oxygen tension in an unpressurized or partially pressurized aircraft interior creates an increased risk compared to extended immobility on the ground (Toff et al., 2006). From an ergonomic standpoint, however, it would be reasonable to allow aircrew some mobility in their seats and design HMDs that would not further impede the ability to move around, thereby lessening the chance of clot formation.

High altitude illness bears mentioning when discussing altitude effects, but rarely occurs in aircrew. There is however the potential that high-altitude illness may occur if a base station for flight operations is established at altitudes greater than 6000 to 8000 ft (1800 to 2450 m). This syndrome is made up of several symptom complexes including fluid build-up in the lungs called high altitude pulmonary edema (HAPE), brain swelling called high altitude cerebral edema (HACE), retinal hemorrhages, and extremity swelling. High altitude illness generally occurs 1 to 4 days after arrival and there is a tendency for previously acclimatized personnel returning to altitude to fall victim more frequently. The rate of ascent, the altitude attained, the amount of physical activity at high altitude, colder temperatures and individual susceptibility contribute to the incidence and severity of this condition (Hackett, Rennie and Levine, 1976).

Dysbarism

Dysbarism or barotrauma refers to medical problems that arise from the pressure differences between areas of the body and the environment and is a particular concern for aircrew and divers. All involve gases trapped in an enclosed area where pressure cannot equalize during ascent or descent causing pain. This can involve actual air spaces that have become blocked off, referred to as “trapped gas disorders,” or be due to the introduction of bubbles in spaces where there should be none, as would be the case is decompression illnesses.

Trapped gas disorders are directly related to Boyle’s law (as the pressure increases, the volume decreases and vice versa). As atmospheric pressure decreases, this volume change in trapped gas-filled spaces and organs within your body accounts for the distortion and damage to surrounding tissues leading to pain and occasionally bleeding. Examples can include external ear squeeze, middle ear squeeze, inner ear barotraumas, sinus squeeze, tooth squeeze and gastric squeeze. Rapid decompression at altitude can lead to pulmonary barotrauma (pulmonary over-pressurization syndrome, or burst lung) if aircrew fail to expel air from the lungs during the event. Externally attached devices, depending on their fit and design, have been known to cause problems as well, e.g., mask squeeze and G-suit squeeze, and should be considered in the development of HMDs.

In decompression illness, gas (mostly nitrogen) that was previously dissolved in solution within body fluids forms bubbles inside tissue causing severe pain and neurologic disorders (if the bubbles form in the brain). This process, also known as the “bends” is usually a risk for divers and is explained by Henry’s law (more gas will be dissolved in a liquid when the gas is pressurized) interacting with the drop in water pressure when surfacing too

---

[16] Throughout this chapter, several different altitude equivalents are given to describe the effects of a number of cigarettes smoked within a certain periods of time. As these values are quoted from different sources, they are subject to variation.

quickly or stay at depth too long. However, it can also be a threat to personnel going to lower atmospheric pressure, resulting in a similar phenomenon.

Other altitude effects

Space borne radiation particles and those ionized particles generated by the collision of these particles with atoms in our atmosphere are collectively referred to as galactic cosmic radiation. In general, our atmosphere combined with the Earth's magnetic field adequately protects us from cosmic radiation, but in certain flight regimes and during disturbances in the Sun's atmosphere, an increased exposure to these charged particles can occur. At higher altitudes there are higher levels of cosmic radiation and aircraft flying at altitudes of 30,000 to 40,000 ft (9,100 to 12,200 m) receive about 100 times greater exposure the ground. The Earth's magnetic field deflects many radiation particles that would otherwise enter the atmosphere and this shielding is most effective at the equator but diminishes at higher latitudes. At the poles, cosmic radiation is about twice as high as at the equator since the magnetic field is essentially nonexistent at the poles. Occasionally, unpredictable solar particle events (SPE), which are ejections of a large mount of charged particles, can lead to sudden increases in radiation levels in the atmosphere. The sun also protects us from cosmic radiation since its heliosphere extends well beyond Neptune and the solar winds intercept many potentially harmful particles. This protection does, however, vary slightly with the 11-year solar cycle; when the sun is at solar max with the greatest number of sunspots, it affords the greatest protection (Friedberg et al., 2000).

Currently, there is no conclusive evidence that ionizing radiation results in significant adverse health effects on long-haul civilian aircrew. However, based on exposure risks and internationally adopted standards, pregnant crewmembers are the only personnel that are at greater risk and these risks are to the fetus only (Chee, Braby and Conroy, 2000). On the other hand, flying higher and lengthier military missions (i.e., reconnaissance, global-reach strategic bombing) might expose aircrew to damaging radiation. Of interest to the HMD developer community is that this exposure potential suggests that the devices must be tested against ionizing radiation failure modes and may require additional hardening for this type of interference.

Lower temperatures at higher altitudes present an obvious thermoregulatory problem for aircrew. At 30,000 ft (9,100 m), for example, the ambient temperature is in the region of -40°C (-40°F). Generally, this is mitigated by having enclosed cockpits with adequate heating systems but in some cases the cabin may not be conditioned or flights may occur in winter conditions. The HMD designer should be aware of this potential hazard and consider effects of vapor precipitation on displays as well as discomfort that could be associated with cold surfaces touching the head.

Also associated with flight into colder and thinner atmosphere is a drop in relative humidity. Although there is no evidence that significant physiologic effects occur due to extended exposure to dry air, substantial subjective complaints of thirst, due to dry mucous membranes, and eye irritation, due to decreased tear film, do occur. In particular, contact lens wearers suffer in these environments and may even lose their lenses or develop corneal ulceration (Dennis, Apsey and Ivan. 1993). This has implications for the design of HMDs that would not allow spectacle use as discussed earlier.

Countermeasures integration

Obviously the various threats that are described here will require some type of countermeasure to lessen the impact of the effect on aircrew and help assure mission success with no loss of life. The mitigation of these threats can take various forms: changes in training, tactics, mission planning, aircraft design and capability and additional life support equipment. Device manufacturers should consider especially the use of oxygen delivery systems and aircrew equipment intended for climatic control. As mention previously in the section on temperature extremes (see *Thermal stress*), this type of integration is key to being able to maintain a useful fit of the device.

An audio device or display that cannot be properly fitted around an oxygen delivery system is equally problematic. As an example, the U.S. Army has recently developed novel oxygen delivery systems like the Portable Helicopter Oxygen Delivery System (PHODS) which utilizes a nasal canulla on a single arm from the helmet (Figure 16-12). As other devices are added to the helmet, minute displacements may cause improper oxygen delivery and lead to subtle hypoxic decrements. More familiar systems of oxygen delivery, each presenting unique human factors issues, include very simple pipe stem systems, various on-demand mask systems and positive pressure systems. Positive pressure systems are more frequently seen in unpressurized fast jet cockpits and usually used above 32,000 ft. It is beyond this altitude that breathing even 100% oxygen is not adequate to properly oxygenate the body due to the low atmospheric pressure and the system begins delivery of oxygen under pressure to avoid arterial desaturation (Ohshund 1991). This type of system must obviously fit very snugly to the face and should not be interfered with by any additional display devices.

Figure 16-12. PHODS device installed on Gentex HGU-56/P US Army helicopter helmet.

Noise

Noise in an acoustical form[17] is defined as any unpleasant or unwanted sound that is unintentionally added to a desired sound. Noise generally is thought of as an auditory problem, e.g., noise can block, distort, or produce a change in the meaning of a communication (see Chapter 13, *Auditory Conflicts and Illusions*). However, noise exposure can have a range of auditory and non-auditory consequences, producing both physiological and psychological effects. At low levels, noise is a distracter and can degrade performance; at high levels, noise can produce temporary and long-term hearing loss. Generally, problems due to noise include hearing loss, stress, high blood pressure, sleep loss and fatigue, distraction, and lost productivity. Considerable effort has been undertaken to develop and provide noise protection to the Warfighter operating in the military environment. However, protective devices often are considered by individuals to be detrimental to the tasks at hand and are not employed as frequently or as effectively as is needed to prevent physiological damage.

Characteristics of auditory noise

Noise is sound, albeit undesirable sound. Noise is a series of changes in sound pressure levels that are created by a source, transmitted via some medium (usually air), and collected and interpreted by the auditory system. Therefore, noise can be defined by all of the general characteristics of a sound wave: frequency, phase, amplitude,

[17] The term noise is often categorized as audio or electronic. Audio noise is used in the music industry to describe unwanted sounds encountered in audio, recording and broadcast systems. Electronic noise refers to unwanted additions to signals in electronic circuits; shot and thermal noise are two of the most common types.

and wave velocity. The response of the human auditory system to the presence of noise will depend on these characteristics (see Chapter 9, *Auditory Function*). Like other sounds, noise is generally complex, consisting of a combination of frequencies.

It is not just the intensity that determines whether noise is hazardous. The duration of exposure is also important. The terms *steady-state* and *impulse* often are used to characterize the duration of sound and hence noise. Steady state noise has negligible fluctuations of level within the period of observation; it is continuous, by definition lasts more then one second, and includes such sources as vacuum cleaners, hair dryers, electrical power generators, lawn mowers and idling engines**.** Impulse noise is very intense and of short duration, usually less than a second. Examples include backfires from motor vehicles, sonic booms and weapons fire. Impulse noise is more difficult to characterize than steady-state noise (Hamernik and Hsueh, 1991).

Noise levels are measured in decibels (dB), a logarithmic unit of measurement that expresses the magnitude of a physical quantity (e.g., sound intensity) relative to a reference level. As the dB expresses a ratio of two quantities with the same unit, it is a dimensionless unit.[18] A normal conversation is at approximately 65 dB SPL;[19] shouting typically can be around 80 dB. To take into account the fact that the human ear has different sensitivities to different frequencies, the intensity of noise is usually measured in A-weighted decibels (dB(A)).[20] The hazard threshold for steady-state noise is 85 dB(A) – the sound of some power lawn mowers – and for impulse noise, 140 dB(P)[21] – the sound made by some machine guns. The noise level for discomfort is 120 dB(A) – the sound of a jet airplane on takeoff, as heard by someone 164 ft (50 m) away. Pain may occur when sounds are louder than 130 dB(A) – the sound of a live rock music concert (Rash, 2006).

Environmental noise

Noise is ubiquitous, with an almost limitless number of natural and artificial sources. The Warfighter can expect to encounter constant and high levels of noise; this is especially true in and around both ground and air military vehicles. In addition, while combat may increase the frequency and intensity of noise exposures, training activities produce equally dangerous noise environments. Paakkonen and Lehtomaki (2005) report that noise episodes in combat and training exercises reaching a peak level of 180 dB. Average noise exposure levels for military exercises were measured outside the ear at approximately 95 to 97 dB and in the ear canal at 82 to 85 dB. Peak levels of 110 to 120 dB for military trainers were measured in the ear canal during the use of small-bore weapons.

The U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM), maintains a list of noise levels for common U.S. Army equipment on its web site (U.S. Army Center for Health Promotion and Preventive Medicine, 2008). Steady-state noise for selected vehicles, aircraft and power equipment are presented in Table 16-11; impulse noise values for selected armament and munitions are presented in Table 16-12. (Note: Impulse noise levels are measured in a “peak”-related decibel form known as dB(P)).

While not the nosiest military vehicles, helicopters do present the most complex noise environments. Noise spectra for helicopters are comprised of aerodynamically-induced noise from the main and tail rotor assemblies, main gearbox and various transmission chains (Rainford and Gradwell, 2006). In addition to the steady-state noise produced by the engines, mechanical components and airflow, there is impulse noise, commonly called blade slap, caused by the blade-vortex interaction (Schmitz, 1995; Widnall, 1971).

---

[18] The decibel (dB) is one-tenth of a bel (B). The dB is used in a variety science and engineering disciplines.

[19] Sound pressure level (SPL) is the term most often used in measuring the magnitude of sound. It is a relative quantity in that it is the ratio between the actual sound pressure and a fixed reference pressure.

[20] A-weighting began with the work of Fletcher and Munson (1933) that resulted in a set of equal-loudness contours corrected for the normal sensitivity profile of human hearing.

[21] Decibel Peak dB(P) is used for peak sound level equal to 20 times the common logarithm of the ratio of the highest instantaneous sound pressure to a reference pressure of 20 micropascals. It is used in the measurement of impulse noise.

Table 16-11.
Measured steady-state noise levels for selected U.S. Army Equipment.
(U.S. Army Center for Health Promotion and Preventive Medicine, 2008)

Model	Name/Condition	Location	Speed km/hr(mph)	Sound Level dB(A)
M996, M997	HMMWV* mini and maxi ambulance, at two-thirds payload	Patient areas	up to 88(55)	Less than 85
M113A3 family	Armored personnel carrier A3 version		Idle 16(10) 32(20) 48(30) 63(40)	85-92 106 109 114 118
M1A2, M1, M1A1	Abrams tank	In vehicle	Idle 16(10) 48(30) 63(40)	93 108 114 117
M2A2	Bradley fighting vehicle	In vehicle	Idle 16(10) 32(20) 61(38)	74-95 110 115 115
MEP-802A	5 kW Tactical quiet generator	Operator panel	Rated load	80
CH-47D	Chinook helicopter	Cockpit	Cruise speed	102.5
UH-60A	Black Hawk helicopter	Pilot Copilot	Cruise speed	106 106
OH-58D	Kiowa helicopter	Right seat Left seat	Cruise speed	101.6 100.3
AH-64	Apache helicopter	Pilot Copilot	Cruise speed	104 101.3

* *High Mobility Multi-wheeled Vehicles (HMMWV)*

An interesting paradox with engine-driven vehicles is that sounds can be perceived as both noise and important information simultaneously. For example, aircraft pilots can tell much about the operating conditions of engines by their sounds. Pilots learn to identify specific engine problems by the change in their sounds.

Naval ships face a common high-noise problem due to the need to conduct operations in closely confined areas. The U.S. Navy Center, Norfolk, VA, reports that older ships were not designed using the noise reduction techniques employed on modern ships. Even for newer ships, a number of high-noise areas cannot be avoided. On aircraft carrier flight decks, flight operations are confined to a 4.5-acre (18,200 m^2) area as compared to land-based flight operations that are normally conducted on 10,000 acres (40.5 square kilometers). Noise levels on the flight deck can exceed 145 dB(A). Noise sources on the flight deck include aircraft engines, catapults, and arresting gear equipment. Below the flight deck is the gallery deck in which approximately 1400 sailors live and work. The high noise levels directly above adversely impact most of the gallery deck. Gallery deck noise levels, often in excess of 100 dB(A), can have the effect of reducing cognitive skill levels and cause miscommunication problems, both frequently identified as causes of fatal accidents (U.S. Navy Safety Center, 2008).

Aside from aircraft-related noise, all ships have noise associated with the ship propulsion, i.e., ship propeller excitation on the ship structure. The excited structure then re-radiates as airborne noise. Ventilation systems are often a significant source of shipboard noise. Because of space constraints, air ducts used aboard ships are often very small and have sharp curves and bends. This results in air moving through the ducts at very high velocities, causing noise and vibration in the ventilation system. Fans can also project noise throughout the ventilation system if they are poorly mounted, not properly isolated from air ducts, or are the wrong size. Finally, noise may be generated at the air duct outlets that distribute air in the work environment if proper design parameters are not followed. Poor ship design can provide transmission paths (e.g., through ventilation ducts) for noise to travel from the noisy machinery spaces to berthing accommodations and workspaces (U.S. Navy Safety Center, 2008).

Table 16-12.
Measured impulse noise levels for selected U.S. Army Armament and munitions.
(U.S. Army Center for Health Promotion and Preventive Medicine, 2008)

Model	Name	Location	Sound Level dB(P)
M16A2	5.56-mm rifle	Shooter	157
M9	9-mm pistol	Shooter	157
M2	0.50 caliber machine gun fired from a HMMWV*	Gunner	153
M26	Grenade	At 50 ft	164.3
M72A3	Light antitank weapon (LAW)	Gunner	182
M109A5/6	Paladin, 155mm self propelled howitzer firing M4A2 zone 7 charge	In fighting compartment, hatches open except driver's	166.1
M29A1	81 mm mortar, M374A3 round with charge 4	1 m from the muzzle, 0.9 m above ground, 135° azimuth	178.8

* *High Mobility Multi-wheeled Vehicles (HMMWV)*

Physiological effects of noise

Noise can produce a host of physiological effects, such as headache, fatigue, nausea and insomnia. The major physiological effect of noise exposure is hearing loss. Noise-induced hearing loss (NIHL) can be either temporary or permanent. A temporary hearing loss is a brief shift in the auditory threshold that occurs after a relatively short exposure to excessive noise (more than 90 dB). Fortunately, for such loss, normal hearing recovers fairly quickly after the noise stops. However, if the noise level is sufficient to damage the tiny hairs in the cochlea – the part of the inner ear that is responsible for transforming sound waves into the electrical signals that go to the brain – the threshold shift can be irreversible, resulting in permanent partial or total hearing loss (Rash, 2006). Research has determined that individuals exposed to steady state sound levels of 85 dB(A) for an 8-hour period or longer are in danger of losing their hearing. Likewise, exposure to impulse noises of 140 dB(P) can result in hearing loss (U.S. Army Center for Health Promotion and Preventive Medicine, 2008).

In a review of noise-induced health effects, Soames-Job and Hatfield (2000) cite the following studies and findings:

- Occupational studies having demonstrated that noise exposure contributes to hearing loss (Morata, 1999; Ward, 1993), and may have a detrimental impact on cardiovascular health (Talbott et al., 1996).

- Noise has also been found to impair performance both in occupational (Smith, 1989) and educational settings (Haines et al., 1998; Hygge, Evans and Bulinger, 1998).
- Community surveys demonstrate negative reactions (Fields, 1994; Hatfield and Job, 1998; Job, 1988) and sleep disturbance (Griefahn, 1992; Griefahn et al., 1998; Ohrstrom, Bjorkman and Rylander, 1990; Pearsons et al., 1995) resulting from noise exposure.
- Noise associated with entertainment (e.g. loud music) has been found to have deleterious effects on hearing (Axelsson and Prasher, 1999).
- The effects of aircraft noise on children's blood pressure are uncertain (Cohen et al., 1980; Morrell et al., 1998).
- Although suggestive of a greater prevalence of psychiatric illness amongst residents of high noise areas, the evidence is inconclusive (Abey-Wickrama et al., 1969; Jenkins, Tarnopolsky and Hand, 1981; Kryter, 1990).

In a study of audiometric data of 54,057 Navy enlisted personnel in the Navy and Marine Corps Hearing Conservation Program database from 1995 to 1999, Bohnker at al. (2002) compared threshold shift patterns with historical literature. The data suggest that 82% of the population did not display significant threshold shift (STS) on the *annual* and *termination* audiograms, which increased to 94% after the follow-up examinations. Compared with historical data, STS rates were significantly lower for the most junior enlisted personnel (E1-E3) but not significantly different for more senior enlisted personnel. STS rates were found not to appear to correlate with expected high- and low-noise exposure Navy enlisted occupations.

Performance effects of noise

Hygge (2003) states that a number of studies on the effects of noise on cognition and human performance report the general finding that the task being performed has to be complex and cognitively demanding in order to be negatively affected by noise (e.g., Smith, 1989; 1992). Tasks that are simple and repetitive are unaffected by noise, and if the task is boring, simple enough, or well learned, that noise may even improve performance. Thus, a search for noise sensitive tasks must focus on tasks having a moderate or greater level of complexity and are demanding on cognitive resources. He also states that noise effects on cognition is a fairly covered area in psychological noise research, citing a number of studies that have compared the relative impacts on attention, reading, memory and learning (Cohen et al., 1986; Evans and Hygge, 2002; Evans and Lepore, 1993).

Many of the recent studies on the performance effects of noise have been conducted on children and are based on exposure to aircraft and traffic noise. One reason for this is that groups of children (mostly school age) serve as convenient and less confounded samples. While children are known to have a greater susceptibility and hence show a greater response to noise, findings of these studies often are extrapolated to adults. A representative study (Stansfeld et al., 2005) reported a linear association between exposure to (external) aircraft noise and impaired reading comprehension and recognition memory, and between exposure to road traffic noise and episodic memory (in terms of information and conceptual recall). Results also showed non-linear and linear associations between aircraft and road traffic noise, respectively; annoyance also showed a linear association with road traffic noise. Neither aircraft noise nor road traffic noise was found to have affected sustained attention, self-reported health, or mental health.

In studies involving adults, high background noise levels (>90 dB(A)) typically are found to reduce the quality of performance. A number of studies have demonstrated that noise hinders performance on cognitive tasks involving vigilance, decision-making, and memory (Broadbent, 1971; Salas, Driskell, and Hughes, 1996; Smith, 1989). However, these studies typically involved artificially-generated noises in artificial settings, and exposure was usually short-term (i.e., hours). However, in an investigation of the effects of background noise over a 70-hour period on cognitive performance of astronauts on the International Space Station showed "little to no effect of noise on reasoning, perceptual decision-making, memory vigilance, mood, or subjective indices of fatigue" (Smith et al., 2003).

In a noise study more relevant to the issues of displays, Choi (1983) investigated whether noise intensity and display orientation had any effect on short-term memory task. Results showed that continuous white noise[22] at intensity levels of 30, 85, and 105 dB had no effect on the short-term memory task.

More specifically to HMDs, noise (with the exception of the example of pilots using engine noise to monitor engine performance) is to be attenuated by a helmet to which the HMD is integrated or attached. Therefore, HMD systems are expected to provide an acceptable level of sound (noise) protection.

Noise protection

Warfighters generally wear protective head gear or helmets. Depending on the application, helmets can provide a combination of impact, penetration, sun, windblast and noise protection. Maintaining the necessary hearing protection for the military noise environments, while providing high performance voice communications, is a goal of the HMD designer. Historically, the quality of voice communication has been reduced as noise protection has been emphasized. An example of this is the use of earplugs which block both unwanted sound (noise) and wanted sound (communication voice). Two newer technologies that overcome this problem are the Communication Enhancement and Protection System (CEPS) and active noise reduction (ANR). (See Chapter 5, *Audio Helmet-Mounted Displays*.)

The CEPS is a system designed to control the sound level that arrives at the ear and provide the user with dual radio communications. An expanding foam earplug attenuates ambient sounds that enter the occluded ear canal. The system integrates highly sensitive microphones, rapid response micro-circuitry that inputs the sound to the ear through the miniature earphone of the communications earplug (CEP) that is attached to the expanding foam earplug. The user can control the volume of the signal reaching the ear by contact switches. The device was designed to provide enhanced sound detection capability and localization in "recon" or "watch" modes; enhanced face-to-face communication for night, Mission Oriented Protective Posture (MOPP) or military operations on urban terrain (MOUT) operations; and two-way radio communications in stealth mode. It provides protection for both hazardous impulse and continuous noise environments and rapid cut-off and recovery protection for weapons firing (Gordon and Houtsma, 2008; Mozo and Murphy, 1998).

ANR, first conceived in the 1930s and refined in the 1950s, did not become prevalent in aviation until the 1990s (Tennyson, 2001). In conventional ANR headsets, the frequency and amplitude of the sound inside the headset cavity are measured by a small microphone, and a 180° out-of-phase copy is produced and fed back into the headset. The result is that the two signals superimpose and cancel each other. This out of phase canceling technique is very effective for low frequencies, below 800 Hz, but is generally ineffective for higher frequencies. In some designs, the ANR device actually increases the noise level inside the ear cup in the region of 1000 Hz. Total hearing protection consists of the passive protection provided by the ear cup and the ANR component provided by the electronic system. Studies show ANR does improve speech intelligibility when worn alone, but both hearing protection and speech intelligibility are degraded when worn with ancillary equipment such as spectacles or chemical-biological mask (Gower and Casali, 1994; Mozo and Murphy, 1997).

An interesting challenge for noise protection has been the introduction of inflatable restraints (airbags) into rotary-wing aircraft (Crowley and Dalgard, 2000). In the civilian community, the effectiveness of airbags in reducing deaths in automobile accidents is well known. However, studies have documented incidents of hearing loss associated with airbag deployment (Huelke et al., 1999; Morris and Borja, 1998). The U.S. Army has studied the use of airbags in helicopters as early as 1991 (Alem et al, 1991a; 1991b; Shanahan, Shannon and Bruckart, 1993). A Cockpit Airbag System (CABS) has been developed for the UH-60A/L Black Hawk and OH-58D Kiowa helicopters. To support airbag fielding, the U.S. Army Aeromedical Research Laboratory, Fort Rucker,

[22] White noise is defined as random noise that has uniform power spectral density at every frequency in the range of interest.

AL, has conducted tests to determine the risks to crewmembers and passengers associated with exposure to high impulse noise levels expected during an inadvertent system deployment (Brozoski et al., 2000). A series of 21 airbag deployment tests were conducted in a static UH-60A helicopter. Peak sound pressure levels ranged from 134 dB to 161 dB. Levels at pilot, copilot, and gunner stations exceeded 140 dB during all 21 deployments. Levels in the passenger compartment exceeded 140 dB during 9 of the 21 deployments. Army policy requires the aircrew in the UH-60 helicopter to wear helmets that provide hearing protection or a combination of helmet and earplugs. Passengers are required to wear protective earplugs or muffs or a combination of muffs and earplugs. This level of hearing protection also meets the requirements for protection against high impulse noise levels created by the deployment of airbags. Therefore, if the required hearing protective devices are worn, the potential of inadvertent deployment of the CABS in the UH-60 helicopter has been determined not to pose an additional risk to the hearing of crew and passengers (Ahroon, Gordon and Brozoski, 2002).

The noise protection provided to many Warfighters is a given, as the need to wear a helmet is integral to the Warfighter's mission. However, many military personnel perform tasks in high-noise work environments where helmets, with their inherent protection,[23] are not employed. In these situations, less sophisticated but equally effective hearing protection devices are made available (e.g., earplugs and earmuffs). For these individuals, the command structure must institute and enforce a policy of requiring effective use of such hearing protection devices. However, Abel (2008) has documented user concerns of hearing protection interfering with detection and localization of auditory target warnings and perception of orders. In addition, users frequently complain that devices were often incompatible with other gear and difficult to fit.

Vibration

Modern-day work is more mechanized than in the past (Kjellberg, 1990). This is equally true in both civilian and military environments. A consequence of this mechanization is increased human exposure to *vibration*. For the moment, vibration will be defined as a *to and fro* motion about a point of equilibrium. For a discussion of human exposure, vibration can be categorized as either *localized* or *whole body* (Mansfield, 2005)

Localized vibration, also referred to as hand-arm vibration (HAV), is most associated with the use of various types of vibrating pneumatic, electrical, hydraulic, and gasoline powered hand-tools. However, such hand or hand-arm transmitted localized vibration can be equally associated with more mundane and common actions, e.g., holding onto a steering wheel. As is obvious from the name, HAV is coupled almost exclusively via the hand-arm combination.

Whole-body vibration (WBV) affects the whole of the exposed individual – all parts from head to toe. Most WBV is related to riding in vehicles, e.g., trucks, buses, and fork-lifts in the civilian community; and tanks, helicopters, personnel-carriers, and boats in the military community. WBV is transmitted via seats, backrests, or through the floor, coupling through the buttocks, back or feet.

Human effects due to vibration depend on whether the exposure is acute (i.e., having a rapid onset and short duration) or prolonged and usually are grouped into three categories: physiological effects (including the very common occurrence of motion sickness), psychological effects, and performance effects.

In the civilian world, WBV has been studied extensively by researchers in the field of occupational medicine. While strong correlations between WBV and long-term physiological consequences have been established, it has been difficult to separate these effects from those associated with straining during heavy lifting and poor posture (Cardinale and Pope, 2003; Hulshof and Veldhuizen van Zanten, 1987; Seidel and Heidel, 1986). Low-back pain is a common complaint after exposure to WBV and has been shown to be a major cause of disability in the population under the age of 45 years and has been linked to WBV exposure encountered in some industrial

[23] Noise protection in aviation flight helmets has been a long-pursued development goal. However, most U.S. Army Warfighters wear the Army Combat Helmet (ACH), which has no added hearing protection.

settings (Cardinale and Pope, 2003). Additional suspected health effects due to prolonged exposure include hemorrhoids, hernias, digestive disorders, and urinary problems (Hedge, 2008).

In the military environment, low-back pain has been a long-standing health problem for helicopter pilots and ground-vehicle drivers exposed to prolonged WBV. Seat design, sitting posture and vehicular vibration are often identified as high-risk factors (Bongers et al., 1990; Ensign et al., 2000; Pelham et al., 2005; Wasserman, 2003). Vibration has been a problem in aviation ever since aircraft were fitted with engines (Lam, 2003). As reciprocating engines became commonplace, multiple nodes of vibration developed in the airframes, with intensity and effects varying by location. In helicopters, however, vibration is aircraft-wide, affecting all crewmembers. As in the civilian industrial community, while the medical impact of vibration has not been proven, it generally is accepted that there is a distinct relationship between the presence of vibration and the chronic low back pain often experienced by helicopter pilots and military vehicle drivers.

In addition to physiological effects, vibration can lead to psychological effects such as discomfort and annoyance. Lam (2003) has emphasized that while vibration causes chronic and acute fatigue in operational aircrew and leads to chronic back pain, its significant effects on operational performance must not be overlooked. For example, the utility of sights and vision-enhancing devices (e.g., HMDs) is degraded in the presence of severe vibration.

The physics of vibration

Before expanding the various effects of vibration on the human in detail, it is necessary to improve on the initial superficial definition of vibration as a “to and fro motion” about a point of equilibrium. This will be accomplished using a more rigorous description of vibration from the point-of-view of physics. In physics, the phenomenon of vibration is a subset of oscillatory motion. Oscillations can be several types, e.g., electrical, electromagnetic, mechanical, electro-mechanical, optical, biological, and chemical. The term *vibration* usually is reserved for mechanical oscillations, i.e., to describe a mechanical movement that oscillates about a fixed point (Figure 16-13). Classical examples include a tuning fork, playground swing, oscillating spring, and simple pendulum. Within the context of this chapter’s discussion of adverse operational factors in the military environment, examples include helicopters and motorized vehicles. As a mechanical form of oscillations, vibrations are propagated via a mechanical coupling.

In the example of Figure 16-13, imagine that a piece of colored chalk is attached to the oscillating mass and the chalk is in contact with a long sheet of paper being pulled along as the mass goes through its oscillation. The result will be a simple waveform like that shown in Figure 16-14. As the mass on the spring moves through its up and down motion, the chalk will trace multiple complete motion paths or *cycles*. The number of cycles per unit time defines the *frequency* of the vibration. The standard unit for expressing frequency is the Hertz (Hz), defined as one cycle per second (cps).

Along with frequency and amplitude, a third characteristic is needed to fully define a waveform – *acceleration*. The speed of a vibrating object varies from zero to a maximum during each cycle. It moves fastest as it passes through its stationary position towards its maximum displacement (amplitude). It slows down as it approaches the full amplitude, where it momentarily stops and then moves in the opposite direction passing again through the equilibrium position toward the other maximum displacement position. Speed is expressed in units of distance per unit time (e.g., meters per second [m/s] or feet per second [ft/s]). Acceleration is a measure of how quickly the vibrating object’s speed changes with time. Acceleration is expressed in units of meters per second per second (or meters per second squared [m/s^2]). The magnitude of the acceleration changes from zero (at the equilibrium point) to a maximum (at full amplitude) during each cycle; it increases as the vibrating object moves further from its normal equilibrium position.

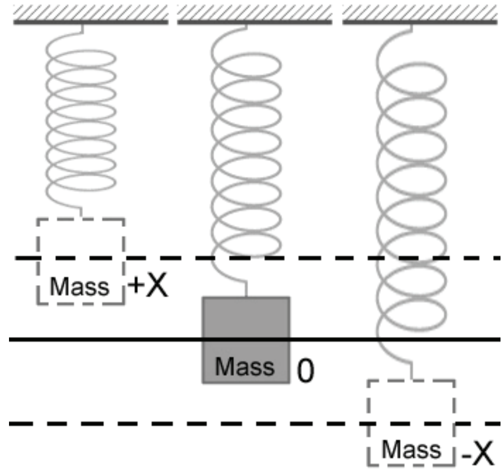


Figure 16-13. Vibration as oscillation about a fixed point, as demonstrated by an oscillating spring.

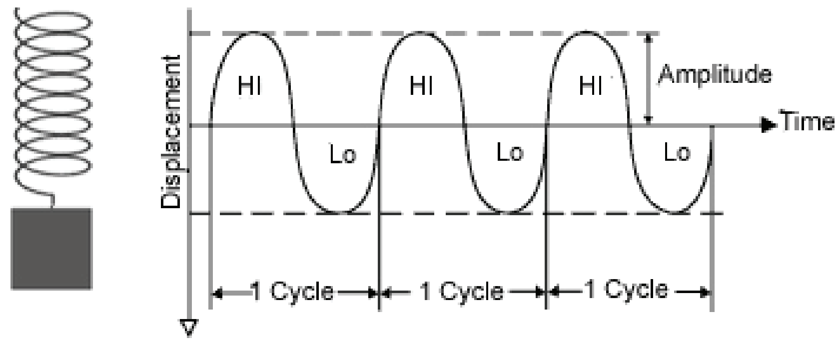


Figure 16-14. The waveform resulting from the motion of a mass attached to an oscillating spring.

Acceleration is the characteristic most frequently used to quantify vibration levels[24] and is measured via a sensor/transducer known as an accelerometer. The piezoelectric accelerometer is the most popular class of these devices; other types are based on piezoresistive, capacitive, and servo transducer technologies.

The vibration waveform presented above is a simple one with only one frequency and amplitude present. It is said to be sinusoidal (i.e., having the form of a sine wave). In the real world, most vibrations are complex, consisting of multiple frequencies and amplitudes.

Complex waveforms (vibrations) are the result of several forcing frequencies occurring at the same time (Figure 16-15). In this case, the resulting vibration will be a summation of the vibration at each frequency. Under these conditions the resulting waveform of the vibration will not be a sinusoid, and may be very complex.

Before leaving the physics discussion of vibration, one special topic stills needs to be introduced – *resonance*. Nearly all objects, when hit or struck, will vibrate, and tend to vibrate at one or more particular frequencies, which depend on the composition of the object, its size, structure, weight and shape. These frequencies of natural vibration are called the *resonant frequencies*. A vibrating object in contact with a second object transfers the maximum amount of energy to the second object when the first object vibrates at the second object's resonant frequencies. At these frequencies, even small oscillating driving forces can produce large amplitude vibrations,

[24] Vibration magnitude is generally measured in terms of the acceleration of the oscillations, rather than the velocity or displacement between peak-to-peak movements. The preferred International System (S.I.) unit for vibration acceleration magnitude is meters-per-second-per-second (m/s^2), and measurements are often expressed as root-mean-squared (rms) values rather than peak values.

because the system stores vibration energy. This phenomenon is called resonance. The resonant frequency of the human body is approximately 4 to 5 Hz.

A classic aviation example of resonance is ground resonance with helicopters having fully-articulated rotor systems. When a helicopter is resting on the ground with its rotor spinning, a condition called ground resonance can develop. This is a destructive harmonic vibration caused by a dynamic reaction of the rotor blades to the lateral motion of the helicopter. The helicopter can be destroyed by this resonance in periods as short as minutes.

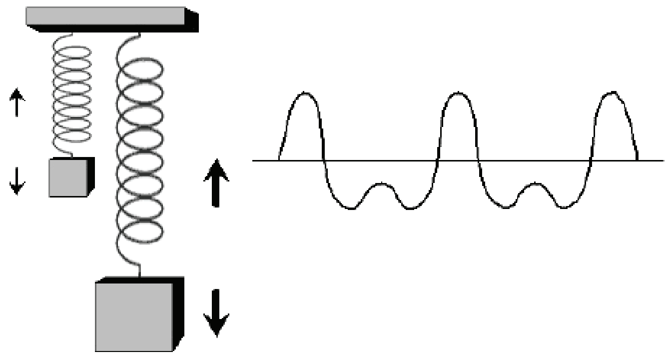


Figure 16-15. A complex waveform resulting from the summation of two sources at two different frequencies and amplitudes.

Human WBV generally occurs in three axes: fore-to-aft (x-axis), lateral (y-axis) and vertical (z-axis). Rotational vibration about the x-, y- and z-axes (called roll, pitch and yaw, respectively) may also occur (Nakashima and Cheung, 2006). The vibration frequency range that is considered important for health, comfort and perception is 0.5 to 80 Hz (International Organization for Standardization-ISO, 1997).[25] Human WBV resonance occurs in the vertical (up-down) direction at frequencies from 4 to 8 Hz. In the lateral and fore-to-aft directions, WBV resonance occurs at 1 to 2 Hz. It is at these resonant frequencies that humans are most vulnerable. The occupational vibration standards attempt to define and compensate for these potentially hazardous human resonant frequencies (Wasserman, 2003).

Physiological effects of noise

The physiological (i.e., health) effects of HAV and WBV are distinctly different, as might be expected based on differing vibration exposure patterns and pathways into the human body (Wasserman, 2003). For this discussion, we will focus on the effects of WBV, and readers are directed to both classical and modern treatises on HAV and its effects (Hamilton, 1918; Pelmear and Wasswerman, 1998; Wasserman, Taylor and Behrens, 1982).

A considerable body of scientific literature exists on the effects of human exposure to WBV. It has been shown that these effects can be either short-term or long-term. Short-term effects include annoyance, discomfort, fatigue, motion sickness, a temporary shift in hearing threshold, reduced motor control, and impaired vision. In addition, although cause and effect has not been proven, long-term and repeated exposure to WBV has been linked to chronic back pain. The degrees to which these effects are manifested depend largely on the characteristics of the vibration and include the frequency, magnitude and duration of exposure; other factors include posture and seating station design (where applicable).

[25] The range of frequencies that is most often associated with whole-body vibration is still a point of conjecture. While the 1997 ISO standard cites 0.5 to 80 Hz, Griffin (1990) uses an approximate range of 0.5 to 100 Hz.

Discomfort

Discomfort and fatigue are the lesser of the physiological effects of WBV. Vibration factors impacting these effects include exposure duration, magnitude, and frequency. Intermittent and random vibration can have a wakening effect, but continuous exposure can lead to increased fatigue or drowsiness. A few studies have shown a possible link between long-term exposure to low-frequency (3 Hz) vibration and fatigue (Mabbott et al., 2001). As not enough data have been collected to establish meaningful exposure limits, ISO 2631-1 (1997) states, "There is no conclusive evidence to support a universal time dependence of vibration effects on comfort."

Nakashima (2004) suggests that "it is intuitive that an increase in vibration magnitude will lead to increased discomfort." However, the same magnitude of vibration will not produce the same level of discomfort at all frequencies. Studies of the combined effects of vibration magnitude and frequency on discomfort have found that the growth in sensation, ψ, with increasing vibration magnitude, Φ, has been found to agree approximately with Stevens' Power Law, given by

$$\psi = kn\ \Phi$$ Equation 16-2[26]

where k is a constant that depends on the system of units, and n is a frequency dependent growth function.

As stated previously, vibration is an ever-present condition in the aviation community – for the pilot, aircrew and ground crew. Aircraft ground and maintenance crew are exposed to noise levels that are sufficient to induce WBV (Smith, 2002). Whether on the ground or in the air, for frequencies between 100 and 1,000 Hz, a 120 dB noise signal will cause tissue vibration. Below 100 Hz, the airborne vibration can cause movement in the body cavities and air-filled or gas-filled spaces; this can induce symptoms such as nausea, coughing, headache and fatigue (cited in Smith, 2002).

Auditory effects

Research into the effects of WBV on auditory functions is sparse, most being retrospective studies that may identify associations but not cause and effect. The reason for this is that such studies would expose human subjects to unacceptable vibration levels. However, Nakashima (2004) does present a summary of the limited research to investigate temporary hearing loss, or temporary threshold shift, due to vibration:

- Early studies by Temkin (1927), Pinter (1973), and Pyykko et al. (1981) suggested that low-frequency hearing loss was intensified in workers who were exposed to both noise and vibration (cited by Hamernik, Ahroon and Davis, 1989; Nakashima, 2004).
- Okada et al. (1972) studied the effect of noise and vibration, both separately and in combination. They found that 5 Hz vibration with an acceleration of 5 m/s^2 produced a threshold shift of more than 7 dB at 1 kHz and 4 kHz after a 1-hour exposure. The 5-Hz vibration is significant because it is approximately equal to the resonance frequency of the human body. Other vibration frequencies (2, 10 and 20 Hz) caused smaller amounts of threshold shift. A greater threshold shift was reported for exposure to vibration in combination with noise than for exposure to noise alone.
- Hamernik et al. (1989), upon a review of the literature, hypothesized that vibration "may potentiate the effects of noise and may thus increase the risk of hearing loss in a variety of exposure situations." However, studies reviewed that involved humans were limited to low levels of exposure and the

[26] This equation is derived from the more conventional form of Stevens' Power Law (1975) expressed as $\psi = k\Phi^B$, *where* ψ is the magnitude of the sensation, Φ is the intensity of the stimulus, B is a characteristic of any given stimulation continuum and indicates how rapidly the magnitude of the sensation (ψ) grows as the stimulus intensity (Φ) increases, and k is a proportionality constant that depends on the type of stimulus and units used.

reported effects measured were relatively small. Animal studies reviewed also shown an enhanced noise-induced hearing loss in the presence of vibration, but the scope of these studies were limited. They reported their own animal studies in which chinchilla were tested using a 30-Hz/3G root-mean-square (rms) and a 20-Hz/1.3-G rms cage vibration separately and in combination with continuous noise (a 95-dB, 0.5-kHz octave band) and impact noise (113--, 119-, or 125-dB peak SPL) exposure paradigms. All exposures had a 5-day duration. Temporary and permanent threshold shifts were measured using evoked potentials, and sensory cell loss was measured using surface preparation histology. The results obtained from some of the noise/vibration paradigms showed that such exposures can alter some of the dependent measures of hearing. This effect was found to be statistically significant only for the stronger vibration exposure conditions and was evident primarily in the extent of the outer hair cell losses and in the shape of the permanent threshold shift (PTS) audiogram.

- Other studies have also reported temporary threshold shift after prolonged exposure to 5-Hz vibration, which is at the resonance frequency of the human body (see review by Griffin, 1990).

Back pain

One of the most commonly reported physiological effects of WBV is back pain. Back pain is a common complaint for industrial vehicle operators and passengers as well as for military personnel (in ground vehicles, most fixed-wing aircraft, and rotary-wing aircraft [helicopters]). Teschke et al. (1999) identifies a number of confounding factors in trying to establish cause and effect between vibration and back symptoms: age, physical condition and working posture. It has been suggested that repeated and long-term exposure can lead to serious back problems such as herniated discs or premature degeneration of the spinal vertebrae (in Nakashima (2004). The human spine has a resonance frequency of approximately 5 Hz, and the frequencies at which vibration is most effectively coupled to the spine are 4.5 to 5.5 Hz and 9.4 to 13.1 Hz.

Back pain frequently is reported by helicopter aircrew. The pain is most likely to be felt by the pilot in-flight and has been attributed to both the vibration of the seat and poor posture. Seat cushions are the only devices that mitigate the direct link between the aircraft and the pilot's body. In addition, pilots often must assume a forward-bending posture in order to achieve maximum visibility and precise control, which places increased pressure on the intervertebral disc. Motivated by complaints of fatigue and back pain during increased frequency of extended flight missions (6 to 8+ hours) by pilots flying during Operation Iraqi Freedom and Enduring Freedom, Harrer et al. (2005) investigated WBV exposure for U.S. Navy MH-60S pilots. Pilots were exposed to continuous (WBV. Pilot fatigue is a growing operational concern due to the increased frequency of extended durations of missions (6 to 8^+ hours) in support of Operations Iraqi Freedom and Enduring Freedom. The then current rotary wing seating systems were not optimized for the longer missions and wide range of pilot anthropometric measurements, which is now typical of naval aviation. The current seating systems were designed primarily to meet crashworthiness requirements, not for the wide range of pilot anthropometry or to mitigate WBV. Current Hazard Reports (HAZREP) indicated that pain in both pilots' legs and backs begin 2 to 4 hours into the flight and increase with time. Situational awareness also decreases with an increase in flight duration due to the constant distraction of pilots shifting in their seats while flying to get comfortable. Froom (1987) reported a dose-response relationship between the length of military helicopter flights and back discomfort. He also concluded that this pain is typically dull, over the lower back, and its prevalence and intensity are dependent on the total flight hours of exposure.

The Harrer et al., (2005) study compared the effectiveness of three different seat cushions, the current seat cushion versus two anti-vibration seat cushions (A and B). The three seat cushions were measured for acceleration levels averaged over five-minute intervals using a triaxial seat pad accelerometer. The recordings were completed for several round-trip straight and level flights. A frequency analysis from 0 to 80 Hz was conducted on all acceleration measurements to determine the dominant axis and frequency of the pilots' vibration exposure. The results were then compared to the applicable Threshold Limit Values (TLVs) established by the American

Conference of Governmental Industrial Hygienists (ACGIH) (2005) to determine the MH-60S pilots' permissible exposure time for all three seat cushions.

The results showed that pilots of the MH-60S could operate the helicopter with the current seat cushion for less than 6 hours and the anti-vibration seat cushion B for approximately 8 hours without being overexposed to WBV. The anti-vibration seat cushion A increased the stay-time to approximately 16 hours. Since the average flight during a deployment or mission could last in excess of 8 hours, exposure with the current seat cushion would place the pilots at an unacceptable risk of injury, lack of mission readiness, and possible equipment damage. As helicopters are to be outfitted with auxiliary fuel tanks to accommodate the long-duration missions, this will further extend a pilot's overall sitting (exposure) time. In order to lower the pilots' exposure to WBV and reduce potential safety mishaps, the study recommended that the current MH-60S's be retrofitted with the anti-vibration seat cushion A.

Of course, WBV is not just an aviation problem; it also is a concern in all tactical vehicles. Army Regulation (AR) 40-10, *Health Hazard Assessment Program in Support of the Army Material Acquisition Decision Process* (1991), requires all new tactical vehicles and aircraft to be evaluated for potential WBV health hazards. Moran and Butler (1993) conducted one such evaluation of the U.S. Army's M9161A1 Truck Tractor. The vehicle was tested in bobtail (no trailer), unloaded, and loaded configurations for each of the three test terrains. The results showed that the lowest tolerance levels were experienced on the Belgian block course,[27] with less severe WBV occurring on the cross-country course, followed by the primary terrain course. The results also show the passenger exposure limits were consistently lower than the driver's. The evaluation recommendation for the M916A1, operating in its intended environment, was that WBV be limited to the following passenger exposure limits for each test condition: WBV is not to exceed 17.1 hours in any 24-hour period on the paved surfaces for all configurations. Exposure limits for the cross-country terrain are 5.5, 5.2, and 6.1 hours in any 24-hour period for the bobtail, unloaded, and loaded configurations, respectively. For the Belgian block terrain, WBV in any 24-hour period should not exceed 1 hour for both bobtail and unloaded conditions and 2 hours for the loaded configuration.

The use of HMDs aggravates the effects of vibration on pilots, vehicle drivers and crew. Originally designed for crash and impact protection, helmets in many applications now serve as platforms for mounting displays, chemical protective masks, oxygen systems, and laser and flashblindness protection systems. All of these add-ons increase head-supported weight (HSW), which in turn contributes to increased biomechanical stress/strain on the muscles of the neck that are responsible for controlling head movements (Butler and Alem, 1997).

The impact of helmet/HMD weight can be characterized by the total system mass and change in center-of-mass (CM) (offset) due to the addition of the helmet/HMD to the normal head-neck CM. The helmet/HMD mass and CM combine to create a torque that must be counterbalanced by the muscles in the back of the neck to maintain upright posture. The head, too, creates a torque that attempts to rotate the head, moving the chin downward towards the chest. The torques from the helmet/HMD system and the head combine to create a torque that is larger than the torque due to the head alone. The pivot point through which this torque operates is on top of the cervical spine and is known as the atlanto-occipital (AO) complex (Sobotta, 1990). Figure 16-16 shows the head, the location of the A0 complex on top of the cervical spine, and the locations of the head center-of-mass and the helmet center-of-mass. Force vectors also are shown located at each center-of-mass. These force vectors must be counterbalanced by the muscles in the back of the neck.

The total head-supported mass is not the only factor affecting the stress on the posterior neck muscles. The presence of WBV, as is always true in tactical vehicles and aircraft, can causes the head to pitch up and down (Paddan and Griffin, 1988). This pitching motion causes an involuntary stretch response in the posterior muscles of the neck that further increases the amount of force produced by these muscles. The duration of both training and combat missions requires the posterior muscles of the neck to exert counterbalancing forces for a greater period of time than required under more natural conditions. These factors affect the amount of biomechanical

[27] The Belgian block was an oval cobblestone road approximately ½-mile long with an irregular pattern of 3-inch crests.

stress experienced by the posterior neck muscles and play a role in determining a reasonable head-supported mass limit for Warfighters.

Psychological effects of vibration

Human response to WBV exposure can be psychological as well as physiological. The psychological aspect deals with the level of tolerance, which depends to a great extent on the on the environment and the task being performed (Nakashima, 2004). For example, higher magnitudes of vibration would likely be tolerated, or deemed acceptable, in a mass-transit railroad train rather than in a luxury automobile.

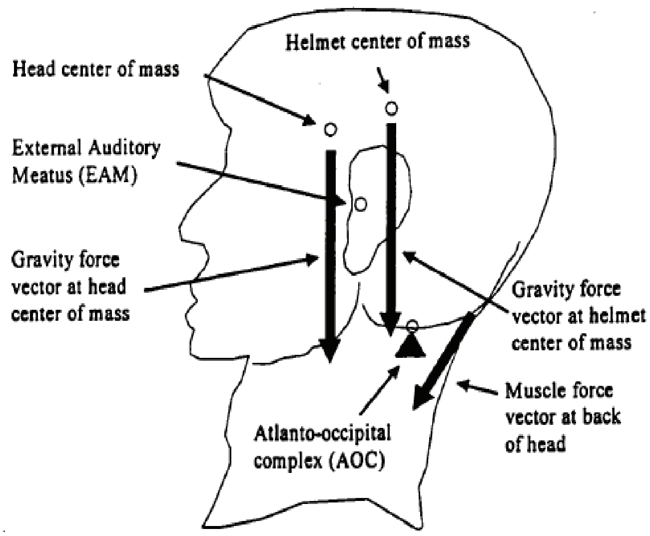


Figure 16-16. Head and neck profile showing the A0 complex, head center of mass, helmet center-of-mass, and force vectors representing the gravity field and the posterior of the neck.

Before the values of a detrimental stimulus become unacceptable to a human from the perspective of degradation in performance or onset of physiological effects, there is a range of levels of the stimulus that may be tolerated but described as *annoying*. Such is the case for short-term or low-amplitude vibrations. It is well known that noise, or unwanted sound, can interfere with speech communication, degrade concentration and interrupt sleep patterns (Nakashima, 2004). However, when the noise is low frequency, a vibration of the body and the surrounding area also may result, which may influence the human perception and acceptance of low-frequency noise. For example, in mechanized environments, the presence of rattle and vibration can increase the level of annoyance (Berglund and Hassmén, 1996; Howarth and Griffin, 1990). Field surveys on the effects of noise and vibration from railway traffic found that residents who were exposed to combined noise and vibration expressed a higher degree of annoyance than those exposed to noise alone (Ohrstrom, 1997; Ohrstrom and Skanberg, 1996).

Performance effects of vibration

If the visual display and its viewer are exposed to vibration that results in an out of-phase oscillation with respect to each other, a blurred image will be seen. Thresholds for this blurring effect depend on the magnitude and frequency of the vibration and can be calculated. In general, the threshold acceleration increases in proportion to the viewing distance and the square of the frequency (Nakashima, 2004). The frequency range of approximately 2 to 20 Hz is associated with such display vibration. Above 20 Hz, the threshold accelerations for blur are rarely encountered (Griffin, 1990). In aircraft, head-up displays and HMDs are collimated by a lens to reduce the image distortion caused by translational vibration. Stabilization systems that move the image on the display can counteract the rotational motion of the head, resulting in greater legibility (Griffin, 1990).

Vibration can affect performance by introducing alignment issues or by interfering with peripheral motor and sensory functions (Kjellberg, 1990). Lewis and Griffin (1976) suggested that interference with kinesthetic feedback mechanisms may be a principal means by which vibration degrades performance in tracking tasks, important in targeting HMDs. While studies have shown that the performance of tasks involving simple reaction time are not significantly affected, vibration has been shown to have a negative affect on more complex cognitive tasks, such as those involving short- and long-term memory. However, the relationships between vibration frequency and magnitude on performance are unclear (Nakashima and Cheung, 2006).

Not surprisingly, as the optics in many HMD designs is frequently levered out from the face, HMDs are particularly susceptible to the effects of WBV (Wells and Haas, 1992). Furness (1981) reported that, at some frequencies, the reading error produced with a panel-mounted display, was present in HMDs at approximately one-tenth of the vibration amplitude. Wells and Griffin (1987a) reported that the number of numerals read correctly from the HMD in a helicopter decreased from 2.4 per second while stationary on the ground to 1.0 per second during in-flight vibration. The reason for the vibration-induced decrement in performance is relative motion between the line-of-sight and the optical axis of the HMD. Rotational oscillation of the head causes vibration of the HMD, but the eyes, under the influence of the vestibular ocular reflex (VOR),[28] remain space-stable (Benson and Barnes, 1978). The VOR, which normally serves to keep images stable on the retina during body movement and vibration, acts to degrade performance with HMDs.

Vibration is also a factor in the use of helmet-mounted sights and head-coupled systems, where head movement is used to direct weapons, sensors, and other systems. Under normal circumstances a person can aim his/her head at a stationary target with pitch and yaw errors as small as 0.1° rms (Wells and Griffin, 1987b). Tracking moving targets with the head is easily learned (Wells and Griffin, 1987c) and, depending on the difficulty of tracking the target motion, can be accomplished successfully. WBV disrupts both head-aiming and head-tracking. With random vibration, aiming at a stationary target is disrupted by the vibration-induced head motion (vibration breakthrough). However, the decrement in head-tracking during vibration is greater than the sum of the decrement caused by tracking and the decrement caused by vibration breakthrough (Wells and Haas, 1992). It is likely that the additional decrement results from attempts to reduce the error between the head-mounted reticule and the target, which is due to lags in the response of the head, result in greater error.

From the perspective of alignment, HMDs require wearers to maintain their eye(s) within the exit pupil(s) of the HMD. Most HMDs have sophisticated and often complex fitting techniques to maintain this alignment. Helmet slippage due to sweat challenges this alignment so does vibration. Smith (2004) conducted a study to characterize cockpit seat and pilot helmet vibration in jet aircraft during aircraft carrier flight operations. Accelerators were used to measure triaxial accelerations at the seat base, seat pan, seat back, and HMD in the F/A-18C (Hornet) jet aircraft. Data were collected during flight operations on two aircraft carriers for a total of 11 catapult launches, 9 touch-and-goes, and 4 arrested landings. Of particular interest was the substantial low frequency seat and helmet vibration observed during the catapult launch. During the stroke period, seat and helmet vertical (Z) accelerations reached 6G and 8G peak-to-peak, respectively, and occurred in the frequency range of 3 to 3.5 Hz. The associated helmet pitch reached peak-to-peak displacements ranging between 9° and 18°. The large helmet rotations were believed to be associated with helmet slippage that can cause partial or complete loss of the projected image on an HMD (vignetting). This is highly undesirable when using the HMD as the primary flight reference. The study recommended that one goal of HMD designers should be “to develop helmet-mounted equipment design guidelines that consider hostile vibratory environments.”

Nakashima (2004) in a discussion of cognitive effects of vibration cites two studies. First, Harris and Shoenberger (1980) studied individual and combined effects of noise and vibration on cognition by monitoring

---

[28] The VOR consists of the constant adjustments of the image in the retina of the eye by the nuclei of the brain stem, which receives information from the eyes, the neck, trunk, cerebellum, and cerebral cortex.

the ability to perform a complex counting task. Twelve male subjects were exposed to 65 or 100 dBA broadband noise, with and without 0.36 rms G-vibration (in the vertical). The vibration was a quasi-random sum-of-sines signal composed of five frequencies: 2.6, 4.1, 6.3, 10 and 16 Hz. The complex counting task involved keeping a simultaneous count of the number of flashes of three lights that flashed at different rates. The study concluded that exposure to broadband noise in combination with a complex vibration signal had a negative effect on the cognitive performance of the subjects, compared to exposure to noise alone.

In the second study, Ljungberg, Neely and Lundstrom (2004) also investigated the effect of combined exposure to WBV and noise on the short-term memory performance. Fifty-four subjects were randomly assigned to low (77 dBA noise and 1.0 m/s^2 vibration), medium (81 dBA noise and 1.6 m/s^2 vibration) or high (86 dBA noise and 2.5 m/s^2 vibration) levels of exposure for duration of 20 minutes. The noise signal was helicopter noise with a dominant 21-Hz component. The memory task involved observing two, four or six letters on a screen for a period of 1, 2 or 3 seconds, after which a probe letter appeared. The subject had to indicate as quickly as possible if the probe was present in the previous list of letters. The subjects stated that it was more difficult to perform the task when exposed to combined noise and vibration, and the high exposure group indicated the highest levels of annoyance. However, no evidence was found for the hypothesis that combined noise and vibration degraded cognitive performance compared to one stimulus on its own. The authors stated that the results were inconsistent with Harris and Shoenberger (1980). Nakashima stated that “the inconsistencies among the results of experimental studies on the effect of combined noise and vibration on cognitive performance are indicative of the complexity of interaction between the two stimuli. There is currently no concrete evidence to support that whole-body vibration exposure has a negative effect on cognition.”

Much work has been done in reducing the modes of vibration in modern helicopters. Nonetheless, except for cruise flight, the helicopter cockpit is still a high vibration environment. This is even more so for rear crew stations that do not have the same seat designs as the front cockpit. This vibration affects both the aircraft and the aircrew. The effects of vibration manifest themselves as retinal blur, which degrades visual performance, and as physiological effects, whose resulting degradation is not fully understood (Biberman and Tsou, 1991). Rotary-wing aircraft differ in their vibrational frequencies and amplitudes and these vibrations are triaxial in nature. However, in general they have a frequency range in all axes of 0.5 to 100 Hz. However, specific frequencies of significant amplitude are associated with the revolution rates of the rotor, gears, engines, and other mechanical components (Boff and Lincoln, 1988). The largest amplitude frequency occurs at the main rotor blade frequency multiplied by the number of blades. Other frequencies having significant amplitude include the main rotor frequency (~7 Hz); two, eight, and twelve times the main rotor frequency; tail rotor frequency (~32 Hz); twice the tail rotor frequency; and the tail rotor shaft frequency (~37 Hz). These vibrations are transmitted to the head through the seat and restraint systems (peak transmission, 3 to 8 Hz). They are typically in the vertical and pitch axes and are affected by posture, body size, and add-on masses, such as HMDs. However, the transfer function of these vibrations to the eye is not straightforward. The activity of the vestibulo-ocular reflex stabilizes some of the vibrational transfer, mostly low frequency. However, visual performance degradation still will be present. To further complicate this scenario, the vibrational transfer function to the helmet and HMD is different from that to the eye. While the general influencing factors are the same, e.g., posture, body size, etc, the helmet/HMD mass is also a factor. The result is a very complex frequency and amplitude relationship between the eye and the HMD imagery, which results in relative motion between the imagery and the eye (Wells and Griffin, 1984).

Vibration standards

There are numerous occupational vibration standards used worldwide for both HAV and WBV. The standards have been established to address: human health and comfort, the probability of vibration perception, and the incidence of motion sickness. In the U.S., the occupational standards used for HAV are:

- ANSI S3.34, Human Exposure to Vibration Transmitted to the Hand, Guide for Measurement and Evaluation of. (American National Standards Institute, 1986).[29]
- ACGIH-HAV, Hand-Arm Vibration Standard. (American Conference of Governmental Industrial Hygienists, 2001).
- NIOSH #89-106, Criteria for a Recommended Standard for Occupational Exposure to Hand-Arm Vibration. (National Institute for Occupational Safety and Health, 1989).

For WBV, the standards in the U.S. are:

- ANSI S3.18, Mechanical vibration and shock – Evaluation of human exposure to whole-body vibration. (American National Standards Institute, 2002).
- ACGIH-WBV, Whole-body vibration: TLV physical agents (American Conference of Governmental Industrial Hygienists, 2001).

International standards include:

- ISO 5349 (for HAV), Mechanical vibration – Measurement and evaluation of human exposure to hand-transmitted vibration. (International Organization for Standardization, 2001).
- European Union Directive 2002/44/EC (for HAV and WBV) – The minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (vibration). (The European Union, 2002).
- ISO 2631-1997 (for WBV) – Mechanical vibration and shock: Evaluation of human exposure to whole-body vibration. (International Organization for Standardization, 1997).

Acceleration

Acceleration is to the aviation system what jolt and vibration are to ground systems, and in the high-G acceleration world of aviation operations, the pilot operator, rather than system design, is the limiting factor in system performance. Modern military aircraft routinely operate in a high-G environment. Additionally, a high sortie[30] rate and sustained operations is the "norm." Therefore, pilots must be in excellent physical and mental condition to perform their duties, both in training and in combat. Acceleration forces on the human body are important to understanding in-flight performance because of their effects on the cardiovascular, pulmonary, and vestibular (orientation) systems. The ability to overcome the effects of acceleration becomes more important as aircraft are designed with greater maneuverability and performance. The ability to combat the adverse effects of G-forces depends directly on one's level of physical condition and ability to reduce negative life stressors.

Before G-forces are discussed in depth, several basic terms are defined below to help the reader understand acceleration and how G-forces are generated.

Physical principles

Speed is the rate of motion (or how far one travels in a certain amount of time), irrespective of direction. An example is flying at 360 knots groundspeed. *Velocity* describes both a rate of motion (speed) and a direction of motion. An example of velocity is 360 knots groundspeed on a heading of 180°. *Acceleration* is a change in velocity per unit time and is generally expressed in feet per second per second (ft/s^2) or meters per second per

[29] While still used, this standard has been replaced by ANSI S2.70-2006: Guide for the Measurement and Evaluation of Human Exposure to Vibration Transmitted to the Hand. (American National Standards Institute, 2006).
[30] Sortie – a mission flown by a military aircraft.

second (m/s^2). Acceleration is produced when either speed or direction (or both) change. One very familiar type of acceleration is that due to gravity. Gravity affects anything on or near the Earth. The acceleration produced by gravity (g) is a constant, 32 ft/s^2 or 9.8 m/s^2. Therefore, a free-falling body will increase its velocity by 32 ft/s or 9.8 m/s for every second it falls. The inertial force resulting from the linear acceleration of gravity acting upon a mass is termed 1G. Therefore, when we discuss G-forces in the flying environment, we are referring to the inertial force resulting from acceleration. Generally, G-forces are dimensionless and expressed as multiples of Earth's gravity, e.g., 5G.

Types of acceleration

There are several types of acceleration. *Linear* acceleration is a change in speed (increase or decrease) without a change in direction. For example, linear acceleration occurs when an aircraft is in a takeoff roll or landing rollout. *Radial* acceleration is a change in direction without a change in speed. When a body moves in a circular path with constant linear speed at each point in its path, it is also being constantly accelerated toward the center of the circle under the action of the force required to constrain it to move in its circular path. This acceleration toward the center of path is called radial acceleration. Radial acceleration occurs when an aircraft pulls out of a dive, pushes over into a dive, or performs an inside or outside turn (and does not change its speed). In these examples, the aircraft's direction changes, but the airspeed remains the same. *Angular* acceleration is a simultaneous change in both speed and direction. Angular acceleration is the most common type of acceleration for aviators and occurs during most aerial maneuvers. For instance, when an aircraft performs a split-S maneuver,[31] the aircraft's speed and direction change simultaneously and the crew experiences angular acceleration.

As an aircraft accelerates in one direction, inertial forces act on the body in the opposite direction of the applied force. The inertial force causes the body to experience a G-force. The following section discusses the types of G-forces a crewmember experiences and the physical factors influencing the effects of G-forces on the body.

Acceleration is experienced primarily across three axes: fore and aft (x-axis), side to side (y-axis) and head to foot (z-axis) (Figure 16-17). The three types of G-forces can be further classified into transverse G, negative G, positive G, and lateral G. By determining the direction of the force and its axis, the type of G can be specified. G-forces can be experienced along other axes as well, but the force applied along the z-axis has the most significant effect on aviator performance. For instance, a force applied from the head towards the feet is a positive Gz force (+Gz) and a force applied from the feet towards the head is a negative Gz force (-Gz). Transverse G-force is the force applied to the front (+Gx) or back (-Gx) of the body. +Gx and -Gx forces are normally encountered during takeoffs, acceleration in level flight, and landing. The maximum transverse G-force tolerable to humans is roughly 15G in the +Gx direction and about 8G in the -Gx direction. Lateral G-forces (the Gy direction) are experienced during spin or roll; however, the effects are negligible. Aircraft are equipped with an accelerometer (G-meter) that monitors G-forces during flight. It displays instantaneous G, maximum positive G, and maximum negative G. The dial also indicates the maximum permissible G-force the aircraft can sustain, both positive and negative.

The maximum tolerance for G acceleration (both in number of G's and time) for each of the different axes at an onset rate of 25G per second (G/s) are (US Navy Aircraft Investigation Handbook, April 1988):

x-axis:	83 +Gx / 0.04 s,	25 -Gx / 2.0 s
y-axis:	9 +Gy / 0.10 s,	9 -Gy / 0.10 s
z-axis:	20 +Gz / 0.10 s,	15 -Gz / 0.10 s

[31] The *split-S* is an air combat maneuver primarily used to disengage from combat. To execute a split-S, the pilot half-rolls his aircraft inverted and executes a descending half-loop, resulting in level flight in the exact opposite direction at a lower altitude.

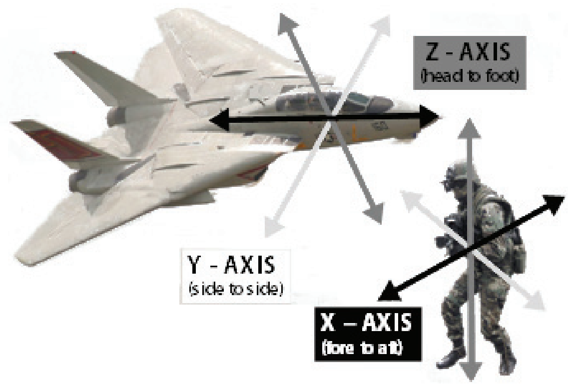


Figure 16-17. Acceleration is a dimensionless parameter and occurs in three axes.

Specific body tolerance is determined by the magnitude of the G-force, the duration of exposure to the G-force and the rate of application (or onset) of those forces. The magnitude of the G-force is the size of the G-force applied to the body. The greater the magnitude of acceleration and accompanying inertial force, the greater the resulting G-force. For instance, a crewmember pulling +6 Gz is being accelerated to six times the gravitational force of the Earth, or 192 ft/s^2. Modern fighter aircraft, like the F-18 and F-16, are capable of exposing the pilot to sustained 8G to 9G. Duration of exposure to the G-Force is another determinate of the effects of the G-force on the body. For example, jumping from a table one meter high results in a decelerate force of about 14G for a fraction of a second, usually with no ill effects. But, being exposed to 14G for over 2 seconds will result in significant physical and physiological effects. Rate of application (or G-onset) directly influences the effect of a G-force. Rate of G-force application is expressed in G per second (G/s). To illustrate the effect of G-onset, imagine dropping a brick on someone's foot versus placing a brick of identical mass on the person's foot. The dropped brick has a greater physical effect on the foot than the brick placed, even though both bricks are identical in mass; the difference is in the rate of acceleration and the resultant inertial force. Acceleration or G-forces along the Z-axis (e.g., accelerated/decelerated turns or maneuvers; positive or negative acceleration) are of special concern in aviation because of the adverse impact on human systems such as the cardiovascular, cerebral, respiratory and visual systems. For example, the average time to a visual symptom (grayout)[32] of +Gz exposure is determined by the rate of G-onset. The slower the onset, the longer the time to grayout in the low to moderate G ranges.

Physical considerations

Several factors or physical considerations (of the operator) determine the effects of G-forces on the body. These factors help explain why certain G-forces have different effects on the body and why the body reacts to certain types of G-forces in different situations. It is important to note that some of these factors are interrelated and have a combined effect on the crewmember.

Previous G-exposure effects G-tolerance. For example, the push-pull effect (PPE) is a phenomena of reduced +Gz tolerance when preceded by exposure to Gz that is less than +1Gz. It is thought that the less than +1Gz exposure causes a cardiovascular relaxation which can affect subsequent +Gz tolerance. A -Gz exposure for a

[32] A grayout (or greyout) is a transient loss of vision characterized by a perceived dimming of light accompanied by a brown hue and a loss of peripheral vision. .It is a precursor to fainting or a blackout and can be caused by hypoxia, a loss of blood pressure or restriction of blood flow to the brain.

duration of less than 2 seconds can significantly affect +Gz tolerance, possibly reducing tolerance by up to 1.5G (dependent upon magnitude and duration of the –Gz exposure). Maneuvers that produce the PPE include dive attacks, extensions, air combat maneuvering guns defense and split-s maneuvers. PPE can reduce G-tolerance by 30% to 40%. However, another aspect of previous G-exposure is the fact that the body can be prompted to prepare for increased G. The G warm-up is a maneuver consists of a very controlled exposure to increased G that prepares the pilot for higher G follow-on maneuvers.

Positive G-force effects G-tolerance. As mentioned earlier, positive G-force is the force applied from the head towards the feet. It is expressed as +Gz. It occurs during turns and dive recoveries and is the G-force most often experienced by crewmembers. Physiological tolerance to positive G is usually indicated by visual symptoms. Blood pooling in the lower extremities usually begins at 1 to 3 +Gz. This decreases head level blood pressure, and at higher +Gz blood flow to the brain ceases (there is generally a 22 mm Hg drop in head level arterial pressure per additional "G"). Initially, the decreased blood pressure results in gray-out of the visual system (between 3 to 4 +Gz). However, the brain has only a 4- to 5-second oxygen reserve and once the oxygen reserve is used, unconsciousness results. The average resting tolerance to +Gz is 5.5G and by 4 to 5 +Gz crewmembers may begin to blackout, with most pilots experiencing gravity-induced loss of consciousness (G-LOC) by 5 to 6 +Gz. With high-G onset rates, unconsciousness can happen without any preceding visual cues, so preventing G-LOC is a blood pressure control game. The pilot must perform an anti-G straining maneuver (AGSM) or unload the Gs immediately. The AGSM sustains blood flow during the critical period of G onset and can provide 3.5 to 4 +Gz of protection provided it is performed correctly. The AGSM is performed by tensing the skeletal muscles (particularly in the lower extremities and the abdomen), cyclic breathing, and exhaling against a closed glottis. The AGSM is started prior to G onset and does not stop until the aircraft returns to 1G flight.

Our bodies are conditioned to live in a positive-G environment; accordingly, we have an increased tolerance to positive G's. However, negative G-forces are not tolerated well by humans, mostly as a result of physical discomfort.[33] A negative G-force is defined as the force being applied from the feet towards the head and is expressed as -Gz. Negative G-force adversely effects G-tolerance; exposures to negative G (between 0 to -1G) for as short as 2 seconds can reduce tolerance by as much as 1.5G during subsequent "pulls" to positive G. Fortunately, -G conditions are seldom experienced in high levels during normal flight. Normally, -Gz is experienced when the nose of the aircraft is lowered during a "pushover" or when experiencing turbulence. In -G maneuvers, the baroreceptors sense the increased blood pressure at the brain level and in response open up the peripheral blood vessels to try to decrease blood pressure with slowing of the heart rate. The physical symptoms of -Gz are a sense of weightlessness, congestion in the head and face, headache, and visual blurring. Blood begins pooling in the head at about 1 -Gz and vision can be affected with as little as 2.5 -Gz. Some flyers have reported a phenomenon called "redout," a reddening of vision during sustained negative Gz flight, however, the causes of redout are not completely understood. The limits of human tolerance (due to physical discomfort) to -Gz begins to appear at - 2.5 to -3Gz , and greater than -3Gz can be physically incapacitating. Currently there is no practical method to counteract the effects of -Gz. Under normal conditions; the only way to combat the effects of -Gz is to reduce aircraft maneuvering and return to a 1-G environment.

Physiological effects and symptoms

Prolonged exposure to G-forces affects the body in four principle ways – restricting mobility, affecting the cardiovascular system, stimulating the vestibular system, and reducing visual acuity. A 150-pound crewmember

---

[33] Children hanging from upside-down by their feet experiencing only -1 Gz notwithstanding. However, in such situations, they are not being called upon to perform demanding physical or cognitive tasks. Inversion tables and similar devices purport to reduce back pain by creating a -Gz environment. The May Clinic cites no scientific evidence to support this claim and cautions individuals with heart disease, high blood pressure and eye diseases (e.g., glaucoma) to avoid the use of these devices (Mayo Clinic, 2007).

weighs 600 pounds when exposed to +4 Gz. This increase in weight severely restricts mobility and movement in the aircraft. For example, the head weighs about 29 pounds when wearing a typical helmet and oxygen mask. At 4 +Gz, the same head-helmet combination has an effective weigh of approximately 116 pounds. This increased weight can force the unprepared pilot's chin into his chest when a loop is initiated. Combined with other physiological effects of +Gz, decreased mobility interferes with the ability to function at peak levels during high-G flight. Additionally, as +Gz forces increase, blood pressure begins to decrease because of the effects of the G-forces on the cardiovascular system. Each +Gz drops blood pressure 22 mm Hg. The cardiovascular system attempts to compensate for the drop in blood pressure by constricting peripheral blood vessels and increasing the heart rate. This compensation is known as the cardiovascular reflex. Vestibular effects and their symptoms also play a critical role in spatial disorientation and balance. The otoliths are stimulated by gravity and linear acceleration forces to provide you a sense of direction. The semicircular canals respond to angular acceleration to provide another sense of direction. If pilots fail to rely on their instruments and visual cues, acceleration forces can provide stimuli that induce disorientation, motion sickness, vomiting and vertigo. As already mentioned, the visual system is affected by high G-forces. For blood to enter the retina, the cardiovascular system must overcome about 13-18 mm Hg of intraocular pressure. As the G-forces increase and the blood pressure in the brain begins to drop, there is insufficient blood pressure to overcome the intraocular pressure. Therefore, the tissue in the eye that detects light (retina) starts losing its blood supply. As the blood supply is decreased, peripheral vision is affected and pilots experience a dimming, misting, or graying of your vision referred to as grayout or they may experience tunnel vision, where the only vision remaining is in the center of the visual field. As the G-force increases, the blood pressure drops to where it cannot overcome the intraocular pressure and all vision is lost, referred to as black-out. It is important to note that black-out does not mean unconsciousness; however, the blacked-out pilot is in imminent danger of G-LOC.

The effects of G-LOC are described as two phases of incapacitation – absolute and relative. In absolute incapacitation the pilot is actually unconscious for roughly 9 to 21 seconds, with an average time of 15 seconds. During this period the body generally relaxes. However, during the latter stages of absolute incapacitation, pilots may experience marked involuntary skeletal muscle contractions and spasms just before regaining consciousness. These contractions can cause the arms to flail, leave the flight controls, or hit other aircraft controls. The second phase of incapacitation is experienced once the pilot regains consciousness. Unfortunately, there is not an instantaneous return to an alert and functional state. Pilots often experience mental confusion, disorientation, stupor, apathy or memory loss. During this time, they are incapable of consciously flying the aircraft, making decisions, taking action against a threat, or communicating effectively. The time of relative incapacitation usually mirrors that of the absolute incapacitation. Auditory stimulation during this period speeds recovery to alertness, although, dissociation, stupor, and feelings of uneasiness often linger after recovery from G-LOC.

Variability in G-tolerance

G-tolerance changes from day to day and hour to hour based on a number of variables. Understanding the reasons for these variables can help maximize tolerance and minimize the threat of acceleration effects. The following section describes some of the physiological factors and their effects on G-tolerance as well as physical protections against acceleration effects.

The role of self-imposed stress

Crewmembers generally drink less water than they need and are slightly dehydrated most of the time. Dehydration reduces G-tolerance markedly by depleting blood plasma volume. Aircrews must drink plenty of noncaffenated, nonalcoholic fluids (even when not thirsty) prior to (and during) flight. The body suffers a 35% decrease in ability to do anaerobic work and a 20% decrease in ability to do aerobic work if you are 3% dehydrated. Therefore, an AGSM can only be maintained for one-half the time it normally would. For instance, if a pilot can normally pull 9G for 10 seconds, the effects of dehydration would limit him to 9G for 5 seconds. Fatigue also significantly decreases G-tolerance. Crewmembers that are fatigued or are lacking sleep tend to experience lapses in mental function and a lower ability to maintain muscle tension during the AGSM. Mental fatigue slows your response and anticipation of high-G maneuvers. Physical fatigue lowers the capability to maintain adequate muscle strain during the AGSM and also lowers the capability to perform subsequent strains. Warfighting aviators should take maximum advantage of crew rest, stay well rested and maintain good sleep patterns prior to flying. Safe flight demands that pilots perform at peak levels in a high-G environment, however, self-medication with over-the-counter drugs can decrease their performance. Those who require medication should not be flying. They are a danger to themselves and their fellow crewmembers. Therefore, pilots are instructed to not self-medicate, to report to the flight surgeon, and to always obtain qualified medical treatment. Alcohol misuse, and the accompanying hangover, drastically reduces G-tolerance. The reduced G-tolerance is primarily due to alcohol's dehydrating effects. In addition, a hangover clouds mental capability, slows the thinking and decision-making processes, as well as the ability to effectively judge situations. Alcohol-use should be avoided prior to flight. Remember from the previous section that although regulations generally restrict alcohol consumption 12 hours prior to flight or mission planning, some detrimental aftereffects can last as long as 48 to 72 hours. Additionally, alcohol can also contribute to fatigue and hypoglycemia. Food is the fuel used to function in a high-G environment. Missing meals or not taking the time to eat correctly directly affects ones ability to withstand increased G-force. Pilots will not have fuel in their system to maintain high levels of activity for extended periods of time if they do not eat or if they eat improperly prior to flight.

Prevention methods

Sometimes referred to as a G-suit, fast pants, or "speed jeans," these devices consist of a pair of pants-like covers fitting tightly over the leg and lower abdomen (Figure 16-18). Air bladders in the thigh, calf and abdomen areas of the suit are automatically inflated by an anti-G valve on the aircraft. However, the G-suit is not the primary means of G-LOC protection and used by itself, only allows for 1 to 1.5G of protection. Pilots must also rely on the AGSM to protect themselves from G-LOC. G warm-up maneuvers also prepare the pilot for subsequent high-G maneuvers. This maneuver consists of a total of 180° of turn and is used to operationally check G-suits and to practice straining maneuvers up to an amount of G approaching the maximum amount anticipated on that particular flight.

Figure 16-18. G-suit.

Physical conditioning mentioned previously is a method to improve muscle strain during the AGSM. Physical conditioning is also important in decreasing the fatigue levels and increasing stamina required for multiple G maneuvers. Both anaerobic and aerobic physical conditionings are encouraged. The AGSM is essentially an anaerobic maneuver. The muscles used to perform the AGSM rely upon anaerobic energy sources (energy sources not requiring oxygen). Crewmembers flying high performance aircraft are encouraged to develop a weight training program to maximize their muscle strain ability. Weight training is the primary method of anaerobic conditioning and decreases your chances of injury, particularly neck injury during high-G maneuvers. Anaerobic conditioning increases the muscle's ability to contract and sustain the contraction throughout the G stress. Without sufficient anaerobic conditioning, the muscles fatigue quickly, and the AGSM loses its efficiency. However, developing a conditioning program based solely on anaerobic exercise is not complete.

Aerobic conditioning must complement your anaerobic conditioning. Pilots need to be aerobically fit to combat fatigue and recover from multiple G-maneuvers. Aerobic exercise programs require oxygen to produce the necessary energy. Aerobic conditioning increases stamina and resistance to fatigue. (G-LOC typically occurs towards the end of engagements during the fatigue period.) Aerobic conditioning does increase cardiovascular fitness, leading to lower heart rates, lower blood pressure, and faster recovery times from aerobic exercise. Unfortunately, these attributes of aerobic exercise are not entirely beneficial and may lead to problems in the high G environment. Therefore, it is not recommended for fighter aircraft crewmembers to pursue an excessive competitive aerobic exercise program; an aerobic exercise program that does not exceed the equivalent of running twenty miles per week is suggested. Overall, for crewmembers that fly in high performance aircraft, a sound anaerobic training program coupled with a sensible aerobic exercise program will help maximize their G-tolerance. However, exercising prior to high-G flight leaves one in a pre-fatigued state and dehydrated, and is not recommended.

Summary: Acceleration

G-forces are the result of inertial forces acting on the body. G is a dimensionless number expressed as a ratio of a body's acceleration to the force of gravity (32 ft/s^2 or 9.81 m/s^2). The magnitude, duration of exposure, rate of application, direction of force applied and previous G exposure are physical factors influencing the body's physiological response to a G-force. These factors define the G-force and can predict the effect the G-force will affect performance. +Gz is the force of greatest concern since it is regularly encountered in-flight. The effects of +Gz are decreased mobility, visual disturbances like grayout and blackout, and finally G-LOC. Physiological factors will increase or decrease your G tolerance. These factors include your physical condition and self-imposed stresses (fatigue, dehydration, self-medication, alcohol use and nutrition). Staying in shape, avoiding self-imposed stresses, and performing an effective AGSM will help increase G-tolerance and decrease the effects of acceleration on performance.

Ambient lighting

Use of HMDs is not confined to nights only. While pilotage imagery mostly may be employed at night and during other periods low illumination (e.g., dawn, dust and periods of weather-related low visibility), HMD symbology may be employed over the entire 24-hour period.

Since HMDs produce visible energy for viewing of their information, night operation offers little problems. A basic understanding of the human visual system's response to low illumination (scotopic and mesopic vision) is most sufficient for HMD designers and can be reviewed in Chapter 7, *Visual Function*. Nonetheless, there is one potential problem for HMD users during night and low-illumination operation. This problem is associated with chromatic aftereffects and first was raised in the early 1970s (Glick and Moser, 1974). This afterimage phenomenon was reported by U.S. Army aviators using NVGs for night flights. It was initially, and incorrectly, called "brown eye syndrome." The reported visual problem was that aviators experienced only brown and white color vision for a few minutes following NVG flight. Glick and Moser (1974) investigated this report and concluded that the aviator's eyes were adapting to the monochromatic green output of the NVGs. When such adaptation occurs, two phenomena may be experienced. The first is a "positive" afterimage seen when looking at a dark background; this afterimage will be the same color as the adapting color. The second is a "negative" afterimage seen when a lighter background is viewed. In this case, the afterimage will take on the compliment color, which is brown for the NVG green. The final conclusion was that this phenomenon was a normal physiological response and was not a concern. A later investigation (Moffitt, Rogers, and Cicinelli, 1988) looked at the possible confounding which might occur when aviators must view color cockpit displays intermittently during prolonged NVG use. Their findings suggested degraded identification of green and white colors on such displays, requiring increased luminance levels.

For HMD designers the more difficult problem is supplying sufficient luminance in the presence of high ambient lighting conditions, specifically high intensity light sources, which are discussed in the following section.

High intensity light sources

The Warfighter can be exposed to a number of high intensity (bright) light sources. These sources can be natural (e.g., the Sun) or artificial (e.g., lasers, explosions, searchlights, and fires). The effects of such exposure can range from glare, through flashblindness (or dazzle), to retinal burns. For HMD users, lasers have always been a major concern. For example, with NVGs, lasers viewed directly will shut down the image intensifier (I^2) tubes, causing loss of imagery. Such shutdowns of I^2 tubes can result from virtually all high "brightness" sources. Of greater concern with lasers are situations where the laser energy may enter the eyes from the periphery. This can result in flashblindness which is a temporary vision impairment that follows a brief exposure to bright light and interferes with the ability to detect or to identify a target, or even retinal injury (Rash and Manning, 2001). However, while laser exposure is still a concern in the military community, it has yet to become the severe operational hazard, as previously anticipated.

For the Warfighter, glare is the most frequently encountered effect from high intensity light sources; this is especially true for HMD users. Glare can be classified into various types by either its source:

- Direct glare – bright light in the FOV
- Reflected glare – bright light reflected from a surface; and further categorized as:
    - Specular (smooth, polished surfaces)
    - Spread (pebbled, brushed surfaces)
    - Diffuse (flat-painted, matte surfaces)
    - Compound

Or, its effect (Hedge, 2008):

- Discomfort glare – produces discomfort, does not impair vision
- Disability – reduces visual performance
- Blinding – causes temporary blindness

Both direct and reflected glare are of issue with HMD use. Blinding glare tends to be transient and infrequent but has the most severe visual impact. However, disability glare is very common, more frequently degrading visual performance. Disability glare reduces visual performance by reducing image contrast or visually distracting an individual. Usually, but not always, glare is transient, being a serious problem only when it happens during some critical moment – a purposely induced or inadvertent inability to detect or identify an object on the side of the road or ahead while driving at night; a temporary inability to read instruments while flying or targeting an enemy; a reduced capability when using night vision devices. Refractive surgery often exacerbates susceptibility to disability glare, increasing its magnitude and frequency of occurrence. Consequently, this increases risk to personnel in the battlespace.

The study of glare

Glare was described by Goethe in 1810 and Purkinje in 1823. Their explanations portended the neural versus physical (scatter) debate that was clearly framed by Helmholtz in 1852. Cobb, in 1911, was the first to quantify disability glare by developing the concept of *equivalent background* (taken from Vos, 2003). The concept was expanded by Holladay (1926; 1927), Stiles (1929), and Stiles and Crawford (1937). Their work, formally presented at the 1939 Commission Internationale de l'Eclairage (*CIE*) meeting, culminated in a formula that clearly implied that intraocular scatter was the main cause of glare:

$$L_{eq} = 10E_{glare} / \theta^2 \qquad \text{Equation 16-3}$$

where, L_{eq} is the equivalent veiling background in candelas per square meter (cd/m^2), E_{glare} is the illuminance of the glare source at the eye measured in lux, and θ is the angular distance between the line-of-sight and the glare source in degrees. For an extended glare sources this formula is integrated over the angular aperture of the glare source. Subsequent research, carefully controlling pupil size and eye movement, substantiated the proportionality of L_{eq} and E_{glare}. In addition, it was shown that the forward scatter from the cornea, crystalline lens and ocular fundus, taken together, are sufficient to explain L_{eq} (Vos, 2003).

The Holladay-Stiles formula is still widely used and considered a good estimate for glare from sources between 1° and 30°. It should also be noted that this formula was widely used during World War II vision research.

Fry and Alpern (1953) referenced a 1939 observation by Schouten and Ornstein, … " that the depression of brightness still persists when the image of the glare source falls on the optic nerve head," an area without receptors and lateral neural connections. Fry and Alpern found that the course of foveal dark adaptation following a peripheral glare source or a direct veiling illumination followed the same pattern. In addition they showed that increasing the glare angle was equivalent to decreasing the direct veiling illuminance. These studies argued that the brightness of the foveal image of a test object was a consequence of forward light scatter in the eye caused by a peripheral glare source and not lateral neural effects.

Around 1965, the *CIE* asked Vos to head a committee to update the Holladay-Stiles formula. He had recently completed a doctoral dissertation on the mechanisms of glare. In a succession of papers that followed he showed that the cornea, lens and fundus contributed about equally, with some variability, to forward scatter in the normal eye (Vos, 1963; Vos and Boogaard, 1963; Vos and Bouman, 1964). Vos also showed that the three sources of scatter alone could account for the L_{eq} in the Holladay-Stiles formula, putting to rest the physical scatter-neural controversy.

There were other major issues regarding scatter that also had to be solved. One had to do with the question of wavelength. In general it has been found that stray light (scatter) in the eye is independent of wavelength (van den Berg, Ijspeert and de Waard, 1991; Vos, 2003; Wooten and Geri, 1987). However, van den Berg et al. (1991) found a small wavelength-dependent scatter with transmission of light through the ocular wall of subjects with blue eyes. The effect was virtually zero for subjects with dark brown eyes. They concluded that "depending on pigmentation, eye-wall transmittance and fundal reflections do introduce some wavelength dependence." This suggests that most scatter in the eye is Mie[34] scatter, due to intraocular and intracellular particles substantially larger than the wavelengths of visible light (Figure 16-19). This is consistent with scatter produced by most cataracts, lens opacities resulting from hereditary factors, trauma, inflammation, ultraviolet radiation, drugs, or disease (de Waard et al., 1992; Kanski, 2003; Klein, Klein and Linton, 1992; Schneck et al., 1993; Smith, 2002;

---

[34] Mie scattering described by the German physicist Gustav Mie is based on an analytical solution of Maxwell's equations for the scattering of electromagnetic radiation by spherical particles (1910).

Thomson, 2001). Cataracts are usually whitish, occasionally brunescent, and are made up of fairly large particles. The amount of scatter in the normal eye that is independent of wavelength has also been shown to be related to eye pigmentation, more pigmented eyes generally showing less scatter (van den Berg, Ijspeert and de Waard, 1991; Vos, 2003; Vos and van den Berg, 1999).

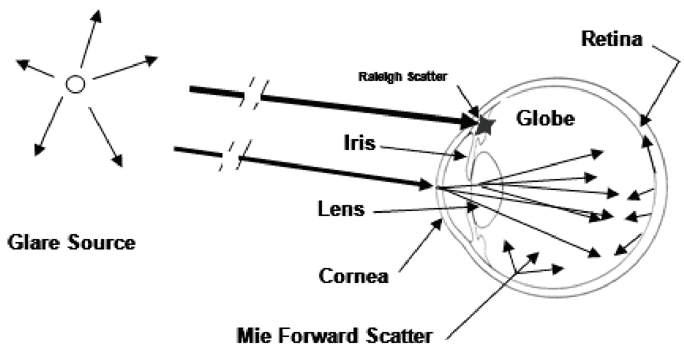


Figure 16-19. The cornea, lens and retina contribute about equally to Mie scatter. There is very little wavelength-dependent scatter, mostly for individuals with little eye pigment.

Another major factor affecting disability glare is age (Ijspeert et al., 1990; Vos, 2003; Vos and van den Berg, 1999). De Waard et al. (1992) found that light scatter increases by a factor of three by age 80. Schieber (1994b, 1995) extensively reviewed the impact of visual aging on driving performance, pointing out that there is not only an increase in glare sensitivity with age, but also an increase in glare recovery time. Swanson (1998) pointed out that scatter increases significantly with age and that as little as one-third of the light reaches the retina in a 65-year-old as in a 25-year-old. Additionally, he pointed out that light scatter in the lens is responsible for a majority of the complaints of disability glare for older adults, often leading to a voluntary cessation of night driving. Haegerstrom-Portnoy et al. (1999) showed that, even though everyone experiences disability glare to some extent, the effect is accelerated after age 65.

The measurement of glare

There have been many attempts, but no universally adopted technique, a gold standard, for measuring disability glare. The technique currently gaining the widest support, particularly in Europe, has adopted the stray light definition of glare and called the problem solved, i.e., glare is the veiling light that results from forward Mie scatter in the eye (van Rijn et al., 2005b). Van Rijn et al. (2005a) stated that:

> "…disability glare is the reduction in visual performance caused by veiling luminance on the retina. It is an effect of intraocular stray light. Measurements of glare and stray light are particularly important for drivers, cataract, and refractive surgery. Glare testing in the elderly may be important in view of the high accident rates in this age group, especially at night. Moreover, glare measurements may predict future decrease of visual acuity."

This would be fine, except that this approach does not, for example, adequately predict the nighttime "glare" experienced by refractive surgery patients when they see the headlights of an oncoming vehicle. Stray light measurement is certainly a major factor in creating glare and works very well for evaluating cataract patients. However, with the advent of refractive surgery, now widely accepted within the military community, other factors have come into play. Van den Berg (1991) questioned the validity of most glare-testing, stating that:

> "…present tests for the stray light type of glare fail on validation research. Also, in clinical use, the reliability of glare testing seems to be questionable. The problem is the absence of a generally accepted reference, a golden standard of glare."

Ghaith et al. (1998) found that disability glare assessment 1, 3, and 6 months refractive surgery techniques of post radial keratotomy (RK) and photorefractive keratectomy (PRK), using measurements from the Brightness Acuity Tester (BAT) and the Multivision Contrast Tester (MCT 8000), did "not accurately reflect patient's subjective assessment of their visual performance in daily life as expressed in a questionnaire." These devices measured high and low contrast visual acuity (VA), which should be affected by veiling glare resulting from forward Mie scatter. In a personal correspondence Barbur (2004) said that visual performance of most refractive surgery patients does not differ significantly from normal subjects having had no surgery. However, he pointed out that there are some significant "outliers" that present with demonstrable vision problems, particularly when under mesopic ambient illumination important to the Warfighter, when a large pupil size favors increased aberrations.

The major visual problems that patients experience are night vision glare, reduced contrast sensitivity, halos and starburst (Bailey et al., 2003; Fan-Paul et al., 2002). These problems are usually reduced a few months after surgery (McLeod, 2001). However, it is not entirely clear whether this is a resolution of the problem, patient adaptation, or simply self-justification, i.e., resolution of cognitive dissonance (Brunette et al., 2000; Chou and Wachler, 2001; Melki, Proano, and Azar, 2003).

The incidence of problems is also unclear and somewhat dependent on the diameter of the ablation zone (Martinez et al., 1998):

> "Depending on the magnitude of the attempted correction and the size of the ablation zone, past PRK studies have reported 15% to 60% of patients complaining of glare, 26% to 78% complaining of halos, and 12% to 45% complaining of difficulty with night vision. As many as one third of patients after PRK have reported to be disappointed with their results, despite good uncorrected visual acuity or even emmetropia. In some studies, up to 10% of patients who underwent PRK with an ablation zone 4.00 mm in diameter considered the problem of halos severe enough to interfere with driving at night."

More recent papers have reported significant reductions in night vision problems with both Laser-Assisted *In Situ* Keratomileusis (LASIK) and PRK (Figure 16-20). This has come with an increase in ablation zone diameter and better ablation techniques. Of 690 questionnaires answered, 55.1% of patients reported an increase in daytime glare and 31.7% reported a decrease in the quality of night vision following surgery (Brunette et al., 2000). In spite of this, they reported that 96.2% said they believed having the surgery was a good choice. Bailey et al. (2003) surveyed 841 patients (returning questionnaires) and found a 117%-increase in reporting starbursts for each 1-mm decrease in ablation diameter. In a report on a single patient Chalita and Krueger (2004) performed wavefront-guided LASIK enhancement surgery after lifting the preexisting flap on a 3-year post-LASIK patient who presented with post-LASIK symptoms of glare, halo and double vision. The retreatment outcome was complete resolution of double image and halos. This outcome coincided with a reduction in both low- and high-order aberrations. Chalita et al. (2004) found a strong correlation between wavefront measurements/aberrations and visual symptoms such as coma, starbursts and glare.

Klein (2001) said that, "Night vision is an embarrassing topic for refractive surgery […] A large percentage of post refractive surgery eyes have large pupils at night that result in disturbing halos." This is particularly true for individuals in their 20s, a prime age for the Warfighter. Currently, there is often a poor correlation between measurements made with current devices to measure post-surgery scatter and visual performance. There is a need for better measurement strategies to evaluate and predict post-refractive surgery vision at night, particularly for young individuals, where virtually all of the vision is under conditions of low illumination with large pupils (Barbur, 2004; Klein, Hoffman and Hickenbotham, 2003; Klein, 2001; Sagawa and Takeichi, 1992).

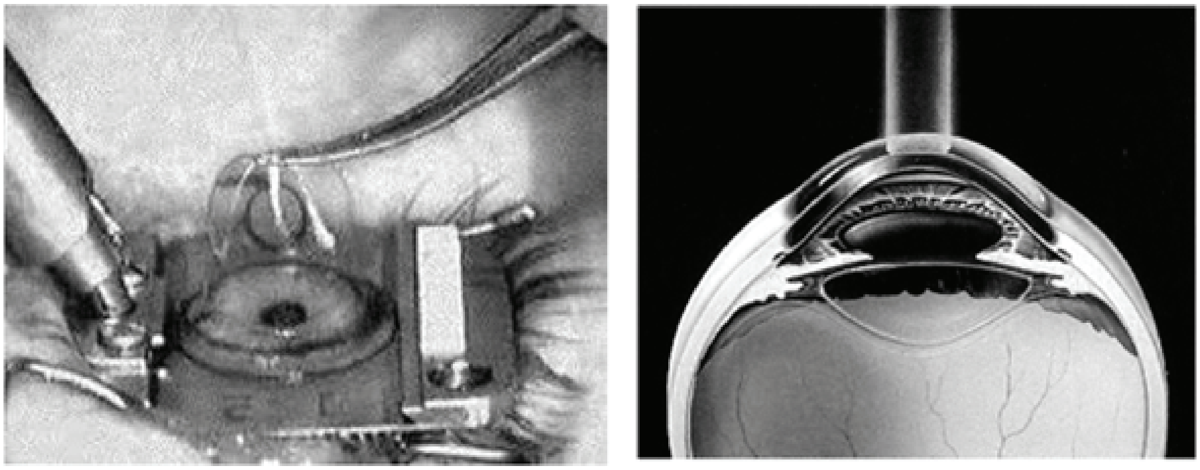


Figure 16-20. LASIK refractive surgery (left) removes a flap from the cornea before laser ablation. After laser ablation the flap is replaced. PRK refractive surgery (right) removes only the top layer of cells (epithelium) from the cornea prior to laser ablation. These cells grow back, from the periphery toward the center. After 6 months the outcome of these surgeries is very similar.

Complexity of the disability glare problem

Vision is a dynamic process. Nowhere is this more obvious than in the long history of efforts to develop a practical definition and measure of disability glare. We all know it when we see it – halos, reduced contrast, starburst patterns that can obscure objects due to oncoming automobile headlights at night. Outwardly it seems simple enough, but definition, quantification and reliable prediction are very difficult. In fact, it is very complex on several levels, the anatomical/structural levels, sensory levels, the physics and environmental levels, and the cognitive and perceptual levels. Disability glare is the overall consequence of a myriad of changing, interactive factors that can increase inter-subject variability and mask retinal image degradation (Chisholm et al., 2003). The factors discussed below are condensed from several sources (Atchison and Smith, 2000; Bron, Tripathi and Tripathi, 1997; Kaufman and Alm, 2002; Korb et al., 2002). All these factors impact susceptibility to disability glare or its measurement.

The anatomy and physiology are themselves very complicated (see Chapter 6, *Basic Anatomy and Physiology of the Human Eye*). The cornea is a multi-layered armature that forms the surface on which the tear layer, the first optical surface, forms and reforms through blinking (Bron, Tripathi and Tripathi, 1997). The tear layer/cornea provides more than two-thirds of the optical power of the eye (Atchison and Smith, 2000). Multiple glands in the lids and around the eye socket form the complex tear layer that has both aqueous and oily components (Bron, Tripathi and Tripathi, 1997; Korb et al., 2002). The power of this aspherical optical surface can and does vary as a function of time. In addition, irregularities in the shape and curvature of the cornea/tear layer can create variations in focus. These variations are called astigmatism and aberrations. The corneal stroma (the central layer of the cornea), scatters about 10% of the visible light striking it.

The crystalline lens is a multi-layered optic just behind the iris/pupil (Bron, Tripathi and Tripathi, 1997). It grows by increased layering throughout life. The crystalline lens is suspended like a trampoline by the zonule fibers. Sphincter muscles attached to the zonule fibers can pull on the lens to change its shape and, therefore, the eye's focus (the process of accommodation). Over the years the lens increases in size and stiffens, losing its ability to change focus/accommodate (presbyopia), as well as some of its clarity (Bron, Tripathi and Tripathi, 1997; Ciuffreda, 1998). Changes in the fibers and proteins of the crystalline lens can produce local areas that scatter light (cataract), sometimes becoming opaque (Hemenger, 1990; Swanson, 1998; Thomson, 2001). Cataracts produce primarily Mie scatter.

Aberrations, caused by shape irregularities of the eye's optical surfaces, are partially counterbalanced by the cornea-lens optical combination (Kelly, Mihashi and Howland, 2004; Artal, Berrio and Guirao, 2002), but this balance can be upset by both age and refractive surgery (Artal, Berrio and Guirao, 2002; Artal et al., 2003). The power of the cornea and the accommodated lens combine with the distance between the lens and the photosensitive retina to determine whether the eye will be emmetropic (requiring no correction to focus an image).

The pigmented iris, between the cornea and the lens, is an aperture stop that reduces intraocular stray light primarily through absorption (Keating, 1988; van den Berg, Ijspeert and de Waard, 1991). The pupil, the physical opening in the iris, can change diameter with changes in illumination, convergence of the two eyes, accommodation, and emotion (Bron, Tripathi and Tripathi, 1997; Ciuffreda, 1998; Lowenstein and Loewenfeld, 1969). In the dark, the pupils of young adults can be well over 7 mm in diameter (Schumer, Bains and Brown, 2000). On-average, they become smaller with age.

Within the eye, between the cornea and the lens is a circulating fluid called the aqueous humor. It provides nutrition and oxygen for the avascular cornea. Posterior to the lens is the vitreous humor, a gel/liquid. With age, this substance begins to have localized areas of different refractive index due to pockets of localized liquefaction. These variations in homogeneity act as local optical surfaces in the main optical pathway of the eye and can cause scatter (Bron, Tripathi and Tripathi, 1997; Smith, 2002).

The multilayered inside-surface of the eye, the retina, is a specialized extension of the brain with some 130 million or more photoreceptors of various kinds (Bron, Tripathi and Tripathi, 1997). The photoreceptors reside in a back layer of the retina behind the retinal nerve cells. Posterior to the photoreceptors is the pigment epithelium, a pigmented layer that helps to reduce stray light in the eye, and the choroid, a highly vascular layer. Complex, laterally interacting neurons lay anterior to the photoreceptors (toward the incoming light), except in the fovea centralis, which is a tiny depression in the retina about 1.5 mm in diameter, where most of the neurons are pushed aside. Most of our sharp vision takes place in this area (Bron, Tripathi and Tripathi, 1997; Schwartz, 1994). The photoreceptor cells of the eye are of two general types. The more densely packed, thinner cells are called cones. They are most concentrated in the fovea and parafovea regions, rapidly diminishing in density peripheral to these areas. They function under brighter lighting conditions, provide our sharpest acuity, and are responsible for the first stage of neural color processing. The larger photoreceptors are called rods. They are absent in the fovea, increase in density peripheral to the fovea and then begin to decrease in density. These cells are very sensitive to light and motion, but provide much poorer acuity. They do process brightness variations, but do not process color information. Photoreceptor cells change their sensitivity and range of sensitivity to light as ambient lighting conditions change – adaptation level. Initially, this process can be very rapid (Boynton, Bush and Enoth, 1954; Bron, Tripathi and Tripathi, 1997; Schwartz, 1994). Cones and rods interact neurally in complex ways. This is particularly important when considering lower illumination levels (Krizaj, 2000; Krizaj and Hawlina, 2002; Stabell and Stabell, 1979, 1998)

As the light environment changes the eye adjusts. Vision, as living, is a dynamic, changing process. When the ambient illumination increases, the eye light adapts; when illumination decreases, the eye dark adapts (Baker, 1949, 1953; Barlow, 1972; Bartlett, 1966, Graham, 1966a-b, Schwartz, 1994). The time course for these processes is also a variable, with the most rapid changes occurring during the first few seconds of an illumination change. These chemical, neural and mechanical changes combine with sensory and perceptual neural processing within the brain to help maintain a relatively stable representation of the world visually (Schwartz, 1994).

The effect of scattered light may be enhanced under conditions of low light adaptation. Intraocular stray light can cause a dark-adapted retina to light-adapt, producing a prolonged reduction in vision after the glare source has been removed. "With pathologically increased dark adaptation the effect can be stronger" (van den Berg, 1991). Steady stimuli, producing scattered light that acts more as an adapting stimulus (altering the state of adaptation), can create a paradoxical increase in contrast sensitivity as ambient light increases at low luminances (Bichao, Yager and Meng, 1995). In general, transient light stimuli are considerably more effective at producing glare and raising thresholds than are steady glare sources (Bichao, Yager and Meng, 1995). Under some conditions there can be a persisting visual after-image (following light stimulation), particularly in a relatively uncluttered FOV (Brown, 1966). This after-image can be a result of retinal or central neural activity (Shinsuke, Kamitani and Nishida, 2001).

In general, cone receptors operate above approximately 3.4 cd/m^2; these brightness levels are photopic. Between about 0.034 to 3.4 cd/m^2, moonlight, vision operates with both rods and cones; these brightness levels are mesopic. Only rods operate at brightness levels below 0.034 cd/m^2; these brightness levels are scotopic. Most of us are using photopic or mesopic vision at night while driving a car. Ultimately, one million neural fibers from each eye are sent to an area of the brain called the lateral geniculate nucleus. From this nucleus on there is a continuing cascade of neural processing within the brain (abstracting and assembling information originating at the retina). This results in a representation of the external world that combines with memory, other senses and emotion, forming a context for behavior appropriate to our biological niche. Glare can interrupt this process.

Intraocular glare results from and combines with many environmental (extraocular) factors. The most obvious are the glare source's color, brightness, temporal characteristics and angle with respect to the observer's line-of-sight (Vos and van den Berg, 1999). But there are many other environmental factors – scratched windshields or windscreens, eyeglasses or goggles, contact lenses, type of contact lenses, fog, rain, snow and ice, time of day, other objects like automobile chrome, flashes at night, use of night vision devices, the context in which glare occurs, and more. Each of these factors or combination of factors can influence the degree and importance of glare (Applegate, 1989; Applegate and Wolf, 1987; de Wit and Coppens, 2003; Elliott, Mitchell, and Whitaker, 1991; Lewis, 1993; Pitts, 1993).

Perceptual and cognitive factors combine with sensory input to play a role in disability glare (Allen et al., 2001; Anderson and Holliday, 1994; Green, 2004; Green and Senders, 2004; Pulling et al., 1980; Schieber, 1994a-b). Issues of target acquisition, recognition and identification depend on contrast sensitivity, context, masking, clutter, and other factors, as well as sensory considerations. A bright headlight may cause a reduction in pupil size, decreasing aberrations of the eye and improving acuity, but the individual may not see an unexpected object due to reduced light gathering ability of the eye and consequent reduced contrast. However, an expected object may be seen under the same conditions.

Disability glare is the combined consequence of a multitude of interacting factors, many of which are nonlinear. And disability glare can be dangerous. It can be dangerous when a Warfighter has only a split second to detect or identify someone or something. It also can be dangerous when a pilot needs to detect or identify another aircraft or is engaged in critical, low-altitude maneuvers. With the advent of refractive surgery and its increasingly wider application in the military, the issue of glare with younger people has become real. Of equal concern is that there is no current gold standard for measuring glare and predicting problems from glare at night.

Ergonomic Issues

The term *ergonomics* refers to the applied science of designing the characteristics of devices that humans use in such a way as to ensure efficient, effective, comfortable and safe use. In this section we will discuss two specific ergonomic issues that are associated with HMD design and use – eyewear and user controls – and then address the global issue of compatibility with a host of components, devices and systems that may be required to be used by Warfighters in combination with their HMD, i.e., system compatibility.

Eyewear

For Warfighters, eyewear includes devices used for vision correction and for eye protection. Therefore, the discussion below is limited to corrective spectacles, sunglasses, ballistic protective goggles, laser protective goggles, and nuclear-biological-chemical (NBC) masks (Figure 16-21). Most of this eyewear is necessary to prevent or reduce injury from dust, wind, shrapnel and debris, laser energy, or the sun. In conflicts involving improvised explosive devices (IEDs), mortars, sand, wind and dust, protective eyewear is essential equipment

(Dawson, 2008). Data from the Iraq conflict shows that 10% of Warfighters injured have injuries to the eye(s). Therefore, it becomes a matter of fact that this eyewear will be used in conjunction with HMDs.

Military Eye Protection System (MEPS)

Sun, Wind, Dust Goggles

Aviator Sunglasses

M-43 NBC Protective Mask

Figure 16-21. Examples of military eyewear.

Although provided with government-issued eyewear, a number of Warfighters elect to purchase their own eyewear products from commercial vendors. Unfortunately, many Warfighters do not possess the knowledge to make selections that meet military requirements. The military has addressed this problem by developing programs that test commercially-available eyewear. One program, the U.S. Army’s Military Combat Eye Protection Program (MCEPP)[35] tests commercial protective eyewear to military ballistic and ANSI Z87.1 (American National Standards Institute, 2003) standards. The program maintains an Authorized Protective Eyewear List (APEL), which is available to Warfighters via the Internet at http://peosoldier.army.mil/pmseq/ eyewearmessage.asp

Until rather recent designs, an HMD’s optics historically has been located very close to the eyes. This close proximity results in a very small distance between the optics and eye(s). This has proven to be an important equipment compatibility issue. The operational requirements of warfare have necessitated that Warfighters be provided with protection against directed energy (e.g., lasers, microwave radiation, etc.) and chemical warfare environments. Protection has been generally in the form of protective spectacles, goggles, or masks. Most of these protective add-on devices must be located between the HMD optics and the eyes. Oxygen masks are an additional requirement for moderate above-sea-level altitudes. Current HMD designs provide little space for incorporation of these additional devices and systems.

In addition, as Warfighters age, they undergo changes in their visual capability. Aviators experience the same sort of refractive error progression as the general population; individuals who are nearsighted or farsighted tend to become more nearsighted or farsighted with age, resulting in increased dependence on glasses or contact lenses. One of the most pronounced effects is the ability to accommodate, i.e., change focus. Human range in accommodation generally decreases with age from a robust 11 diopters at age 20 years to a limiting 2 diopters by

---

[35] The MCEPP was created by the Program Executive Office (PEO) Soldier, an organization with the U.S. Army whose primary purpose is to develop equipment that can be rapidly fielded.

age 50 years (Records, 1979). As a result, to retain the experience of older aviators, there is a requirement to provide visual correction, and this correction must be useable while wearing HMDs.

HMDs are examples of optical systems. Simply stated, an optical system consists of one or more optical elements. These optical elements include lenses, mirrors, prisms, filters, etc. One of the simplest optical systems is a magnifying glass, which consists of a single lens encased in a ring that may have a handle attached. Beyond the simple magnifier, practically all optical systems contain multiple optical elements. HMD optical systems are generally quite complex and can consist of a dozen or more optical elements.

Exit pupil and eye relief

As optical systems, most HMDs have their optical elements fixed in place within a housing. Furthermore, these systems are designed to be viewed by the human eye. Figure 16-22 shows the optical design of a simple pupil-forming compound microscope and the path of light rays through the system. For ease of discussion, the optical elements are presented only. Light rays passing through the system form an image at the exit pupil. Simply defined, the exit pupil is where the eye must be placed in order to optimally view the image. The exit pupil can be thought of as the area through which all of the image-forming rays pass. [Note: Technically, the exit pupil is a volume in space.] If the eye is placed behind or in front of the exit pupil, the eye will not capture some of the rays. This results in a reduced FOV (Rash et al., 2003).

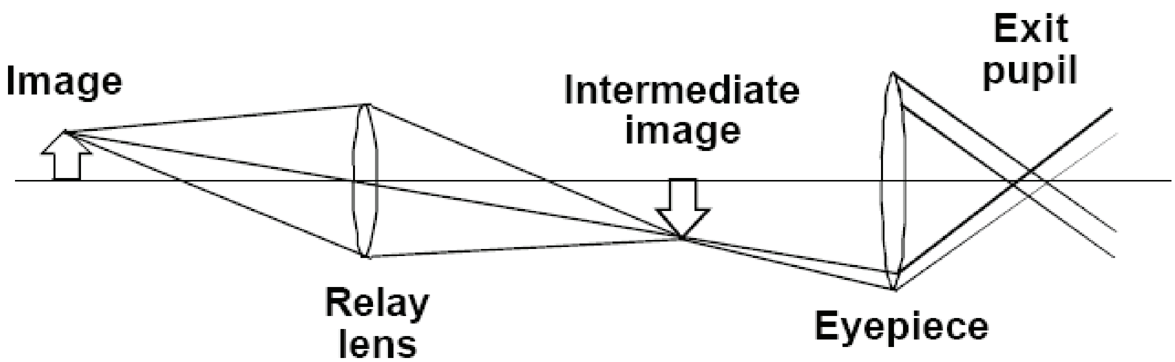


Figure 16-22. The path of light rays through a pupil-forming system.

An important characteristic of the system in Figure 16-22 is the distance along the optical axis between the last optical element and the exit pupil (where the eye would be positioned). This distance is known as the “optical eye relief.” Figure 16-23 expands the final element of the system and presents this distance. In addition, it further refines the definition of the optical relief as the distance along the optical axis from the last optical surface to the cornea of the eye. [Note: Often, the entrance pupil of the eye, which is approximately 3 mm behind the surface of the cornea, is used as the reference point in the definition of optical eye relief.] When an optical system is defined, the optical eye relief distance is often cited as an important parameter.

In Figures 16-22 and 16-23, the optical system was depicted as exposed optical elements. But, in practice, one cannot ignore the system housing. Figure 16-24 shows a side cut-away view of the example optical system showing the last optical element as enclosed in the housing. The most noticeable difference when the housing is considered is the extension of the housing beyond the final surface of the last optical element. This difference impacts the available (or usable to the viewer) optical eye relief distance. A new distance requires definition, that of the distance from the plane through the outer edge of the housing to the cornea of the eye. This new distance is often referred to as “physical eye relief.” Physical eye relief is, at best, equal to optical eye relief. In practice, physical eye relief almost always is less than optical eye relief.

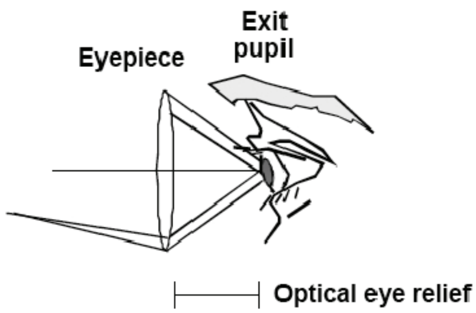


Figure 16-23. Defining the optical eye relief distance as the distance along the optical axis from the last optical surface to the cornea of the eye.

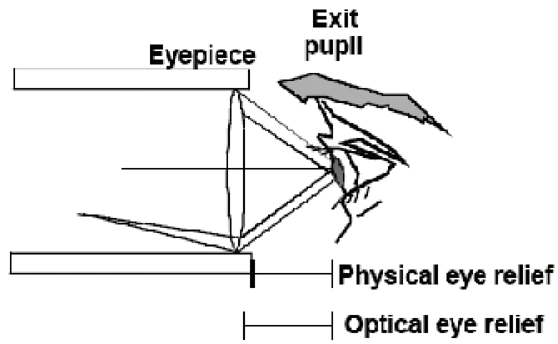


Figure 16-24. A side cut-away view of the example optical system showing the last optical element as enclosed in the housing.

Consider the display unit of the Integrated Helmet and Display System (IHADSS) HMD fielded on the AH-64 Apache helicopter. The IHADSS is a monocular design. Imagery obtained from a nose-mounted thermal sensor is reproduced on a miniature, 1-inch diameter, cathode-ray-tube mounted on the right side of the helmet. The image on the face of the CRT is optically relayed to the pilot’s right eye. The relay optics and CRT are referred to as the Helmet Display Unit (HDU) (Figure 16-25). Two optical elements of the HDU should be noted. The first is the objective lens that is positioned almost perpendicular to the pilot’s face. This lens is approximately $1^3/_4$ inches in diameter and is mounted as the last element in the HDU housing barrel. The second is the beam splitter mounted on the side of the HDU barrel farthest from the face. The beam splitter, also referred to as the combiner, reflects the IHADSS imagery into the pilot’s eye.

Figure 16-25 demonstrates the difference in the concepts of optical and physical eye relief for the IHADSS HMD. The IHADSS design optical eye relief is 10mm. By definition, this distance is the distance along the optical axis from the last optical element (center of the combiner) and the exit pupil. Figure 16-25 shows why the optical relief distance is not a functional (practical) parameter. The center of the combiner is located well back behind the lip of the HDU barrel housing. The HDU barrel and the interaction between the barrel and the wearer’s cheekbone limit how close the combiner can be placed in front of the eye. This situation severely reduces the available distance between the pilot’s eye and the plane that passes through the closest physical HDU structural element, and it is this distance that defines the physical eye relief.

In summary, *optical* eye relief distance is an optical system design parameter. However, it is a misleading parameter when the optical design is intended for use in systems such as HMDs where intervening devices must be placed between the optical system and the users eye, e.g., corrective spectacles, oxygen masks, nuclear, biological and chemical (NBC) protective mask, etc. A more useful parameter is *physical* eye relief. Physical eye relief distance, usually less than optical eye relief distance, takes into consideration the physical features of the structure and housing of the optical system’s elements and these features’ impact on reducing the “real” distance available between the optical system “HMD” and the viewer’s eye.

Before leaving the description of eye relief, it may be worth addressing why the HMD design cannot simply provide a greater physical eye relief distance. For any HMD design, the two starting parameters that must be decided upon are exit pupil size and eye relief. However, these two parameters have considerable impact on the size of the last optical element in the HMD’s eyepiece and the focal length of the system. All of these factors combined have additional impact on packaging size and total head-supported weight, very important parameters for HMD use in the military environment. In conclusion, the designer simply cannot make the eye relief distance as large as may be desired.

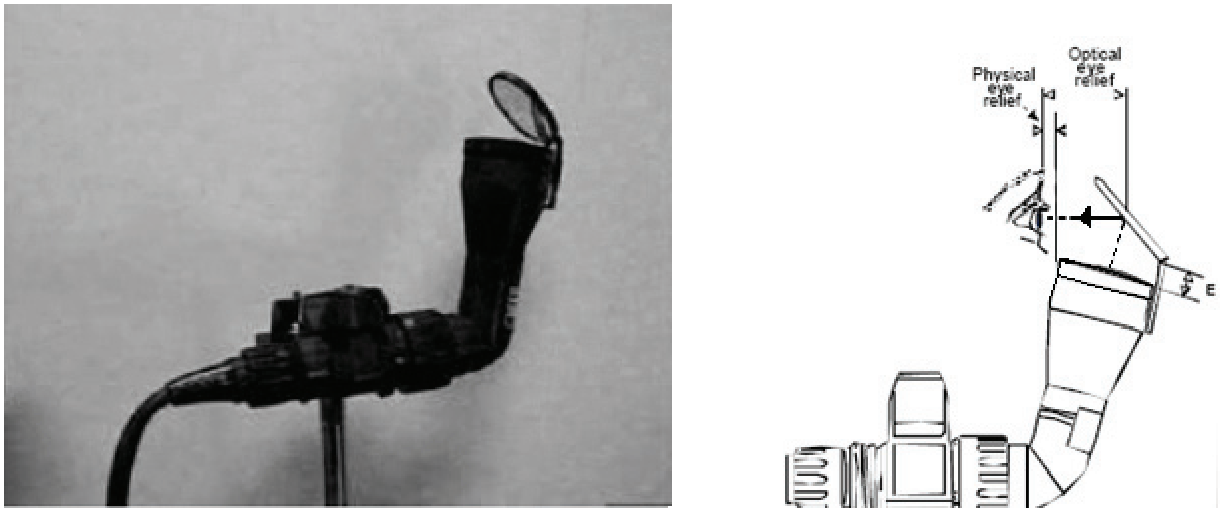


Figure 16-25. The IHADSS HMD helmet display unit.

Vision correction and HMDs

It is estimated that up to one-third of active duty U.S. Army, Navy, and Air Force Warfighters require corrective lenses (Madigan and Bower, 2004).[36] This percentage also is applicable to both the aviation and ground/ship Warfighter communities. Spectacles have been the traditional solution for visual correction. However, they pose complex issues for Warfighters using HMDs. Among these are: discomfort when worn with the helmet, slippage, reduced FOV, and interfering reflections off the lens and frame. Incompatibility with NBC protective masks and some forms of laser eye protection also have been problems.

<u>Spectacles</u>

Spectacle lenses are held in a frame, supported on the nose. Two arms attached to the frame hold the assembly to the head. The distance from the back part of the correcting spectacle lens to the eye's corneal surface is called the vertex distance (Figure 16-26). It ranges from 8 to 18 mm. Lens thickness is usually several millimeters, depending on material, type of glass or plastic, and lens shape, determined largely by the amount of correction required (Benjamin, 1998). Consequently, the front of the lens forms a surface that is some distance from the face, limiting ability to position or place the eye in the exit pupil. This necessitates increasing the required physical eye relief for many optical devices including HMDs (Licina, 1998; Melzer and Moffitt, 1997; Rash and McLean, 1998). If a pilot uses bifocals, the position of the head and eyes are restricted so that the region of the lens having the appropriate power (correction) is centered between the object being viewed and the pupil of the eye. However, in spite of some obvious limitations, eyeglasses work very well in most situations. They are cheap, durable, can be worn for extended periods of time, and are easy to manufacture and maintain. They provide excellent vision and a wide range of vision corrections.

---

[36] As would be expected, requirement for visual correction increases with age, e.g., U.S. Air Force normalized, age-related data showed that a fairly constant percentage (over time) of 21- to 40-year-old Air Force pilots wear spectacles, but that almost 50% of ages 41 to 45 years do and approximately 90% of pilots over age 45 years wear spectacles.

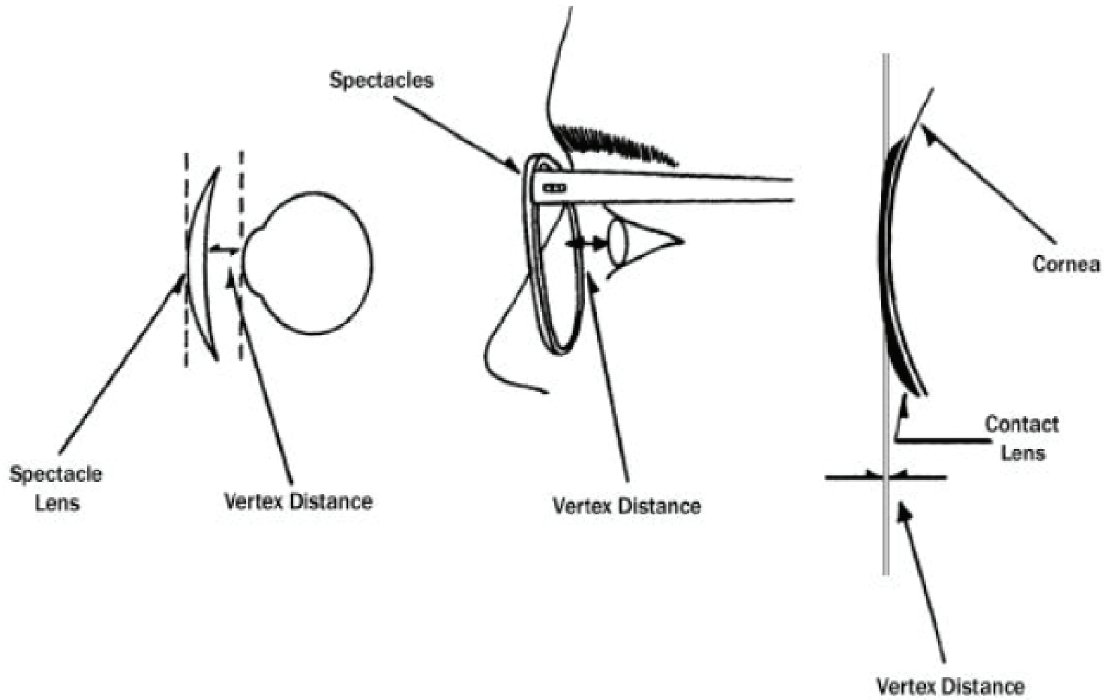


Figure 16-26. Vertex distance for spectacle lenses and contact lenses.

Nonetheless, the problem with spectacles has driven the search for other solutions to refractive error correction for pilots. Eyeglasses “present serious compatibility problems with many advanced optical systems, life support equipment, night vision or laser protective goggles, chemical protective hoods, and other personal protective gear” (Polse, 1990). In addition, there can be problems with perspiration and fogging, G-forces, reflections, seeing in foul weather, and comfort problems/hot spots when worn with helmets.

Contact lenses

To reduce the human factors issues associated with vision correction via spectacles, contact lens (CL) use and refractive surgery techniques have increased in acceptance by all of the military services. Use of CLs would appear to solve many of the problems experienced with spectacles, and, as it turns out, they do. Use of CLs to correct for refractive error solves the eye relief, spectacle comfort and reflection problems. Contact lenses are, simply put, more compatible with current optical devices like HMDs. However, CLs have their own set of problems, making their use less than universal among Warfighters.

CLs are formed, circular pieces of bell- or dome-shaped, transparent material that will maintain their shape while being held to the cornea of the eye by fluid attraction forces or the lid (Mandell, 1988). These small pieces of material vary in diameter from a little less than 7 mm to a little greater than 20 mm. A CL surface largely replaces the cornea optically providing refractive correction of the eye. Use of CLs to correct for refractive error solves the eye relief and spectacle comfort and reflection problems. The reason CLs solve the eye relief problem is because they are very thin, tenths of a millimeter, and rest on the cornea. This makes the vertex distance effectively zero (Figure 16-26). From the standpoint of eye relief, a CL on the cornea is virtually indistinguishable from the cornea without a CL. There is no frame used to support CLs, eliminating this source of discomfort and obstruction. The reflection characteristics of an in-place CL are very close to those of the natural, exposed cornea and do not provide any unusual viewing problems. There are two basic types of CLs, hard and soft.

Molded, hard plastic CLs were first made in 1938. Some were made of polymethyl methacrylate, a lens material that is still used, albeit not often. These lenses do not absorb water and are impermeable to oxygen. They rest on a layer of tears between the back surface of the CL and the corneal epithelium (the surface cells on the cornea). The hard CL moves when the wearer blinks, providing a pumping action that forces fresh, oxygenated tear between the cornea and the lens. Today, hard lenses called rigid gas permeable (RGP) CLs are more generally used. They are made from a variety of materials and, as their name suggests, are permeable to oxygen.

A soft, flexible hydrophilic CL was first conceived in Czechoslovakia in the 1950s and introduced in 1968. Since that time, many CL materials have been developed with varying flexibility, durability, water content, and oxygen transmissibility. Although soft CLs also move on the cornea, they tend to move less than hard C's and do not perform a tear pumping action. Consequently, the cornea depends, in part, on soft CL gas permeability to supply it with oxygen. Soft contact lenses have a certain water content that, along with thickness, is related to gas permeability. This is of concern, because hypoxia of the cornea, insufficient oxygen supplied to the tissue, can change its clarity and power. Hydration of the CL is also important in maintaining soft CL shape and directly related to its optical power. Environmental effects can cause changes in CL water content, particularly with hydrogel lenses (Refojo, 1991). These CLs can dehydrate in dry air until they reach equilibrium with tear absorption. If the air is very dry and the individual is in a draft, the water content at equilibrium may be too low, resulting in reduced CL performance and reduced oxygen transmission (O'Neal, 1991; Polse, 1990). Thick CLs with moderate to low water content have a slower rate of evaporation (O'Neal, 1991; Refojo, 1991). Consequently, the U.S. Air Force approved CLs are 58% water content or less.

The advantages of CL use were outlined by Crosley, Braun and Bailey (1974), Tredici and Flynn (1987), and revisited by the *Committee on Vision Commission on Behavioral and Social Sciences and Education National Research Council* (Polse, 1990) and the *Considerations in Contact Lens Use Under Adverse Conditions: Proceedings of a Symposium* (Flattau, 1991). Some of these advantages are: no interference with optical instruments (increased eye relief), increased FOV, no lens fogging, elimination of reflections from spectacle lenses, elimination of some perspiration problems, and use for treatment of specific medical/optical conditions. Tredici and Flynn (1987) went on to list 16 disadvantages, which include: CL intolerance, dislodging (for a variety of reasons, including G-force), increased chance of corneal edema, often poorer visual acuity than with spectacles, added health care burden, and difficulty of lens hygiene and professional care in the field (Table 16-13).

Even though great technical strides have been made in CL materials and design, the military has taken a very conservative stance regarding CL use in aviation, and they have done so for good reasons (Wiley, 1993). Military

Table 16-13.
Rationale for and disadvantage of contact lens use in U.S. Army aviation.
(Adapted from Crosley, Braun and Bailey, 1974; Tredici and Flynn, 1987)

RATIONALE FOR CONTACT LENS USE IN U.S. ARMY AVIATION

1. Increased field-of-vision
2. Good vision in inclement weather outside aircraft
3. No lens fogging
4. Elimination of reflections from spectacle lens
5. No interference with use of optical instruments (reduced physical eye relief)
6. Reduced perspiration problem
7. Compatibility with protective masks
8. Treatment of some medical/optical conditions

Table 16-13. (Cont.)
Rationale for and disadvantage of contact lens use in U.S. Army aviation.
(Adapted from Crosley, Braun and Bailey, 1974; Tredici and Flynn, 1987)

DISADVANTAGES OF CONTACT LENS USE IN U.S. ARMY AVIATION

1. Some individuals cannot tolerate contact lenses (newer materials have improved comfort and accommodation/adaptation)
2. Often poorer visual acuity than with spectacles
3. Lenses can be dislodged (a greater problem with hard contact lenses)
4. Bubbles can form beneath the contact lens at altitude (central vision with hard lenses, peripheral vision with soft contact lenses)
5. High G-forces can dislodge contact lenses, particularly hard lenses (a greater problem with Air Force aviation than Army aviation)
6. More difficult and time-consuming to fit than spectacles (particularly binocular and toric lenses)
7. Added health care burden (increased cost from professional fitting, follow-up, care)
8. Foreign body problems (particularly with hard contact lenses in high-particulate environments, smoke)
9. Lens hygiene and professional care difficult in field
10. Increased corneal infection risk (greatest with extended wear lenses that are necessary in the field)
11. Edema with extended wear and altitude (can reduce visual acuity and comfort)
12. Extended wear can be a problem (corneal edema, increased infection risk, comfort, etc.)
13. Can act as a sink in chemical environment (increasing toxicity, irritation, allergic reactions)
14. Allergic reactions (GPC, increased concentration of environmental allergens)

aviators must be able to perform continuously under very adverse conditions. Flight crews can be exposed to a variety of adverse environmental conditions: chemicals, dust, heat, cold, altitude changes, high and low humidity, G- forces, and adverse weather. The list is lengthy. Further, CLs may have to be cared for under very primitive conditions and worn for extended periods of time. There may not be optometrists or ophthalmologists in a field environment to care for the variety of eye problems that can arise from CL wear. Some of these conditions and some of the more general problems associated with CL wear restricts their use in the military to this day.

There are a number of excellent papers chronicling the history of CL use in military aviation Wiley, 1993; Lattimore, 1991b; Lattimore and Cornum, 1992; Tredici and Flynn, 1987). These papers give extensive reviews of the problems with CL wear in the military. The number of injuries and diseases associated with CL wear is unclear; these problems, although generally rare, include scratches and abrasions, dry eye and infection.

Soft CLs are relatively resistant to minor dust problems. However, scratches and abrasions do occur. CLs can be a barrier to some chemical exposures, but can also absorb chemicals, such as organic solvents, after a short period of exposure (Dennis et al., 1989a; Nilsson and Anderson, 1982). Consequently, chemical exposure can result in a toxic or allergy problem or simply an irritation. Even tearing of a CL can be a serious problem if it occurs at the wrong time in flight when both hands and feet are required to maintain control. As noted by the working group on Contact Lens Use Under Adverse Conditions (Polse, 1990), the cockpit (and the battlespace in general) can be a dusty and polluted environment. Although soft CLs are generally not recommended for highly polluted environments, they do seem acceptable in the cockpit (Lattimore and Cornum, 1992; Polse, 1990; Josephson, 1991; Dennis, 1988; Dennis et al., 1989b; Dennis et al., 1988; Kok-van Aalphen et al., 1985).

As a summary for the use of CLs as a potential solution to the physical eye relief problem in HMD applications, consider the statement from the working group on Contact Lens Use Under Adverse Conditions (Polse, 1990): "…helicopter personnel currently face the greatest spectacle incompatibility problems of any aviators, even as they face the greatest possible stumbling blocks to the successful use of contact lenses." CL use solves the eye relief, reflection and discomfort problems arising from spectacle use with HMDs. However, CLs do not provide a particularly good, general solution to presbyopia or astigmatism, and present new issues of their own, i.e., logistics, hygiene, use under extreme conditions, etc. At best, CLs can be used in situations where spectacles do not work well. At worst, they create more problems than they solve. In any case, they are here, probably to stay. However, refractive surgery is the latest refractive option emerging, and correction may provide an additional solution.

Refractive surgery techniques

Refractive surgery includes any procedure that surgically modifies the optical power of the human eye in order to eliminate or reduce the need for spectacles or contact lenses. In the latter half of the 20th century, the most common use of surgical intervention to change the power of the eye was the use of incisions in the cornea to correct unwanted or induced astigmatism after cataract surgery. The large peripheral corneal incisions needed to extract the cataractous lens often led to significant unequal power of the postoperative eye and a few accurate incisions in the peripheral cornea easily reduced astigmatism with minimal additional intervention. However, it was not until the advent of radial keratotomy (RK) in the 1970s that refractive surgery entered the popular mainstream. Today's arsenal of refractive surgery techniques includes everything from incisions and laser reshaping of the cornea to ocular implants. Although most techniques have been successful in reducing the individual's need for spectacles or contacts, almost all techniques have side effects that to varying degrees may affect visual performance in the operational environment.

Great leaps have been made in the technologies surrounding refractive surgery, and the outcomes have been much more precise, however, there are still problems associated with refractive surgery. Most notably, individuals may experience problems with night vision, the presence of halos or glare at night, increases in dry eye symptoms (especially after PRK or LASIK), and risks associated with surgeries that expose the eye to possible infections or reactions to agents used in the surgery (such as anesthetics). The problems with night vision, halos and glare have been mainly associated with an increase in the aberrations of the eye after refractive surgery. Aberrations due to changes in the shape of the cornea are most pronounced if the refractive correction is high, the ablation zone is small or the pupil is large (Martínez et al., 1998; Oshika et al., 1999). Aberrations are generally minimal when the refractive correction is less than 6.00 diopters of myopia or 4.00 diopters of hyperopia. Most lasers ablate a zone larger than the daylight pupil size; however, in some cases, pupil size under low light conditions may exceed the ablation area and cause visual disturbances. In a normal eye, the aberrations of the anterior surface of the cornea are balanced by opposite aberrations of the remaining refractive surfaces in the eye, including the posterior corneal surface and the crystalline lens. The anterior surface of the cornea is the primary refracting surface of the eye; therefore, modifications at this surface have the greatest effect on the quality of the image formed by the eye. If the aberration balance of the eye is modified, there are various impacts on visual performance ranging from subtle visual disturbances to severe distortions.

A significant amount of work is being done to improve the outcome of refractive surgery. One main technology has contributed towards this effort – the capability to measure the higher order aberrations of the eye. Most refractive surgery technologies have increased the basic aberration level of the eye through the induction of shape changes or a mismatch between the optics of the added components and the optics of the eye. The most promising procedures for reducing the amount of induced higher order aberrations are the corneal refractive surgery procedures or custom implants. Using a scanning spot laser, very precise ablations can be applied to the cornea in either PRK or LASIK. The problem with PRK is that the cornea undergoes a certain amount of unpredictable healing as the epithelium regrows over the corneal surface and the anterior corneal stroma responds to the laser insult. This can result in an undoing of the precise ablations and a reduction in the overall desired effect of the correction. With LASIK, the replacement of the flap over the ablated area is much like putting a thick blanket over a precisely sculpted surface; the end result is not as finely sculpted as anticipated.

The military has been a leader in studying the impact of refractive surgery techniques, especially in terms of performance under highly visually demanding conditions. Navy studies of PRK have been ongoing since 1993 and more recent efforts have been aimed towards evaluation of advanced LASIK technologies (Stanley, Tanzer and Schallhorn, 2008). Air Force efforts to evaluate refractive surgery have concentrated on determination of the effects of altitude, G-forces, and disability glare (Ivan, 2002). Studies show that moderate altitude does not cause PRK and LASIK corneas to undergo the same significant corneal thickening and curvature changes as previously seen with RK (Davidorf, 1997; van de Pol et al., 2000). Army studies have evaluated the impact of both PRK and LASIK on military operations including the helicopter flight environment (Hammond, Madigan and Bower, 2005; van de Pol et al., 2007) . Overall, study results completed by all three services show that PRK and LASIK are effective alternatives to spectacles or CLs in the aviation environment, with the caveat that not all refractive surgery procedures are appropriate for all aviation specialties. This fact is reflected in the differences in specific approved procedures from one service to the other.

User adjustments

One last but very important topic is that of the most direct interface the user has with the HMD, i.e., the controls that provide the user the ability to make adjustments to the display's characteristics. Despite the trend and the various arguments for automatic or self-adapting circuits and systems, the unique environments and situations encountered by the Warfighter, coupled with the potentially severe outcomes, argue for providing the user with the capability to make control inputs for the purpose of optimizing HMD information. Until advances in a number of scientific fields allow what would currently be considered as "futuristic" user-directed inactive control over HMD functions, (see Chapter 19, *The Potential of an Interactive HMD*), such adjustments most likely will be accomplished by hands-on controls.

On HMD devices, both monocular and binocular, there should be mechanical, electronic, or optical adjustment mechanisms available for the user to optimize the attributes of the imagery and selection of displayed information. The mechanical adjustments are used primarily to align the optical axes and exit pupils of the device to the entrance pupils and primary lines of sight of the user, if required by the inherent design. The electronic adjustments may include display brightness, contrast, electronic focus, sizing, sensor sensitivity characteristics (gain and off-set for thermal sensors), etc. The optical adjustments may include the focus adjustments for the eyepieces and sensor objective lens, and magnification selection for targeting and pilotage sensors.

Adjustment control concepts

Before discussing the various control types, there are a few higher order principles for display controls worth reviewing. First is the principle of location compatibility (or the *collocation principle* as described in Wickens and Hollands, 2000). This principle is most closely associated with the human tendency to move or orient toward a source of stimulation within the design environment. A physical interpretation of this principle is to actually position adjustment controls near the stimulus to which they are related, e.g., collocating a radio volume control on the radio itself. Touch screen controls are the ultimate realization of the collocation principle. Unfortunately, this principle is not always possible to implement, and in military vehicles or on the individual Warfighter, where space is at a premium, it is rarely achieved. Automobile designers recently have elected to ignore this principle by placing radio controls on the steering wheel, although ostensibly for safety consideration, i.e., to minimize time spent looking down away from the road.

Wickens and Hollands (1999) suggest that when the collocation principle cannot be adhered to, the consequences may be minimized by employing other compatibility principles, such as *congruence* and *rules*. Congruence is based on the concept that the spatial array of controls should have the same configuration or "be congruent with" the spatial array of the objects (stimuli) being controlled. Figure 16-27 shows the classical stovetop burner example often used to illustrate the collocation and congruent principles of control layout.

When congruence is not achievable, the designer has to fall back on a set of *rules*, a rule being a definite plan used to map controls to stimuli.

The U.S. Air Force has been exploring new spatial arrangement paradigms, along with information modality and temporal organization through the application of adaptive controls for their next-generation crew stations

(Haas et al., 2001). Their goal is to develop and evaluate interface concepts that will enhance overall performance by embedding knowledge of the Warfighter's state inside the interface, enabling the interface to make informed, automated decisions regarding many of the interface's information management display characteristics. These characteristics include information modality, spatial arrangement, and temporal organization. It is hypothesized that by increasing the ability of the interface to adapt to the changing requirements of the Warfighter in real time the interface will provide intuitive information management to the Warfighter.

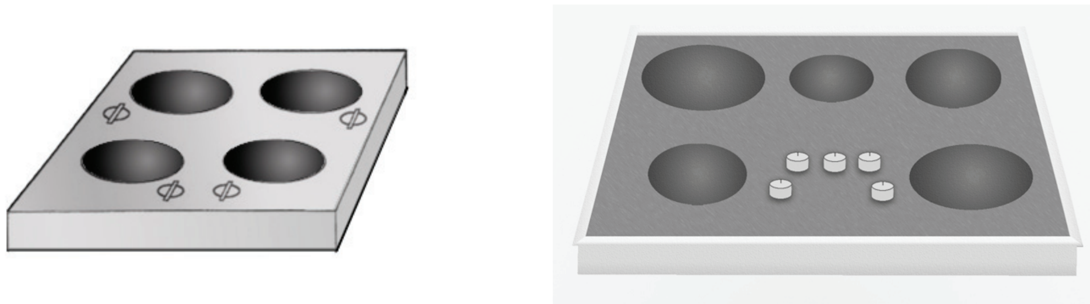


Figure 16-27. Use of stovetop burner arrays to illustrate the collocation (left) and congruent (right) principles of control layout.

A second higher order principle for display controls is *movement compatibility*. The relationship between a control movement and the effect most expected by a user population is known as a *direction-of-motion stereotype*; such a relationship is said to be compatible (Chan and Chan, 2007). Neurocognitive research has reported the strong relationship between movement observation and movement execution (Brass et al., 2000).

As an example, consider the typical *brightness* and *contrast* controls on many displays. These two controls are highly associated with adjustments of image quality and are often adjusted in a back-and-forth manner or clockwise rotational manner, which is typical of many controls (Figure 16-28).

The ISO developed a standard (ISO 9241-410:2008, *Ergonomics of Human-system Interaction – Part 410: Design Criteria for Physical Input Devices*) that specifies criteria based on ergonomics factors for the design of physical input devices for interactive systems, which includes keyboards, mice, pucks, joysticks, trackballs, track pads, tablets and overlays, touch sensitive screens, styli and light pens, and voice- and gesture-controlled devices. It provides guidance on the design of these devices, taking into consideration the capabilities and limitations of users, as well as specific criteria for each type of device.

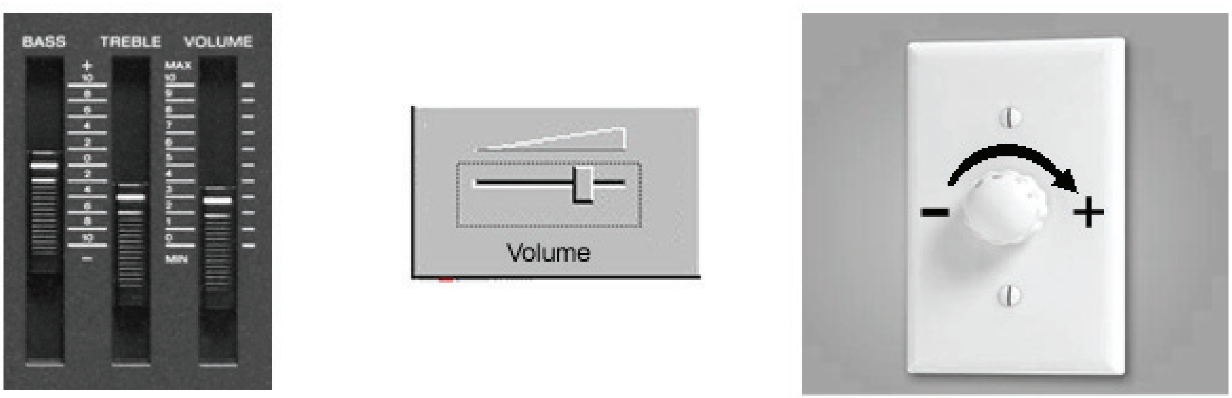


Figure 16-28. Examples of movement compatibility with various control designs: Left – Moving up increases variable, moving down decreases. Middle – Moving right increases variable, moving left decreases. Right – Turning clockwise increases variable, turning counterclockwise decreases.

Control physical design

There are a huge number of human factors and ergonomic issues associated with the physical implementation of input controls. Control device types run a wide gamut that include switches, knobs, handles, wheels, pointers, levers, trackballs, pedals, touch screens and computer mice (Sanders and McCormick, 1993). However, for HMDs, the adjustment input controls are well-defined in function and relatively narrow in selection. In the following sections, design, human factors and ergonomic issues of current HMDs are discussed.

Mechanical adjustments

Except for some early hand-held head-up displays (HUDs) used in helicopter gun ships for rocket and mini-gun alignment, fixed HUDs require no mechanical user adjustments except for seat height. For other HMD types, the mechanical adjustments may include interpupillary distance (IPD), fore-aft, vertical, tilt, roll, yaw, etc. The mechanical adjustment components may range from fine-threaded individual adjustments for one axis or plane to friction locks with ball-joints that include all axes and planes. The mechanical range of adjustments has typically been based on the 1st to 99th percentile male user.

Each potential mechanical misadjustment will affect some visual characteristic, but the adjustments are interrelated (King and Morse, 1992; McLean et al., 1997). For example, with the nonpupil-forming Aviator's Night Vision System (ANVIS), when the fore-aft adjustment is set exactly at the optimum sighting alignment point (OSAP) which is the maximum viewing distance that provides a full FOV, increasing the fore-aft distance from the eye along the optical axis proportionally decreases the ANVIS FOV (Kotulak, 1992; McLean, 1995). From the OSAP, misalignment of the IPD will decrease the FOV in the opposite direction of display movement for each ocular, thereby reducing the binocular FOV, but will not reduce the total horizontal FOV.

Misalignment of the IPD of the NVGs has been blamed for disrupting depth perception (Sheehy and Wilkinson, 1989) and inducing vergence errors (Melzer and Moffitt, 1997). However, when the eyepieces are adjusted to infinity, vergence changes do not occur (McLean et al., 1997).

For a pupil-forming system, when the pupil is moved forward or aft of the eye box that is formed around the exit pupil location along the optical axis, the FOV will be reduced. If the pupil of the eye is moved laterally from the edge of the eye box, the full FOV of the image will be extinguished within the distance of the width of the eye pupil.

For NVGs, the displacements of the right and left oculars together or relative to each other around the roll, tilt, and yaw axes will not displace the viewed image when focused at infinity, since the sensor and display are physically bound together and located near the eye. The individual FOV will be displaced in the direction of movement, but not the image. However, for HMDs with remote sensors, any relative movement between oculars around the axes will displace the images and change the convergence, divergence, or cyclo-rotation to the eyes. For the monocular HDU of the IHADSS, the mechanical adjustments are fore-aft and roll. The combiner can be moved up and down for eye alignment with the optical axis of the HDU, but most of the alignment is obtained with proper helmet fit to keep the combiner at the lowest position to obtain the maximum eye clearance and FOV. Misalignment of the HDU and IHADSS helmet outside a specific value will not allow a proper boresight with the total system.

Activation, adjustment, or movement of any mechanism on the HMD or associated instrumentation must be accomplished by the user through tactile identification and activation through any required personal protective equipment (PPE), e.g., the aviator's flight gloves, as well as, the chemical protective over-glove currently used. Removing gloves for adjustments is not a viable option.

Electronic adjustments

On present night vision imaging systems such as ANVIS, there are no user electronic adjustments provided. The tube amplification and automatic brightness control (ABC) level are set at the factory according to specifications. Since the 2^{nd} and 3^{rd} generation intensifier tubes are basically linear amplifiers with a gamma approaching unity (Allen and Hebb, 1997; Kotulak and Morse, 1994a), the imaged contrast should remain constant for changes in light level and between right and left tubes. A field study at a U.S. Army NVG training facility measured the differences in ANVIS luminance output between the right and left tubes for 20 pairs of ANVIS and found 15% of the sample had luminance differences greater than 0.1 log unit (30%) below the ABC level and none had differences greater than 0.1 log unit above the ABC level (McLean, 1997). The recent AN/PVS-14 monocular night vision device for ground troops has a user adjustable gain control, which may be incorporated in future aviation NVG designs.

For HMDs with remote sensors, both the displays in the HMD and sensor usually have user adjustments for optimization of the image. For the monocular HDU with the IHADSS, the pilot can adjust the contrast and brightness of the CRT display with the aid of a grey scale test pattern. The thermal sensors can be optimized by adjusting the gain and bias levels, where the gain refers to the range of temperatures, and the bias the average or midpoint temperature. The sensor can electronically transmit approximately 30 grey levels, where the HDU can only show about 10 grey levels (Rash, Verona, and Crowley, 1990). This means that scenes containing objects with large temperature differences would either cause loss of details from the saturation of hot objects or no contrast for cooler objects from the background. Thermal sensors are used for both pilotage and target detection. The gain and bias adjustments to optimize the contrast between the trees and sky for pilotage are considerably different than the “hot spot” technique used for the copilot/gunner for target detection. Therefore, the user will desire both manual and automatic sensor adjustment options to obtain specific information for a given scene. Thermal sensors also have an option to electronically reverse the contrast (polarity) from either white hot or black hot to either improve target detection or provide a more natural visual scene for pilotage.

Optical adjustments

For NVGs, the user has both eyepiece and objective lenses to adjust for optimum resolution. The objective lens focus is independent of the eyepiece focus and is similar to the focusing of a camera lens. The eyepiece focus adjusts the spherical lens power to compensate for the user’s refractive error (hyperopia or myopia) or induced accommodation. The standard objective lenses for ANVIS and the AN/PVS-5 NVGs adjust from approximately 10 inches (4.0 diopters) to infinity for the AN/PVS-5s and slightly beyond infinity for the ANVIS. This 4-diopter objective lens adjustment range is obtained with approximately a one-third (120-degree) rotational turn of the focusing knob. This means 1 degree of objective lens rotation equates to approximately 0.03 diopters. With the very fast objective lens for ANVIS (f#/1.2), detectable blur was found with as little as 0.05 diopter of objective lens misfocus (McLean, 1996). The latest fielded I^2 version (ANVIS-9) incorporates a fine focus objective lens where two turns (720 degrees rotation) change the focus from infinity to 1 meter (1 diopter). Objective lens focus with the ANVIS-9 or the Air Force 4949 is both more precise and much more stable during flight.

Eyepiece diopter focus: Fixed or adjustable? The most controversial subject for night imaging devices has been the eyepiece focus for I^2 devices and HMDs. Previous literature has suggested that dark focus, instrument myopia, and night myopia could play a significant part in determining the optimum lens power for night vision devices. A study by Kotulak and Morse (1994b) includes an extensive review of this literature. One group of visual scientists (Moffitt, 1991; Task and Gleason, 1993) suggests using fixed focused systems with a diopter value from 0.00 to -1.00 (infinity to 1 meter). Using aviators labeled emmetropic, other researchers have found better visual resolution with user focus adjustable eyepieces than with infinity fixed-focused eyepieces (Kotulak and Morse, 1994a; Task and Gleason, 1993). Using the most plus lens power focusing monocular technique, Kotulak and Morse (1994b) reported that 13 aviator subjects adjusted the eyepiece focus an average of -1.13 diopters (0.63

SD) with a mean difference between right and left eye focus of 0.57 diopters (0.47 SD). Using the same focusing technique with 12 subjects, Task and Gleason (1993) found an average eyepiece setting of -1.05 diopters (0.24 SD) and with a mean difference between right and left eye focus of 0.40 diopter (0.29 SD).

With the HDU monocular system of the IHADSS, Behar et al. (1990) found the average diopter eyepiece setting by 20 Apache pilots was -2.28 diopters, range 0 to -5.25 diopters. The frequently reported symptoms of asthenopia and headaches were attributed to over stimulating accommodation. [This was attributed to the failure of the IHADSS to provide a zero diopter detent or marking on the HDU focus knob.] However, CuQlock-Knopp et al. (1997) found an average diopter setting for a monocular NVG and the biocular AN/PVS-7 for 22 subjects to be 1.47 diopters and -1.54, respectively, with standard deviations of approximately 1 diopter. CuQlock-Knopp et al. (1997) also evaluated the relationship between the value of the eyepiece diopter setting and the reported eyestrain, and found no significant correlations with either the monocular or the biocular NVG.

For the classical HUD that is mounted on the glare shield and used for an aiming device, the crosshair or pipper must be collimated at infinity to retain alignment with small head and eye movements. For the monocular and binocular night imaging devices, the infinity eyepiece focus will result in some nonspectacle wearing users having less than optimum resolution. Several visual scientists (e.g., Task, Gleason, McLean, et al.) believe that some of the so called emmetropic aviators that do not wear corrective lenses are actually low myopes (-0.25 to -0.75 D) (Kotulak and Morse, 1994b) that will show reduced resolution with decreasing light levels which increase the pupil size and blur circle on the retina. The eyepiece lens power that provides most users with the best resolution with NVGs and HMDs appears to be slightly minus power between approximately -0.25 and -0.75 diopters. To ensure that optimum resolution is obtained by the aviation population of all of the nonspectacle wearing and spectacle wearing personnel using night imaging devices, a small range of adjustment would be desired, and better training in focusing procedures, to include a binocular focusing method to control accommodation with vergence. A problem found with some fixed-focused viewing devices such as the “Cats eyes NVGs” has been the ability of the factory to precisely set the eyepiece focus within a 0.12 diopter tolerance. The zero position on the diopter scale of newly received ANVIS was found to vary by up to 1.25 diopters on 10 sets of NVGs. The military specification for the zero scale tolerance for NVGs is 0.50 diopters. This would result in blurred vision for emmetropic users if the errors are on the plus lens power side. With the newer generation of image intensifiers and thermal sensors, resolution has improved to approximately 20/25 (Snellen acuity) for optimum conditions. Therefore, the focus adjustments for both the objective and eyepiece are more critical than with previous night imaging devices. Thus, a small range of user adjustable eyepiece and objective lens focus capability for the image intensifier systems and for the eyepieces of HMDs is recommended.

System integration

An HMD may be considered as a subsystem (i.e., a single component) that is intended to be used in coordination with other subsystems. The concept of *system integration* as used in this chapter is one of integrating the multiple subsystems into one system and ensuring that the subsystems function together as a system (Georgia State University, 2007). System integration issues will vary depending on the user’s functional environment.

Equipment compatibility

All HMD designs must be physically and functionally compatible with all existing mission and life support (e.g., survival) equipment. Each military branch identifies a list of equipment with which new subsystems must be compatible. Examples include corrective/protective eyewear, protective masks, oxygen masks, shoulder harnesses, survival vests, flotation equipment and components, body armor, vehicle or aircraft seat armor, and cabin interior structures and systems. The difficulty of achieving HMD-equipment compatibility is demonstrated in Figure 16-29 (left), which shows a frontal view of an Apache aviator wearing a full aviator life-support equipment (ALSE) ensemble with M-43 protective mask. Figure 16-29(right) shows the potential for interior aircraft compatibility problems by depicting an aviator in the Apache front seat with the HMD optics attached.

Figure 16-29. A frontal view of an Apache aviator wearing a full, aviator life-support equipment (ALSE) ensemble with M-43 protective mask (left) and in the Apache front seat with the HMD optics attached.

Egress[37]

In enclosed environments (e.g., ground vehicles and aircraft), emergency *egress* is considered one of the most important system integration issues. During pre- and post-crash emergency situations, the HMD user must be able to disengage from some or all components of the HMD system. In most military ground vehicles, fixed- and rotary-wing aircraft, the presence of an HMD adds another level of complexity to the Warfighter escape sequence. For ground vehicles and rotary-wing aircraft, it is essential that a quick-disconnect capability be provided. In addition, in the event that the user is unable to reach the disconnect mechanism or has insufficient time to do so, the HMD cables must be designed to provide a hands-free break-away capability (i.e., allow their breaking away by moderate brute force).

For fixed-wing aircraft, emergency egress typically involves ejection and parachuting. While the requirements for quick disconnect and an alternative hands-free break-away still must be met, this demanding scenario places additional aerodynamic requirements on the initial HMD design in order to prevent injury and HMD/helmet loss during the ejection and parachuting processes, e.g., ejection performance was a major concern during the design of the Joint Helmet-Mounted Cueing System (JHMCS). Barnaba and Kirk (1999) evaluated JHMCS performance parameters during ejection that included structural integrity, facial and head protection, neck tensile loads, ejection seat and crew equipment compatibility, and mechanical functionality.

Summary

The major goal of this chapter is to make HMD designers and users aware of the difficult and demanding environment in which HMDs must operate. Paper designs and laboratory prototypes must take into consideration the multitude of operational factors with which the HMD user must contend. These factors range widely in type and scope. Whether environmental (external) or self-imposed (internal) in nature, they invariably affect human performance. Technology, no matter how great, is only as good as its effectiveness in the hands (or on the heads) of the user in the actual operating environment.

[37] The terms *ingress* and *egress* are defined as entering and exiting, respectively.

References

Abey-Wickrama, I., A'Brook, M.F., Gattoni, F.E.G., and Herridge, C.F. (1969). Mental-hospital admissions and aircraft noise, *Lancet*, 2, 1275-1277.

Acheson, S.K., Stein R.M., and Swartzwelder H.S. (1998). Impairment of semantic and figural memory by acute ethanol: Age-dependent effects, *Alcoholism, Clinical and Experimental Research*, 22(7), 1437-1442.

Ahroon, W.A., Gordon, E., and Brozoski, F.T. (2002). Impulse noise levels of the UH-6OA/L Black Hawk helicopter cockpit air bag system - enhanced lateral air bags. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2002-16.

Akerstedt, T. (1995). Work hours, sleepiness and the underlying mechanisms, *Journal of Sleep Research,* 4(Suppl. 2), 15-22.

Akerstedt, T. (1988). Sleepiness as a consequence of shift work, *Sleep*, 11(1), 17-34.

Akerstedt, T. (2000). Consensus statement: Fatigue and accidents in transport operations, *Journal of Sleep Research,* 9, 395.

Akerstedt, T., and Ficca, G. (1997). Alertness-enhancing drugs as a countermeasure to fatigue in irregular work hours, *Chronobiology International*, 14(2), 145-158.

Alem, N.M., Shanahan, D.F., Barson, J.V., and Muzzy, III, W.H. (1991a). The airbag as a supplement to standard restraint systems in the AH-1 and AH-64 attack helicopters and its role in reducing head strikes of the copilot/gunner, volume I. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 91-6.

Alem, N.M., Shanahan, D.F., Barson, J.V., and Muzzy, III, W.H. (1991b). The airbag as a supplement to standard restraint systems in the AH-1 and AH-64 attack helicopters and its role in reducing head strikes of the copilot/gunner, Volume II. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 91-6a.

Allain, H., Bentue-Ferrer, D., Le Breton, S., Polard, E., and Gandon, J. (2003). Preference of insomniac patients between a single dose of zolpidem 10 mg versus zaleplon 10 mg, *Human Psychopharmacology: Clinical and Experimental*, 18(5), 369-374.

Allen, J.H., and Hebb, R. C. 1997. Determining the gamma of a night vision device. Orlando, FL: Naval Air Warfare Center. Report No. NAWCTSD-TR-95-003.

Allen, M.J., Abrams, B.S., Ginsburg, A.P., and Weintraub, L. (2001). *Forensic Aspects of Vision and Highway Safety*. Tucson, AZ: Lawyers and Judges Publishing Company.

American Conference of Governmental Industrial Hygienists. (2001). ACGIH-HAV, Hand-arm vibration standard. Cincinnati, OH: American Conference of Governmental Industrial Hygienists.

American Conference of Governmental Industrial Hygienists. (2001). Whole-body vibration: TLV physical agents. Cincinnati, OH: American Conference of Governmental Industrial Hygienists.

American Conference of Governmental Industrial Hygienists (2005). Whole-body vibration. In: *Threshold limit values, TLVs and biological exposure indices for 2005*. Cincinnati, OH: American Conference of Governmental Industrial Hygienists.

American National Standards Institute. (2002). ANSI S3.18, Mechanical vibration and shock – Evaluation of human exposure to whole-body vibration. American National Standards Institute.

American National Standards Institute. (2003). ANSI Z87.1, Occupational and Educational Personal Eye and Face Protection Devices. American National Standards Institute.

American National Standards Institute. (1986). ANSI S3.34, Human Exposure to Vibration Transmitted to the Hand, Guide for Measurement and Evaluation of. American National Standards Institute.

American National Standards Institute. (2006). ANSI S2.70-2006: Guide for the Measurement and Evaluation of Human Exposure to Vibration Transmitted to the Hand. (Revision/replacement of ANSI S3.34-1986)

American Sleep Disorders Association. (1994). *Coping with Shiftwork*. Rochester, MN.

Anderson, S.J., and Holliday, I.E. (1994). Night driving: Effects of glare from vehicle headlights on motion perception, *Ophthalmology and Physiological Optics*, 15, 545-551.

Applegate, R.A. (1989). Disability glare and hydrogel lens wear revisited, *Optometry and Vision Science*, 66(11), 756-759.

Applegate, R.A., and Wolf, M. (1987). Disability glare increased by hydrogel lens wear, *American Journal of Optometry and Physiological Optics*, 64(5), 309-312.

Arendt, J. (1999). Jet-lag and shift work: Therapeutic use of melatonin, *Journal of the Royal Society of Medicine*, 92, 402-5.

Arendt, J., and Deacon, S. (1997). Treatment of circadian rhythm disorders – Melatonin, *Chronobiology International*, 14, 185-204.

Arendt, J., and Marks, V. (1982). Physiological changes underlying jet lag, *British Medical Journal*, 284, 144-46.

Arendt, J., Deacon, S., English, J., Hampton, S., and Morgan, L. (1995). Melatonin and adjustment to phase shift, *Journal of Sleep Research*, 4(Suppl. 2), 74-79.

Arendt, J., Skene, D.J., Middleton, B., Lockley, S.W., and Deacon, S. (1997). Efficacy of melatonin treatment in jet lag, shift work, and blindness, *Journal of Biological Rhythms*, 12, 604-17.

Armstrong, S.M. (1999). Melatonin as a chronobiotic for circadian insomnia: Clinical observations and animal models, *Advanced Experiments in Medical Biology*, 460, 283-297.

Armstrong, S.M, Cassone, V.M., Chesworth, M.J., Redman, J.R., and Short, R.V. (1986). Synchronization of mammalian circadian rhythms by melatonin, *Journal of Neural Transmission*, 21(Suppl.), 375-394.

Artal, P., Berrio, E., and Guirao, A. (2002). Contribution of the cornea and internal surfaces to the change of ocular aberrations with age, *Optical Society of America*, 19(1), 137-143.

Artal, P., Guirao., A, Berrio, E., Piers, P., and Norrby, S. (2003). Optical aberrations and the aging eye, *International Ophthalmology Clinics*, 43(2), 63-77.

Aschoff, J., and Wever, R. (1981). The circadian system of man. In: Aschoff, J. (Ed.), *Handbook of Behavioral Neurobiology, Volume 4, Biological Rhythms*. New York, USA: Plenum Press.

Atkinson, G., Drust, B., Reilly, T., and Waterhouse, J. (2003). The relevance of melatonin to sports medicine and science, *Sports Medicine*, 33, 809-31.

Atchison, D.A., and Smith, G. (2000). *Optics of the Human Eye*. Oxford: Butterworth-Heinemann.

Axelsson, A., and Prasher, D.K. (1999). Tinnitus: A warning signal to teenagers attending discotheques? *Noise and Health*, 2, 1-2.

Bailey, M.D., Mitchell, G.L., Dhaliwal, D.K., Boxer Wachler, B.S., and Zadnik, K. (2003). Patient satisfaction and visual symptoms after laser in situ keratomileusis, *Ophthalmology*, 110(7), 13711378.

Baker, H.D. (1953). The instantaneous threshold and early dark adaptation, *Journal of the Optical Society of America*, 43(9), 798-803.

Baker, H.D. (1949). The course of foveal light adaptation measured by the threshold intensity increment, *Journal of the Optical Society of America*, 39(2), 172-179.

Baldeweg, T., Wong, D., and Stephan, K.E. (2006). Nicotinic modulation of human auditory sensory memory: Evidence from mismatch negativity potentials *International Journal of Psychophysiology*, 59(1), 49-58.

Balter, M.B., and Uhlenhuth, E.H. (1992). New epidemiologic findings about insomnia and its treatment, *Journal of Clinical Psychiatry*, 53(12 Suppl.), 34-39.

Banderet, L., and Burse, R. (1988). Effects of high terrestrial altitude on military performance. In: Gal, R., and Mangelsdorff, D. (Eds.), *Handbook of Military Psychology*. New York: John Wiley and Sons.

Banderet, L., and Shukitt-Hale, B. (2002). Cognitive performance, mood, and neurological status at high terrestrial elevation. In: Lounsbury, D.E., Bellamy, R.F., and Zajtchuk, R. (Eds.), *Medical Aspects of Harsh Environments.* Washington, DC: Office of the Surgeon General, Borden Institute.

Banerjee, D., Vitiello, M.V., and Grunstein, R.R. (2004). Pharmacotherapy for excessive daytime sleepiness, *Sleep Medicine Review*, 8, 339-354.

Barnaba, J.M., and Kirk, W.F. (1999). Evaluation of ejection safety for the Joint Helmet-Mounted Cueing System, *Proceedings of SPIE, Helmet- and Head-Mounted Displays IV*, 3689, 210-219.

Baranski, J.V., Gil, V., McLellan, T.M., Moroz, D., Buguet, A., and Radomski, M. (2002). Effects of modafinil on cognitive performance during 40 hr of sleep deprivation in a warm environment, *Military Psychology*, 14(1), 23-47.

Baranski, J.V., and Pigeau, R.A. (1997). Self-monitoring cognitive performance during sleep deprivation: Effects of modafinil, d-amphetamine, and placebo, *Journal of Sleep Research*, 6, 84-91.

Barbur, J.L (2004). (Personal correspondence).

Barlow, H.B. (1972). Dark and light adaptation: Psychophysics. In: Jameson, D., and Hurvich, L.M. (Eds.), *Visual Psychophysics*. New York: Springer-Verlag.

Bartholini, G. (1988). Growing aspects of hypnotic drugs. In: Sauvanet, J.P., Langer, S.Z., and Morselli, P.L. (Eds.), *Imidazopyridines in Sleep Disorders*. New York: Raven Press. p. 1-9.

Bartlett, N.R. (1966). Dark adaptation and light adaptation (Chapter 8). In: Graham, C.H. (Ed), *Vision and Visual Perception*. New York: John Wiley and Sons.

Batchelor, A.J. (2006). Medication and aircrew. In: Rainford, D.J., and Gradwell, D.P (Eds.), *Ernsting's Aviation Medicine***.** London: Hodder Arnold.

Belenky, G., Wesensten, N.J., Thorne, D.R., Thomas, M.L., Sing, H.C., Redmond, D.P., Russo, M.B., and Balkin, T.J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study, *Journal of Sleep Research*, 12(1), 1-12.

Belyavin, A., and Wright, N.A. (1987). Changes in electrical activity of the brain with vigilance, *Electroencephalography and Clinical Neurophysiology*, 66, 137-144.

Benjamin, W.J. (Ed.) (1998). *Borish's Clinical Refraction*. Philadelphia, PA: W.B. Saunders Company.

Benson, A.J., and Barnes, G.R. (1978). Vision during angular oscillation: The dynamic visual and vestibular mechanisms, *Aviation, Space and Environmental Medicine*, 49, 340-345.

Berglund, B., and Hassmén, P. (1996). Sources and effects of low-frequency noise, *Journal of the Acoustical Society of America*, 99(5), 2985-3002.

Bernaards, C.M., Twisk, J.W., van Mechelen, W., Snel, J., and Kemper, H.C. (2003). A longitudinal study on smoking in relationship to fitness and heart rate response, *Medicine and Science in Sports and Exercise*, 35(5), 793-800.

Biberman, L.M., and Tsou, B. (1991). Image display technology and problems with emphasis on airborne systems. Alexandria, VA: Institute for Defense Analysis. IDA Paper P-2448.

Bichao, I.C., Yager, D., and Meng, J. (1995). Disability glare: Effects of temporal characteristics of the glare source and of the visual field location of the test stimulus, *Journal of the Optical Society of America*, A12(10), 2252-2258.

Blois, R., Gaillard, J., Attali, P., and Coqueline, J. (1993). Effect of zolpidem on sleep in healthy subjects: A placebo controlled trial with polysomnographic recordings, *Clinical Therapeutics*, 15(5), 797-809.

Blomqvist, G., Saltin, B., Mitchell, J.E., and Vastagh, G.F. (1970). Acute effects of ethanol ingestion on the response to submaximal and maximal exercise in man, *Circulation*, 42, 463-470.

Boff, K.R., and Lincoln, J.E. (1988). Vibration characteristics of rotary-wing aircraft. *Engineering Data Compendium: Human Perception and Performance*. Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory. p. 2072-2136.

Bohnker, B.K., Page, J.C., Rovig, G., Betts, L.S., and Sack, D.M. (2002). Navy Hearing Conservation Program: threshold shifts in enlisted personnel, 1995-1999, *Military Medicine*, 167(1), 48-52.

Bongers, P.M., Hulshof, C.T., Dijkstra, L., Boshuizen, H.C., Groenhout, H.J., and Valken, E. (1990). Back pain and exposure to whole body vibration in helicopter pilots, *Ergonomics*, 33(8),1007-1026.

Bonnet, M.H. (1994). Sleep deprivation. In: Kryger, M.H., Roth, T., and Dement, W.C., (Eds.), *Principles and practice of sleep medicine* (2nd Ed.). Philadelphia: W.B. Saunders Company.

Bonnet, M.H., Gomez, S., Wirth, O., and Arand, D.L. (1995). The use of caffeine versus prophylactic naps in sustained performance, *Sleep*, 18(2), 97-104.

Bourgeois-Bougrine, S., Carbon, P., Gounelle, C., Mollard, R., and Coblentz, A. (2003). Perceived fatigue for short- and long-haul flights: A survey of 739 airline pilots, *Aviation Space and Environmental Medicine*, 74, 1072-77.

Boynton, R.M., Bush, W.R., and Enoch, J.M. (1954). Rapid changes in foveal sensitivity resulting from direct and indirect adapting stimuli, *Journal of the Optical Society of America*, 44(1), 56-60.

Brass, M., Bekkering, H., Wohlschlager, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial and imitative cues, *Brain and Cognition*, 44(2), 124-143.

Bricknell, M.C.M. (1991). Sleep manipulation prior to airborne exercises, *Journal of the Regimental Army Medical Corps*, 137, 22-26.

Broadbent, D.E. (1971). *Decision and Stress.* New York: Academic Press.

Bron, A.J., Tripathi, R.C., and Tripathi, B.J. (1997). *Wolff's Anatomy of the Eye and Orbit. 8th Edition*. London: Chapman and Hall Medical.

Brown, J.L. (1966). Afterimages. In: Graham, C.H. (Ed.), *Vision and Visual Perception*. New York: John Wiley and Sons, Inc.

Brown, E.L. (2004). The effects of fatigue on position determination and cognitive workload using a visual and 3-dimensional auditory display. Monterey, CA: Naval Postgraduate School.

Brozoski, F.T., Johnson, P.A., Crowley, J.S., and McEntire, B.H. (2000). Effects of inadvertent UH-60 cockpit airbag system deployment on flight control. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2000-15.

Brunette, I, Gresset, J., Boivin, J., Pop, M, Thompson, P, Lafond, G.P., and Makni, H. (2000). Functional outcome and satisfaction after photorefractive keratectomy, *Ophthalmology*, 107(9), 1790-1796.

Buguet, A., Montmayeur, A., Pigeau, R.A., and Naitoh, P. (1995). Modafinil, d-amphetamine, and placebo during 64 hours of sustained mental work. II. Effects on two nights of recovery sleep, *Journal of Sleep Research*, 4, 229-241.

Butler, B.P., and Alem, N.M. (1997). Long duration exposure criteria for head-supported mass. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 97-34.

Cabon, P., Bourgeois-Bougrine, S., Mollard, R., Coblentz, A., and Speyer, J.J. (2003). Electronic pilot-activity monitor: A counter measure against fatigue on long-haul flights, *Aviation Space and Environmental Medicine*, 74(6 Pt 1), 679-82.

Cabon, P.H., Coblentz, A., Mollard, R., and Fouillot, J.P. (1993). Human vigilance in railway and long-haul flight operations, *Ergonomics*, 36(9), 1019-1033.

Cajochen, C., Krauchi, K., and Wirz-Justice, A. (2003). Role of melatonin in the regulation of human circadian rhythms and sleep, *Journal of Neuroendocrinology*, 15, 432-7.

Caldwell, J.A. (2001). Efficacy of stimulants for fatigue management: The effects of Provigil and Dexedrine on sleep-deprived aviators, *Transportation Research*, F4, 19-37.

Caldwell, J.A. (2005). Fatigue in aviation, *Journal of Travel Medicine and Infectious Disease*, 3(2), 85-96.

Caldwell, J.A., and Caldwell, J.L. (1997). An in-flight investigation of the efficacy of dextroamphetamine for sustaining helicopter pilot performance, *Aviation, Space, and Environmental Medicine*, 68(12), 1073-1080.

Caldwell, J.A., and Caldwell, J.L. (1998). Comparison of the effects of zolpidem-induced prophylactic naps to placebo naps and forced-rest periods in prolonged work schedules, *Sleep*, 21(1), 79-90.

Caldwell, J.A., and Caldwell, J.L. (2005). Stimulants. In: Kushida, C.A. (Ed.), *Sleep Deprivation: Clinical Issues, Pharmacology, and Sleep Loss Effects*. New York: Marcel Dekker, Inc.

Caldwell, J.A., and Caldwell, J.L. (2003). *Fatigue in aviation: A guide to staying awake at the stick*. Burlington, VT: Ashgate Publishing Company.

Caldwell, J.A., Caldwell, J.L., and Crowley, J.S. (1997). Sustaining female helicopter pilot performance with Dexedrine during sustained operations, *International Journal of Aviation Psychology*, 7(1), 15-36.

Caldwell, J.A., Caldwell, J.L., Crowley, J.S., and Jones, H.D. (1995). Sustaining helicopter pilot performance with Dexedrine during periods of sleep deprivation, *Aviation Space and Environmental Medicine*, 66(10), 930-937.

Caldwell, J.A., Caldwell, J.L., Smith, J.K., Alvarado, L., Heintz, T., Mylar, J.T., Mandichak, M., and Schroeder, C. (2004). *The efficacy of modafinil for sustaining alertness and simulator flight performance in F-117 pilots during 37 hours of continuous wakefulness*. Technical. Brooks City-Base, TX: U.S. Air Force Research Laboratory; January. Report No. AFRL-HE-BR-TR-2004-0003.

Caldwell, J.A., Caldwell, J.L., Smythe, N.K., and Hall, K.K. (2000a). A double-blind, placebo-controlled investigation of the efficacy of modafinil for sustaining alertness and performance of aviators: A helicopter simulator study, *Psychopharmacology* (Berlin), 150, 272-282.

Caldwell, J.L., and Gilreath, S.R. (2001). Work and sleep hours of U.S. Army aviation personnel working reverse cycle, *Military Medicine*, 166(2), 159-166.

Caldwell, J.A., and Gilreath, S.R. (2002). A survey of aircrew fatigue in a sample of Army aviation personnel, *Aviation, Space, and Environmental Medicine*, 73, 472-480.

Caldwell, J.L., Prazinko, B.F., Rowe, T., Norman, D., Hall, K.K., and Caldwell, J.A. (2003). Improving daytime sleep with temazepam as a countermeasure for shift lag, *Aviation Space and Environmental Medicine*, 74(2),153-163.

Caldwell, J.A., Smythe, N.K., LeDuc, P.A., and Caldwell, J.L. (2000b). Efficacy of dextroamphetamine for the maintenance of aviator performance during 64 hours of sustained wakefulness, *Aviation Space and Environmental Medicine*, 71(1), 7-18.

Cardinali, D.P., Bortman, G.P., Liotta, G., Albornoz, L.E. ,Cutrera, R.A., Batista, J., and Gallo, P.O. (2002). A multifactorial approach employing melatonin to accelerate resynchronization of sleep- wake cycle after a 12 time-zone westerly transmeridian flight in elite soccer athletes, *Journal of Pineal Research*, 32, 41-46.

Cardinale, M., and Pope, M.H. (2003). The effects of whole-body vibration on humans: Dangerous or advantageous? *Acta Physiologica Hungarica*, 90(3), 195-206.

Center for Science in the Public Interest (1996). *Caffeine: The inside scoop*. In: Nutrition Action Health Letter, Dec. Washington DC: Center for Science in the Public Interest.

Centers for Disease Control and Prevention. (2004). 2004 Surgeon General's report – The health consequences of smoking. Retrieved from: http://www.cdc.gov/ tobacco/data_statistics/sgr/sgr_2004/chapters.htm

Cephalon (1998). *Clinical Investigator's Brochure*. West Chester: Cephalon, Inc.

Chagan, L., and Cicero, L.A. (1999). Zaleplon: A possible advance in the treatment of insomnia, *PandT*, 24(12), 590-599.

Chalita, M.R., and Krueger, R.R. (2004). Correlation of aberrations with visual acuity and symptoms, *Ophthalmology Clinics of North America*, 17, 135-142.

Chalita, M.R., Chavala, S., Xu, M., and Krueger, R.R. (2004). Wavefront analysis in post-LASIK eyes and its correlation with visual symptoms, refraction, and topography, *Ophthalmology*, 111(3), 447-453.

Chan, W.H., and Chan, A.H.S. (2007). Movement compatibility for rotary control and digital display, *Engineering Letters*, 14(1) Retrieved from; http://www.engineeringletters.com/issues_v14/issue_1/EL_14_1_2.pdf)

Chee, P.A., Braby, L.A., and Conroy, T.J. (2000). Potential doses to passengers and crew during supersonic transport, *Health Physics*, 79, 547-553.

Chisholm, C.M., Evans, A.D.B., Harlow, J.A., and Barbur, J.L. (2003). New test to assess pilot's vision following refractive surgery, *Aviation, Space, and Environmental Medicine*, 74(5). 551-559.

Chisick, M.C., Poindexter, F.R., and York, A.K. (1998). *Comparing tobacco use among incoming recruits and military personnel on active duty in the United States*. Tobacco Control, 7, 236-240

Cho, K, (2001), Chronic 'jet lag' produces temporal lobe atrophy and special cognitive deficits, *Natural Neuroscience*, 4, 567-68.

Cho, K., Ennaceur, A., Cole, J.C., and Suh, C.K. (2000). Chronic jet lag produces cognitive deficits, *Journal of Neuroscience*, 20(RC66), 1-5.

Choi, S.H. (1983). The effect of noise and display orientation on cognitive performance. Master's Thesis. Monterey, CA: Naval Postgraduate School.

Chou, B., and Wachler, B.S.B. (2001). The role of pupil size in refractive surgery, *Refractive Source*. Retrieved from: http://www.refractivesource.com/

Ciuffreda, K.J. (1998). Accommodation, the pupil, and presbyopia. Chapter 4. In: Benjamin, W.J. (Ed.), *Borish's Clinical Refraction*. Philadelphia: W. B. Saunders Company.

Claustrat, B., Kayumov, L., and Pandi-Perumal, S.R. (2002). Dual role of melatonin in humans: A hormone and a chronobiotic drug. In: Haldar, C., Singaravel, M., and Maitra, S.K. (Eds.), *Treatise on Pineal Gland and Melatonin* (pp. 543-553). Enfield, NH: Science Publishers, Inc.

Co, E.L., Gregory, K.B., Johnson, J.M., and Rosekind, M.R. (1999). *Crew factors in flight operations XI: A survey of fatigue factors in regional airline operations*. NASA Technical Memorandum No. 1999-208799. Moffett Field, CA: NASA Ames Research Center.

Cobb, P.W. (1911). The influence of illumination of the eye on visual acuity, *American Journal of Physiology*, 29, 76-99.

Cohen, S., Evans, G.W., Krantz, D.S., and Stokols, D. (1986). Physiological, motivational, and cognitive effects of aircraft noise on children: Moving from the laboratory to the field, *American Psychologist*, 35, 231-243.

Cohen, S., Evans, G.W., Stokols, D., and Krantz, D.S. (1980). *Behavior, Health, and Environmental Stress*. New York, NY: Plenum Press.

Coleshaw, S.R., Van Someren, R.N., Wolff, A.H., Davis, H.M., and Keatinge, W.R. (1983). Impaired memory registration and speed of reasoning caused by low body temperature, *Journal of Applied Physiology*, 55, 27-31.

Committee on Military Nutrition Research (2002). *Caffeine for the Sustainment of Mental Task Performance: Formulations for Military Operations*. Washington, DC: National Academy Press.

Connolly, D.M., Barbur, J.L., Hosking, S.L., and Moorhead, I.R (2008). Mild hypoxia impairs chromatic sensitivity in the mesopic range, *Investigative Ophthalmology and Visual Sciences*, 49(2), 820-827.

Conway, T.L. (1998). *Tobacco use and the United States military: A longstanding problem* [Editorial]. Tobacco Control, 7, 219-221.

Cornum, R., Caldwell, J.A., and Cornum, K.G. (1997). Stimulant use in extended flight operations, *Airpower Journal*, 11(1), 53-58.

Cornum, K.G., Cornum, R., and Storm, W. (1995). *Use of psychostimulants in extended flight operations: A Desert Shield experience*. In: Advisory Group for Aerospace Research and Development Conference Proceedings No. 579, Neurological Limitations of Aircraft Operations: Human Performance Implications. Nuilley-sur-Seine, France: NATO Advisory Group for Aerospace Research and Development; p. 371-374.

Costa, G. (1997). The problem: Shiftwork, *Chronobiology International*,14(2), 89-98.

Crosley, J.K., Braun, E.G. and Bailey, R.W. 1974. Soft (hydrophilic) contact lenses in U.S. Army aviation: An investigative study of the Bausch and Lomb Soflens, *American Journal of Optometry and Physiological Optics*, 51(7), 470-477.

Crowley, J.S., and Dalgard, C.L. (Eds.) (2000). Proceedings of the Technical Cooperative Program Workshop: Inflatable restraints in aviation. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2000-21.

Crowley, J.S., Wesensten, N., Kamimori, G., Devine, J., Iwanyk, E., and Balkin, T. (1992). Effect of high terrestrial altitude and supplemental oxygen on human performance and mood, *Aviation Space and Environmental Medicine*, 63, 696-701.

Cruickshanks, K.J., Klein, R., Klein, B.E.K., Wiley, T.L., Nondahl, D.M., and Tweed, T.S. (1998). The epidemiology of hearing loss study, *Journal of the American Medical Association*, 279, 1715-1719.

CuQlock-Knopp, V.G., Sipes, D.E., Torgerson, W., Bender, E., Merritt, J.O., McLean, W., and Myles, K. (1997). Extended use of night vision goggles: An evaluation of comfort for monocular and biocular configurations. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory. Report No. ARL-TR-1427.

Curley, M.D., and Hawkins, R.N. (1983). Cognitive performance during a heat acclimatization regimen, *Aviation, Space, and Environmental Medicine*, 54, 709-713.

Cymerman, A., Fulco, C.S., Muza, S.R., Zupan, M., Kambis, K.W., Elliot, L., Udy, J., and Rock, P.B. (2005). Neurocognitive function in moderate altitude residents (2200 m) exposed for 3 days to 4300 meters (Abstract), *Banff, 10th International Symposium*.

Cymerman, A., Muza, S.R., Beidleman, B.A., Ditzler, D., and Fulco, C.S. (2006a). Postural stability and acute mountain sickness during exposure to 24 hours of simulated altitude (4300 m), *High Altitude Medicine and Biology*, 2, 509-514.

Cymerman, A., Muza, S., Fulco, C., Beidleman, B., Elliott, L., Elliott, S., Tapia, M., Money, A., and Root, E. (2007). Normobaric intermittent hypoxic exposures do not improve psychomotor performance during hypobaric hypoxia, *High Altitude Medicine and Biology*, 7(4), 324.

Cymerman, A., Muza, S. R., Fulco, C. S., Friedlander, A., Zupan, M., and Elliott, L. (2006b). Oculometrics as a measure of altitude acclimatization of moderate-altitude residents compared to sea-level residents (Abstract), *Journal of the Federation of American Societies for Experimental Biology*, 20, A1431.

Czeisler, C.A., Johnson M.P., Duffy, J.F., Brown, E.N., Ronda, J.M., and Kronauer, R.E. (1990). Exposure to bright light and darkness to treat physiologic maladaptation to night work, *The New England Journal of Medicine*, 322(18), 1253-1259.

Dawson, D., and Armstrong, S.M. (1996). Chronobiotics – drugs that shift rhythms, *Pharmacology and Therapeutics*, 69, 15-36.

Dawson, D. (2008). Army offers choices in protective gloves and eyewear. Retrieved 30 October 2008 from: http://peosoldier.army.mil/docs/gloveseyewear08.pdf

Dawson, D., and Reid, K. (1997). Fatigue, alcohol and performance impairment, *Nature*, 388, 235.

de Waard, P.W.T., Ijspeert, J.K., van den Berg, T.J.T.P., and de Jong, T.V.M. (1992). Intraocular light scattering in age-related cataracts, *Investigative Ophthalmology and Visual Science*, 33(3), 618-625.

de Wit, G.C., and Coppens, J.E. (2003). Stray light of spectacle lenses compared with stray light in the eye, *Optometry and Vision Science*, 80(5), 395-400.

Dehart, R L, and Davis, J.R. (Eds.) (2002). *Fundamentals of Aerospace Medicine*. Philadelphia: Lipcott Williams and Wilkins.

Dennis, R.J. (1988). Wear of soft contact lenses by tactical aircrew. Considerations in contact lens use under adverse conditions - a symposium. Brooks AFB, TX.

Dennis, R.J., Apsey, D.A., and Ivan, D.J. (1993). Aircrew soft contact lens wear: A survey of USAF eye care professionals, *Aviation Space and Environmental Medicine*, 64, 1044-1047.

Dennis, R.J., Flynn, W. J., Oakley, C. J., and Block, M. G. (1989b). Soft contact lenses: Sink or barrier to chemical warfare agents? Brooks AFB, TX: USAF School of Aerospace Medicine. USAFSAM-TR-89-2.

Dennis, R.J., Woessner, W.M., Miller, R.E., and Gillingham, K.K. (1989a). Effect of fluctuating +G sub z exposure on rigid gas-permeable contact lens wear. Brooks AFB, TX :U.S. Air Force School of Aerospace Medicine Technical Report No. USAFSAM-TR-89-23.

Department of the Air Force. (2006). Air Force Instruction AF48-123, Medical Examinations and Standards. Retrieved 20 September from: http://www.e-publishing.af.mil/index.asp

Department of the Army. (2006). Aeromedical Policy Letters. Retrieved 28 August 2008 from: https://aamaweb.usaama.rucker.amedd.army.mil/AAMAWeb/ policyltrs/Army_APLs_Mar06_v3.pdf

Department of the Army. (2007). Army Regulation AR 40-8, Temporary Flying Restrictions Due to Exogenous Factors Affecting Aircrew Efficiency. Retrieved 1 September 2008 from: http://www.army.mil/ usapa/epubs/pdf/r40_8.pdf

Department of the Army. (1991). Army Regulation AR-40-10. Health hazard assessment program in support of the Army material acquisition decision. Washington, DC: Department of the Army.

Department of the Army. (2008). Army Regulation AR-40-501, Standards of Medical Fitness. Retrieved 10 September 2008 from: http://www.usapa.army.mil/ pdffiles/r40_501.pdf

Department of Defense. (2005). Department of Defense survey of health related behaviors among active duty military personnel; A component of the defense lifestyle assessment program (DLAP). RTI International, December 2006. Retrieved from: http://www.ha.osd.mil/special_reports/ 2005_Health_Behaviors_Survey_1-07.pdf

Department of the Navy. (2001). OPNAV Instruction 3710.7T, NATOPS General Flight and Operating Instructions. http://doni.daps.dla.mil/Directives/03000%20Naval%20Operations%20and%20Readiness/03-700 % 20Flight%20 and%20 Air%20Space%20Support%20Services/3710.7T-1.pdf

Department of Transport South Africa. (2007). Arrive alive…don't drink and drive. Retrieved 10 August 2008 from: http://www.transport.gov.za/projects/arrive/alcohol98.html

Dinges, D.F. (1992). Probing the limits of functional capability: The effects of sleep loss on short-duration tasks. In: Broughton, R.J., and Ogilvie, R.D. (Eds.), *Sleep, arousal, and performance*. Boston, Birkhauser.

Dinges, D.F. (1995). An overview of sleepiness and accidents, *Journal of Sleep Research*, 4(Suppl. 2), 4-14.

Dinges, D.F., Graeber, R.C., Rosekind, M.R., Samel, A., and Wegmann, H.M. (1996). *Principles and guidelines for duty and rest scheduling in commercial aviation*. Moffett Field, CA: NASA Ames Research Center, Report No. 110404.

Dinges, D.F., and Kribbs, N.B. (1991). Performing while sleepy: Effects of experimentally-induced sleepiness. In: Monk, T.H. (Ed.), *Sleep, sleepiness and performance* (pp. 97-128). Chichester: John Wiley and Sons.

Dinges, D.F., and Powell, J.W. (1988). Sleepiness is more than lapsing. *Sleep Research*, 17, 84.

Dyer, F.N. (1986). Smoking and soldier performance. Fort Rucker, AL: U.S. Army Aeromedical research Laboratory. USAARL Report No. 86-13.

Edwards, R.J., Sanders, M.G., and Price, D.R. (1988). Comparison of Army flight school performance in smokers and nonsmokers. Fort Rucker, AL: U.S. Army Aeromedical research Laboratory. USAARL Report No. 88-6.

Elie, R., Ruther, E., Farr, I., Emilien, G., and Salinas, E. (1999). Sleep latency is shortened during 4 weeks of treatment with zaleplon, a novel nonbenzodiazepine hypnotic, *Journal of Clinical Psychiatry*,*60*(8), 536-544.

Elliott, D.B., Mitchell, S., and Whitaker, D. (1991). Factors affecting light scatter in contact lens wearers, *Optometry and Vision Science*, 68(8), 629-633.

Emonson, D.L., and Vanderbeek, R.D. (1995). The use of amphetamine in U.S. Air Force tactical operations during Desert Shield and Storm, *Aviation Space and Environmental Medicine*, *66*, 260-263.

Enander, A.E., and Hygge, S. (1990). Thermal stress and human performance, *Scandinavian Journal of Work, Environment & Health*, 16, 44-50.

Ensign, W., Hodgdonl, J.A., Prusaczyk, W.K., Ahlers, S., and Shapiro, D. (2000). A Survey of self-reported injuries among Special Boat Operators. San Diego, CA: Naval Health Research Center.

Ernsting, J., Nicholson, A.N., and Rainford, D.J.(1999). *Aviation Medicine*. Oxford: Butterworth Heinemann.

Evans, G.W., and Hygge, S. (2003). Noise and performance in children and adults. In: Luxon, L., and Prasher, D., (Eds.), *Noise and Its Effects*. London: Whurr Publishers.

Evans, G.W., and Lepore, S.J. (1993). Nonauditory effects of noise on children: A critical review. *Children's Environments*, 10, 31-51.

Faerevik, H., Markussen, D., Oglaend, G.E., and Reinertsen, R.E. (2001). The thermoneutral zone when wearing aircrew protective clothing, *Journal of Thermal Biology*, 26, 419-425.

Fan-Paul, N.I., Li, J., Miller, J.S., and Florakis, G.J. (2002). Night vision disturbances after corneal refractive surgery, *Survey of Ophthalmology*, 47(6), 533-546.

Federal Aviation Administration. (2008). Guide for Aviation medical Examiners. http://www.faa.gov/ about/office_org/headquarters_offices/avs/offices/aam/ame/guide/media/guide06.pdf

Fields, J.M. (1994). A review of an updated synthesis of noise/annoyance relationships. NASA CR-194950. NASA, Hampton, VA.

Finnoff, J.T. (2008). Environmental effects on brain function, *Current Sports Medicine Reports*, 7, 28-32.

Flattau, P.E. (Ed.) (1991). Considerations in contact lens use under adverse conditions: Proceedings symposium. Washington, DC: National Academy Press.

Forbes, A.J., and Berkahn, C. (1999). Does zaleplon help you sleep and wake refreshed? *Drugs and Therapy Perspectives*, 14(4), 1-4.

Friedberg, W., Copeland, K., Duke, F.E., and O'Brien III, E.B. (2000). Radiation exposure during air travel: Guidance provided by the FAA for air carrier crew, *Health Physics*, 79, 591-595.

Froom, P, Hanegbi, R., Ribak, J., and Gross, M. (1987). Low back pain in an AH-1 Cobra helicopter, *Aviation Space Environmental Medicine*, 58, 315-8.

Fry, G.A., and Alpern, M. (1953). The effect of a peripheral glare source upon the apparent brightness of an object, *Journal of the Optical Society of America*, 43(3), 189-195.

Fry, J., Scharf, M., Mangano, R., and Fujimori, M. (2000). Zaleplon improves sleep without producing rebound effects in outpatients with insomnia, *International Clinical Psychopharmacoly,15*, 141-152.

Furness, T.A. (1981). The effects of whole-body vibration on the perception of the helmet-mounted display, Unpublished Doctoral Thesis, University of South Hampton.

Gabriel, G. (2006). Hans Selye: The discovery of stress. Retrieved on 15 October 2008 from: http://www.brainconnection.com/topics/?main=fa/selye

Galinsky, T.L., Swanson, N.G., Sauter, S.L., Hurrell, J.J., and Schleifer, L.M. (2000). A field study of supplementary rest breaks for data-entry operators, *Ergonomics, 43(5)*, 622-638.

Georgia State University. (2007). Systems integration (Course syllabus). Retrieved 24 February 2009 from: http://www2.cis.gsu.edu/cis/program/syllabus/graduate/CIS8020

Ghaith, A.A., Daniel, J., Stulting, R.D., Thompson, K.P., and Lynn, M. (1998). Contrast sensitivity and glare disability after radial keratotomy and photorefractive keratectomy, *Archives of Ophthalmology*, 116(1), 12-18

Giesbrecht, G.G., Arnett, J.L., Vela, E., and Bristow, G.K. (1993). Effect of task complexity on mental performance during immersion hypothermia, *Aviation, Space, and Environmental Medicine*, 64, 206-211.

Gilbert, S.S., Van Den Heuvel, C.J., and Dawson, D. (1999). Daytime melatonin and temazepam in young adult humans: Equivalent effects on sleep latency and body temperature, *Journal of Physiology*, 51, 903-14.

Gillberg, M. (1984). The effects of two alternative timings of a one-hour nap on early morning performance, *Biological Psychology,19*, 45-54.

Glick, D.D., and Moser, C.E. (1974). Afterimages associated with using the AN/PVS-5, night vision goggle. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL LR 75-1-7-1.

Goethe, J.W. (1810). Zur Farbenlehre I, Tübingen, Ch.8. Deutsch: Taschenb Verlag, 1963.

Gonzalez-Alonso, J., Teller, C., Andersen, S.L., Jensen, F.B., Hyldig, T., and Nielsen, B. (1999). Influence of body temperature on the development of fatigue during prolonged exercise in the heat, *Journal of Applied Physiology*, 86, 1032-1039.

Goode, J.H. (2003). Are pilots at risk of accidents due to fatigue? *Journal of Safety Research, 34*, 309-313.

Gordon, E., and Houtsma, A.J.M. (2008). Performance of the Communication Enhancement and Protective System (CEPS) under sustained high and low temperature. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2008-10.

Gower, D.W., and Casali, J.G. (1994). Speech intelligibility and protective effectiveness of selected active noise reduction and conventional communications headsets, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 36(2), 350-367.

Graham, C.H. (1966a). Some basic terms and methods (Chapter 3). In: Graham, C.H. (Ed.), *Vision and Visual Perception*. New York: John Wiley and Sons, Inc.

Graham, C.H. (1966b). Some fundamental data (Chapter 4). In: Graham, C.H. (Ed.), *Vision and Visual Perception*. New York: John Wiley and Sons, Inc.

Green, M. (2004). *Seeing Pedestrians at Night*. Retrieved from: www.visualexpert.com/Resources/ pedestrians.html

Green, M., and Senders, J. (2004). *Human Error in Road Accidents*. Retrieved from: www.visualexpert.com/ Resources/ roadaccidents.html

Grether, W.F. (1973). Human performance at elevated environmental temperatures, *Aerospace Medicine,* 44, 747-755.

Griefahn, B. (1992). Noise control during the night: Proposals for continuous and intermittent noise, *Acoustics Australia,* 20, 43-47.

Griefahn, B., Deppe, C., Mehnert, P., Moog, R., Mohler, U., and Schumer-Kohrs, A. (1998). Intermittent versus continuous noises. Their influence on sleep latency and motility during the night, *Proceedings of Internoise '98,* Christchurch. (on CD.) Australia: Causal Productions.

Griffin, M.J. (1990). *Handbook of Human Vibration*. London: Academic Press.

Griffiths, R.R., and Mumford, G.K. (1995). *Caffeine - A drug of abuse?* In: Bloom, F.E., Kupfer, D.J., (Eds.). Psychopharmacology: The Fourth Generation of Progress. New York: Raven Press.

Guardiola-Lemaitre, B. (1997). Toxicology of melatonin, *Journal of Biological Rhythms, 12,* 697-706.

Haas, M.W., Nelson, W.T., Repperger, D., Bolia, R., and Zacharias, G. (2001). Applying adaptive control and display characteristics to future Air Force crew stations, *International Journal of Aviation Psychology,* 11(2), 223-235.

Hackett, P.H., Rennie, D., and Levine, H.D. (1976). The incidence, importance, and prophylaxis of acute mountain sickness, *Lancet*, 2(7996), 1149-55.

Haddock, C.K., Klesges, R.C., Talcott, G.W., Lando, H., and Stein, R.J. (1998). Smoking prevalence and risk factors for smoking in a population of United States Air Force basic trainees, *Tobacco Control,* 7, 232- 235.

Haegerstrom-Portnoy, G., Schneck, M.E., and Brabyn, J.A. (1999). Seeing into old age: Vision function beyond acuity, *Optometry and Vision Science*, 76(3), 141-158.

Haimov, I., and Arendt, J. (1999). The prevention and treatment of jet lag, *Sleep Medicine Reviews, 3,* 229-40.

Haines, M.M., Stansfeld, S.A., Job, R.F.S., and Berglund, B. (1998). Chronic aircraft noise exposure and child cognitive performance and stress. In: Carter, N.L., and Job, R.F.S. (Eds.), *Noise Effects, Proceedings of the 7th International Congress on Noise as a Public Health Problem*. Sydney, Australia: Noise Effects '98.

Hamernik, R.P., and Hsueh, K.D. (1991). Impulse noise: Some definitions, physical acoustics and other considerations, *Journal of the Acoustical Society of America*, 90, 189-196.

Hamernik, R.P., Ahroon, W.A., Davis, R.I., and Axelsson, A. (1989). Noise and vibration interactions: Effects on hearing, *Journal of the Acoustical Society of America*, 86(6), 2129-2137.

Hamilton, A. (1918). A Study of Spastic Anemia in the Hands of Stonecutters: The Effect of the Air Hammer on the Hands of Stonecutters. U.S. Dept of Labor Report 236, No. 19.

Hammond, M.D., Madigan, Jr., W.P., and Bower, K.S. (2005). Refractive surgery in the United States Army, 2000–2003, *Ophthalmology*, 112, 184–190.

Hancock, P.A. (1986). Sustained attention under thermal stress, *Psychological Bulletin*, 99, 263-281.

Harkrider, A., and Champlin, C. (2000). Acute effects of nicotine on physiological responses of the auditory systems of nonsmokers, *Acoustical Society of America*, 107(5), 2900.

Harkrider, A.W., Champlin, C.A., and McFadden, D. (2001). Acute effect of nicotine on non-smokers: I. OAEs and ABRs, *Hearing Research*, 160(1-2), 73-88.

Harma, M. (1995). Sleepiness and shift work: Individual differences, *Journal of Sleep Research*, 4(Suppl. 2), 57-61.

Kristin, L., Harrer, K.L., Yniguez, D., Majar, M., Ellenbecker, D., Estrada, N., and Geiger, M. (2005). Whole body vibration for MS-60S pilots, *Proceedings of the Forty Third Annual SAFE Association Symposium*, 303-314.

Harris, C.S., and Shoenberger, R.W. (1980). Combined effects of broadband noise and complex waveform vibration on cognitive performance, *Aviation, Space, and Environmental Medicine*, 51(100), 1-5.

Hatfield, J., and Job, R.F.S. (1998). Evidence of optimism bias regarding the health effects of exposure to noise. In: Carter, N.L., and Job, R.F.S. (Eds.), *Noise Effects, Proceedings of the 7th International Congress on Noise as a Public Health Problem*. Sydney, Australia: Noise Effects '98, p. 251-254.

Hedge, A. (2008). Human vibration. Retrieved 16 August 2008 from: http://ergo.human.cornell.edu/studentdownloads/DEA325pdfs/Human%20Vibration.pdf. Cornell University.

Hedge, A. (2008). Glare and visual performance. Retrieved 16 July 2008 from Cornell University website: http://ergo.human.cornell.edu/studentdownloads/DEA350pdfs/glare.pdf

Helmholtz H. (1852). Ueber Herrn Brewsters Analyse des Sonnenlichtes, *Ann Physik und Chemie,* 86, 501- 523.

Hemenger, R.P. (1990.) Light scatter in cataractous lenses, *Ophthalmological Physiological Optics*, 10, 394-397.

Herxheimer, A., and Petrie, K.J. (2002). Melatonin for the prevention and treatment of jet lag, *Cochrane Database System Reviews, 2*, CD001520.

Hoffman, R.G. (2001). Human psychological performance in cold environments. In: Pandolf, K.B. (Ed.), *Medical Aspects of Harsh Environments*. Washington, DC: Borden Institute.

Holladay, L.L. (1926). The fundamentals of glare and visibility, *Journal of the Optical Society of America,* 12(2), 271-319.

Holladay, L.L. (1927). Action of a light-source in the field of view in lowering visibility, *Journal of the Optical Society of America*, 14(1), 1-15.

Horne, J.A. (1988). Sleep loss and 'divergent' thinking ability, *Sleep,* 11, 528-536.

Horne, J.A. (1993). Human sleep, sleep loss and behaviour: Implications for prefrontal cortex and psychiatric disorder, *British Journal of Psychiatry,* 162, 413-419.

Howarth, H.V.C., and Griffin, M.J., (1990). Subjective response to combined noise and vibration: Summation and interaction effects, *Journal of Sound and Vibration*, 143(3), 443-454.

Huelke, D.F., Moore, J.L., Compton, T.W., Rouhana, S.W., and Kileny, P.R. (1999). Hearing loss and automobile airbag deployments, *Accident Analysis and Prevention*, 31(6), 789-792.

Hulshof, C.T.J., and Veldhuizen van Zanten. (1987). *International Archives of Occupational Environment and Health*, 59, 205-220.

Hygge, S. (2003). Noise exposure and cognitive performance – Children and the elderly as possible risk groups. Retrieved on 18 August 2008 from University of Gavle, Sweden, Laboratory of Applied Psychology web site: http://www.euro.who.int/document/NOH/ paperbyhygge2003.pdf

Hygge, S., Evans, G.W., and Bullinger, M. (1998). The Munich airport noise study- effects of chronic aircraft noise on children's cognition and health. In: Carter, N.L., and Job, R.F.S. (Eds.), *Noise Effects, Proceedings of the 7th International Congress on Noise as a Public Health Problem*. Sydney, Australia: Noise Effects '98. p. 268-274.

Iglesias, R., Terres, A., and Chavarria, A. (1980). Disorders of the menstrual cycle in an airline stewardess, *Aviation Space and Environmental Medicine,* 51, 518-20.

Ijspeert, J.K., De Waard, P.W.T., van den Berg, T.J.T.P., and De Jong, P.T.V.M. (1990). The intraocular stray light function in 129 healthy volunteers; dependence on angle, age and pigmentation, *Vision Research*, 30(5), 699-707.

International Organization for Standardization. (1997). Mechanical vibration and shock-Evaluation of human exposure to whole-body vibration, Part 1: General requirements. ISO 2631-1:1997. *International Organization for Standardization*.

International Organization for Standardization. (2001). ISO 5349 (for HAV), Mechanical vibration – Measurement and evaluation of human exposure to hand-transmitted vibration. *International Organization for Standardization*.

Jaaskelainen , I.P., Alho, K., Escera , C., Winkler, I., Sillanaukee, B., and Naatanen, R. (1996). Effects of ethanol and auditory distraction on forced choice reaction time, *Alcohol*, 13(2), 153-156.

Jehue, R., Street, D., and Huizenga, D. (1993). Effect of time zone and game time changes on team performance: National Football League, *Medical Science, Sports, and Exercise,* 25(1), 127-131.

Jenkins, L., Tarnopolsky, A., and Hand, D. (1981). Psychiatric admissions and aircraft noise from London Airport: Four-year, three-hospitals' study, *Psychological Medicine*, 11, 765-782.

Job, R.F.S. (1988). Community response to noise: A review of factors influencing the relationship between noise exposure and reaction, *Journal of the Acoustical Society of America*, 83, 991-1001.

Johnson, R.F., and Kobrick, J.L. (2001). Psychological aspects of military performance in hot environments. In: Pandolf, K.B. (Ed.) *Medical Aspects of Harsh Environments*, Washington, DC: Borden Institute, 135-159.

Josephson, J.E. (1991Ocular occupational health concerns: Considerations for pilots wearing contact lenses. In: Flattau, P.E. (Ed.), Considerations in contact lens use under adverse conditions: Proceedings symposium. Washington, DC: National Academy Press.

Kanski, J.L. (2003). *Clinical Ophthalmology: A Systematic Approach (5th Ed.)*. Philadelphia: Butterworth Heinemann.

Kaufman, P.L., and Alm, A. (Ed.s). (2002). *Adler's Physiology of the Eye: Clinical Application (10th Ed.)*. Philadelphia: C. V. Mosby.

Keating, M.P. (1988). *Geometrical, Physical, and Visual Optics*. Boston: Butterworth-Heinemann.

Kelly, J.E., Mihashi, T., and Howland, H.C. (2004). Compensation of corneal horizontal/vertical astigmatism, lateral coma, and spherical aberration by internal optics of the eye, *Journal of Vision,* 4(4), 262-271.

Kennaway, D.J., and Wright, H. (2002). Melatonin and circadian rhythms, *Current Topics in Medical Chemistry,* 2, 199-209.

Kenagy, D.N., Bird, C.T., Webber, C.M., and Fischer, J.R. (2004). Dextroamphetamine use during B-2 combat missions, *Aviation Space and Environmental Medicine,* 75(5), 381-386.

Kjellberg, A. (1990). Psychological aspects of occupational vibration, *Scandinavian Journal of work, Environment and Health,* 16(Supplement 1), 39-43.

King, J.M., and Morse, S.E. (1992). Interpupillary and vertex distance effects on field-of-view and acuity with ANVIS. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 93-9.

Kirsch, A.D. (1996). Report on the statistical methods employed by the U.S. FAA in its cost benefit analysis of the proposed Flight Crewmember Duty Period Limitations, Flight Time Limitations and Rest Requirements, Docket No. 28081. *Comments of the Air Transport Association of America to FAA notice 95-18, FAA Docket No. 28081*. Appendix D, pp. 1-36.

Klein, B.E.,K., Klein, R., and Linton, K.L.P. (1992). Prevalence of age-related lens opacities in a population, *Ophthalmology*, 99(4), 546-552.

Klein, K.E., Bruner, H., Holtmann, H., Rehme, H., Stolze, J., Steinhoff, W.D., Wegmann, H.M. (1970). Circadian rhythm of pilots' efficiency and effects of multiple time zone travel, *Aerospace Medicine*,41(2), 125-132.

Klein, S.A. (2001). Problems with wavefront aberrations applied to refractive surgery: Developing standards, *Ophthalmic Technologies XI, SPIE Proceedings*, 4245, SPIE: Bellingham, Washington.

Klein, S.A., Hoffmann, S., and Hickenbotham, S.C. (2003). Improved glare (halos and scattered light) measurement for post-LASIK surgery. Vision Processing Laboratory, Vision Science Program UC Berkeley. http://cornea.berkeley.edu/presentations/

Knauth P. (1995). Speed and direction of shift rotation, *Journal of Sleep Research*, 4(Suppl. 2), 41-46.

Kok-van Aalphen, C., van der Linden, J.W., Visser, R., and Bol, A.H. (1985). Protection of the police against tear gas with soft lenses, *Military Medicine*, 8, 451-454.

Korb, D.R., Craig, J., Doughty, M., Guillon, J., Smith, G., and Tomlinson, A. (2002). *The Tear Film: Structure, Function and Clinical Examination*. London: Butterworth-Heinemann.

Koslowsky, M., and Babkoff, H. (1992). Meta-analysis of the relationship between total sleep deprivation and performance, *Chronobiology International*, 9(2), 132-136.

Kotulak, J.C. (1992). In-flight field-of-view with ANVIS. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 93-8.

Kotulak, J.C., and Morse, S.E. (1994a). Factors that determine visual acuity through night vision goggles for emmetrones. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 94-l 6.

Kotulak, J.C., and Morse, S.E. (1994b). Relationship among accommodation, focus and resolution with optical instruments, Journal of the Optical Society of America, II, 71-79.

Krizaj, D. (2000). Mesopic state: Cellular mechanisms involved in pre- and post-synaptic mixing of rod and cone signals, *Microscopy Research and Technique*, 50, 347-359.

Krizaj, D., and Hawlina, M. (2002). Cellular basis for rod-cone interactions in the outer retina, *ZDRAV VESTN*, 71(Supplement II), 137-145.

Kryter, K.D. (1990). Aircraft noise and social factors in psychiatric hospital admission rates: A re-examination of some data, *Psychological Medicine*, 20, 395-411.

Kumari, V., Gray, J.A., Ffytche, D.H., Mitterschiffthaler, M.T., Das, M., Zachariah, E., Vythelingum, G.N., Williams, S.C.R., Simmons, A., and Sharma, T. (2003). Cognitive effects of nicotine in humans: An fMRI study, *NeuroImage*, 19(3), 1002-1013.

Kunz, D., and Bes, F. (2001). Exogenous melatonin in periodic limb movement disorder: An open clinical trial and a hypothesis, *Sleep*, 24, 183-7.

Lagarde, D., and Batejat, D. (1995). Disrupted sleep-wake rhythm and performance: Advantages of modafinil, *Military Psychology*, 7(3), 165-191.

Lam, D.M. (2003). Aviation medicine. In: Kelly, P.W. (Ed.), *Military Preventive Medicine: Mobilization and Deployment*, *Volume I*. Department of the Army. Washington, DC: Office of the Surgeon General.

Lattimore, M.R., and Cornum, R.L.S. (1992). The use of extended wear contact lenses in the aviation environment : An Army wide study. Fort Rucker, Alabama: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 92-35.

Lavie, P. (1986). Ultrashort sleep-waking schedule. III. 'Gates' and 'forbidden zones' for sleep, *Electroencephalography and Clinical Neurophysiology*, 63, 414-425.

Lee, K.A., Vaillant, G.E., Torrey, W.C., and Elder, G.H. (1997). A 50-year prospective study of the psychological sequelae of World War II combat, *Year Book of Psychiatry and Applied Mental Health*, 1997(7), 301-302.

Leger, D., Badet, D., and de la Giclais, B. (1993). The prevalence of jet-lag among 507 traveling businessmen, *Sleep Research*, 22, 409.

Levin, E.D., McClernon, F.J., and Rezvani, A.H. (2006). Nicotinic effects on cognitive function: Behavioral characterization, pharmacological specification, and anatomic localization, *Psychopharmacology (Berlin)*.184(3-4), 523-39.

Levin, E.D., and Rezvani, A.H. (2002). Nicotinic treatment for cognitive dysfunction, *Current Drug Targets (CNS Neurological Disorders*), 1(4), 423-31.

Lewy, A.J., Ahmed, S., Jackson, J.M., and Sack, R.L. (1992). Melatonin shifts human circadian rhythms according to a phase-response curve, *Chronobiology International*, 9, 380-92.

Lewy, A.J., Bauer, V.K., Ahmed, S., Thomas, K.H., Cutler, N.L., Singer, C.M., Moffit, M.T., and Sack, R.L. (1998). The human phase response curve (PRC) to melatonin is about 12 hours out of phase with the PRC to light, *Chronobiology International*, 15, 71-83.

Lewis, A.L. (1993). Basic concepts in environmental lighting. Chapter Five. In: Pitts, D.G., and Kleinstein, R.N (Eds.), *Environmental Vision: Interactions of the Eye, Vision, and the Environment*. Boston: Butterworth-Heinemann. p. 137-150.

Lewis, C.H., and Griffin, M.J. (1976). The effects of vibration on manual control performance, *Ergonomics*, 19(2), 203-216.

Licina, J.R. (2000). Human factors engineering (HFE) issues. In: Rash, C.E. (Ed.), *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft. Bellingham*. WA: SPIE Press.

Lieberman, H.R., Wurtman, R.J., Emde, G.G., Roberts, C., and Coviella, I.L.G. (1987). The effects of low doses of caffeine on human performance and mood, *Psychopharmacology (Berlin)*, 92, 308-312.

Lino, A., Silvy, S., Condoroelli, L., and Rusconi, A.C. (1993). Melatonin and jet lag: Treatment schedule, *Biological Psychiatry*, 34, 587-88.

Ljungberg, J., Neely, G., and Lundstrom, R. (2004). Cognitive performance and subjective experience during combined exposures to whole-body vibration and noise, *International Archives of Occupational Environmental Health,* 77, 217-211.

Lowenstein, O., and Loewenfeld, I.E. (1969). The pupil. In: Davson, H. (Ed.), *The Eye*. New York: Academic Press. p. 255-337.

Mabbott, N., Foster, G., and McPhee, B. (2001). Heavy vehicle seat vibration and driver fatigue. Report to the Department of Transport and Regional Services, Australian Transport Safety Bureau.

Madigan, W., and Bower, K. (2004). Refractive surgery and protective eyewear in the military, *Ophthalmology,* 111(5), 855-856.

Mandell, R.B. (1988). *Contact Lens Practice.* Springfield, IL: Charles Thomas.

Mansfield, N.J. (2005). *Human Response to Vibration*. Boca Raton, FL: CRC Press.

Martin, F.C., and Peters, T.J. (1985). Assessment in vitro and in vivo of muscle degradation in chronic skeletal muscle myopathy of alcoholism, *Clinical Science (Colch.)*, 68, 693-700.

Martinez, C.E., Applegate, R.A., Klyce, S.D., McDonald, M.B., Medina, J.P., and Howland, H.C. (1998). Effect of pupillary dilation on corneal optical aberrations after photorefractive keratectomy, *Archives of Ophthalmology*, 116, 1053-1062.

Mattila, M.J., Aranko, K., Mattila, M.E., and Strömberg, C. (1984). Objective and subjective assessment of hangover during subacute administration of temazepam and nitrazepam to healthy subjects, *European Journal of Clinical Pharmacology,* 26, 375-380.

McKenzie, R.E., and Elliot, L.L. (1965). Effects of secobarbital and d-amphetamine on performance during a simulated air mission, *Aerospace Medicine,* 36, 774-779.

McLean, W.E. (1995). Video method of measuring field-of-view of electro-optical devices versus eye clearance. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 95-30.

McLean, W.E. (1996). ANVIS objective lens depth of field. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 96-17.

McLean, W.E. (1997). An assessment of luminance imbalance with ANVIS at an Army helicopter training airfield. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 97-21.

McLeod, S.D. (2001). Beyond Snellen acuity, *Archives of Ophthalmology*, 119, 1371-1373.

Meir, R. (2002). Managing transmeridian travel: Guidelines for minimizing the negative impact of international travel, *Strength and Conditioning Journal*, 24(4), 28-34.

Melki, S.A., Proano, C.E., and Azar, D.T. (2003). Optical aberrations after LASIK. In: Azar, D.T., and Koch, D.D. (Eds.) *LASIK Fundamentals, Surgical Techniques, and Complications*. New York: Marcel Dekker. p. 387-402.

Menkes, D.B. (2000). Hypnosedatives and anxiolytics. In: Dukes, M.N.G., and Aronson, J.K (Eds.), *Meyler's Side Effects of Drugs.* Amsterdam: Elsevier Science.

Melzer, J.E., and Moffitt, K. (1997). *Head-Mounted Displays: Designing for the User*. McGraw-Hill, New York, New York.

Miller, II, R.E., and Tredici, T.J. (2002). Night Vision Manual for the Flight Surgeon. Brooks AFB, TX: Armstrong Laboratory. AL-SR-1992-0002.

Minors, D.S., and Waterhouse, J.M. (1990). Circadian rhythms in general. In: Scott, A.J. (Ed.) *Shiftwork: Occupational Medicine: State of the Art Reviews*. Philadelphia, PA: Hanley and Belfus, Inc.

Mitler, M.M. (2000). Nonselective and selective benzodiazepine receptor agonists—where are we today? *Sleep*, 23(Suppl. 1), S39-S47.

Mitler, M.M., and Aldrich, M.S. (2000). *Stimulants: Efficacy and adverse effects*. In: Kryger, M.H., Roth, T., and Dement, W.C. (Eds.). *Principles and Practice of Sleep Medicine* (3rd Ed.). Philadelphia: W.B. Saunders Co.

Minors, D.S., and Waterhouse, J.M. (1988). Avoiding jet lag again, *Nature,* 32, 23-24.

Mizoue, T., Miyamoto, T., and Shimizu, T. (2003). Combined effect of smoking and occupational exposure to noise on hearing loss in steel factory workers, *Occupational and Environmental Medicine*, 60, 56-59.

Moffitt, K. (1991). HDU image focus and visual accommodation. San Jose, CA: Kaiser Electronics. Position paper, dated 7 Oct 1991.

Moffitt, K., Rogers, S. P., and Cicinelli, J. (1988). Chromatic aftereffects associated with a night vision goggle simulation, *Aviation, Space, and Environmental Medicine*, 59(2), 125-128.

Monk, T.H. (1990). The relationship of Chronobiology to sleep schedules and performance demands, *Work and Stress,* 4(3), 227-236.

Moore-Ede, M. (1993). *Aviation Safety and Pilot Error, in Twenty-Four Hour Society*. Reading, MA: Addison-Wesley Publishing Co.

Moore-Ede, M.C., Kaas, D.A., and Herd, J.A. (1977). Transient circadian internal desynchronization after light-dark phase shift in monkeys, *American Journal of Physiology,* 232, R31-R37.

Al W., Moran, A.W, and Butler, B.P. (1993). Whole-body vibration assessment of the M916A1 truck tractor. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 93-34.

Morata, T.C. (1999). Is it still necessary to conduct research on noise-induced hearing loss? *Noise and Health*, 2, 3-6.

Morell, S., Taylor, R., Carter, N.L., Job, R.F.S., and Peploe, P. (1998). Cross-sectional relationship between blood pressure of school children and aircraft noise. In: Carter, N.L., and Job, R.F.S. (Eds.), *Noise Effects. Proceedings of the 7th International on Noise as a Public Health Problem*. Sydney, Australia: Noise Effects Inc. p. 275-279.

Morris, M.S., and Borja, L.P. (1998). Noise levels associated with airbag deployment may result in occupants experiencing irreversible hearing loss, *Trauma*, 44(1), 238-239.

Mozo, B.T., and Murphy, B.A. (1997). The assessment of sound attenuation and speech intelligibility of selected active noise reduction devices and the communications earplug when used with the HGU- 56/P aviator helmet. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 97-08.

Mozo, B.T., and Murphy, B.A. (1998). Novel communications and hearing protection for helmet systems: Communications Earplug. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 98-23.

Muller, F.O., Van Dyk, M.V., Hundt, H.K.L., Joubert, A.L., Luus, H.G., Groenewoud, G., and Dunbar, G.C. (1987). Pharmacokinetics of temazepam after day-time and night-time oral administration, *European Journal of Clinical Pharmacology,* 33, 211-214.

Naitoh, P. (1975). Sleep deprivation in humans. In: Venables, P.H., and Christie, M.J. (Eds.), *Research in Psychophysiology*. London: John Wiley.

Nakashima, A.M. (2004). The effect of vibration on human performance and health: A review of recent literature. Toronto: Defence Research and Development Canada. Technical Report DRDC TR 2004-089.

Nakashima, A.M., and Cheung, B. (2006). The effects of vibration frequencies on physical, perceptual and cognitive performance. Toronto: Defence Research and Development Canada, Technical Report 2006-218.

Nakanishi, N., Okamoto, M., Nakamura, K., Suzuki, K., and Tatara, K. (2000). Cigarette smoking and risk for hearing impairment: A longitudinal study in Japanese male office workers, *Journal of Occupational and Environmental Medicine*, 42(11), 045-1049.

National Cancer Institute. (2008). Overview of prevention. Retrieved from: http://www.cancer.gov/cancertopics/pdq/prevention/control-of-tobacco-use

National Cancer Institute. (2008). Purpose of PDQ summary. Retrieved from: http://www.cancer.gov/cancertopics/pdq/prevention/lung/HealthProfessional

National Institute for Occupational Safety and Health. (1989). Criteria for a recommended standard for occupational exposure to hand-arm vibration, *National Institute for Occupational Safety and Health*. NIOSH #89-106

National Transportation Safety Board. (1994). *A review of flightcrew-involved, major accidents of U.S. air carriers, 1978 through 1990* (NTSB Safety Study No. SS-94-01). Washington, DC.

Naval Aerospace Medical Institute. (2007a). Waiver Guide: Medications. http://navmedmpte.med.navy.mil/nomi/nami/waiverguidetopics.cfm

Naval Aerospace Medical Institute. (2007b). U.S. Navy Aeromedical Reference and Waiver Guide. http://navmedmpte.med.navy.mil/nomi/nami/WaiverGuideTopics/pdfs/Waiver%20Guide%20-%20Physical%20Standards.pdf

Nawrot, M., Nordenstrom, B., and Olson, A. (2004). Disruption of eye movements by ethanol intoxication affects perception of depth from motion parallax, *Psychological Science*, 15(12), 858-865.

Neri, D.F., Oyung, R.L., Colletti, L.M., Mallis, M.M., Tam, P.Y., and Dinges, D.F. (2002). Controlled activity as a fatigue countermeasure on the flight deck, *Aviation, Space and Environmental Medicine*, 73(7), 654-664.

Newhouse, P.A., Belenky, G., Thomas, M., Thorne, D., Sing, H.C., and Fertig, J. (1989). The effects of d-amphetamine on arousal, cognition, and mood after prolonged total sleep deprivation, *Neuropsychopharmacology,* 2(2), 153-164.

Nicholson, A.N. (1983). Antihistamines and sedation, *Lancet,* 2, 211-12.

Nicholson, A.N. (1990). Hypnotics and occupational medicine, *Journal of Occupational Medicine,*32(4), 335-341.

Nicholson, A.N., Pascoe, P.A., Spencer, M.B., Stone, B.M., Roehrs, T.A., and Roth, T. (1986). Sleep after transmeridian flights, *Lancet,* 2, 1205-1208.

Nicholson, A.N., Roth, T., and Stone, B.M. (1985). Hypnotics and aircrew, *Aviation Space and Environmental Medicine,* 56, 299-303.

Nicholson, A.N., and Stone, B.M. (1982). *Sleep and Wakefulness Handbook for Flight Medical Officers*. Neuilly sur Seine, France: Advisory Group for Aerospace Research and Development.

Nicholson, A.N., Stone, B.M., and Pascoe, P.A. (1980). Efficacy of some benzodiazepines for day-time sleep, *British Journal of Clinical Pharmacology,* 10, 459-463.

Nilsson, S**.**E., and Anderson, L. (1982). The use of contact lenses in environments with organic solvents, acids or alkali, *Acta Ophthalmologica*, 60(4), 599-608.

Nunneley, S.A., Reader, D.C., and Maldonado, R.J. (1982). Head-temperature effects on physiology, comfort, and performance during hyperthermia, *Aviation, Space, and Environmental Medicine*, 53, 623-628.

O'Brien, C.P., and Lyons, F. (2000). Alcohol and the athlete, *Sports Medicine*, 29(5), 295-300.

Ogilvie, R.D., Simons, I.A., Kuderian, R.H., MacDonald, T., and Rustenburg, J. (1991). Behavioral, event-related potential and EEG/FFT changes at sleep onset, *Psychophysiology*, 28, 54-64.

Ogilvie, R.D., Wilkinson, R.T., and Allison, S. (1989). The detection of sleep onset: Behavioral, physiological and subjective convergence, *Sleep*, 12, 458-474.

Ohrstrom, E. (1997). Effects of exposure to railway noise – a comparison between areas with and without vibration, *Journal of Sound and Vibration*, 205(4), 550-560.

Ohrstrom, E., Bjorkman, M., and Rylander, R. (1990). Effects of noise during sleep with reference to noise sensitivity and habituation: Studies in laboratory and field, *Environment International*, 16, 477-482.

Ohrstrom, E., and Skanberg, A.B. (1996) A field survey on effects of exposure to noise and vibration from railway traffic, Part I: Annoyance and disturbance effects, *Journal of Sound and Vibration*, 193(1), 39-47.

Ohshund, R.K., Dalton, C.I., Reams, G.C., Rose, J.W., and Oswald, R.E. (Eds.) (1991). *United States Naval Flight Surgeon's Manual*. Pensacola, FL: Naval Aerospace Medical Institute.

Okada, A., Miyake, H., Yamamura, K., Minami, M., 1972. Temporary hearing loss induced by noise and vibration, *Journal Acoustical Society of America*, 51(4), 1240-1248.

O'Keefe, J.H., Bybee , K.A., and Lavie, C.J. (2007). Alcohol and cardiovascular health: The razor-sharp double-edged sword, *Journal of the American College of Cardiology*, 50(11), 1009-1014.

O'Neal, M.R. (1991). Effect of aircraft cabin altitude and humidity on oxygen tension under soft and hard gas-permeable contact lenses. In: Flattau, P.E. (Ed.), *Considerations in Contact Lens Use under Adverse Conditions: Proceedings of a Symposium.* Washington, DC: National Academy Press. Retrieved 1 Oct 2008 from: http://www.nap.edu/catalog/1773.html

Orford, R.R., and Silberman, W.S. (2008). Pilot health and aeromedical certification. In: Davis, J.R., Johnson, R., Fogarty, J.A., and Stepanek, J. (Eds.), *Fundamentals of Aerospace Medicine,* 4th Edition Philadelphia: Lippincott Williams and Wilkins.

Oshika, T., Klyce, S.D., Applegate, R.A., Howland, H.C., and El Danasoury, M.A. (1999). Comparison of corneal wavefront aberrations after photorefractive keratectomy and laser in situ keratomileusis, *American Journal of Ophthalmology,* 127, 1-7.

O'Toole, K.J. (2004). *U.S. Air force aircrew fatigue survey*. Abstracts of the Aerospace Medical Association 75th Annual Scientific Meeting, No. 293, pp. B64, Ankorage, AK. Arlington, VA: Aerospace Medical Association.

Oxenkrug, G.F., and Requintina, P.J. (2003). Melatonin and jet lag syndrome: Experimental model and clinical implications, *CNS Spectrum,* 8, 139-48.

Paakkonen, R., and Lehtomaki, K. (2005). Protection efficiency of hearing protectors against military noise from handheld weapons and vehicles, *Noise and Health*, 7(26), 11-20.

Paddan, G.S., and Griffin, M.J. (1988). The transmission of translational vibration to the head, *Journal of Biomechanics*, 21, 195-197.

Palminteri, R., and Narbonne, G. (1988). Safety profile of zolpidem. In: Sauvanet, J.P., Langer, S.Z., and Morselli, P.L. (Eds.). *Imidazopyridines in Sleep Disorders*. New York: Raven Press.

Pandi-Perumal, S.R., Seils, L.K., Kayumov, L., Ralph, M.R., Lowe, A., Moller, H., and Swaab, D.F. (2002). Senescence, sleep, and circadian rhythms, *Ageing Research Reviews,* 1, 559-604.

Papathanasiou, G., Georgakopoulos, D., Georgoudis, G., Spyropoulos, P., Perrea, D., and Evangelou A. (2007). Effects of chronic smoking on exercise tolerance and on heart rate-systolic blood pressure product in young healthy adults, *European Journal of Cardiovascular Prevention and Rehabilitation*, 14(5), 646-52

Parry, B.L. (2002). Jetlag: Minimizing its effects with critically timed bright light and melatonin administration, *Journal of Molecular Microbiology Biotechnology,* 4, 463-466.

Parsons, K.C. (2003). *Human Thermal Environments*, London: Taylor and Francis Incorporated.

Paul, M.A., Gray, G., MacLellan, M., and Pigeau, R.A. (2004). Sleep-inducing pharmaceuticals: A comparison of melatonin, zaleplon, zopiclone, and temazepam, *Aviation Space and Environmental Medicine,* 75(6), 512-519.

Pearson, P., and Timney, B. (1998). Effects of moderate blood alcohol concentrations on spatial and temporal contrast sensitivity, *Journal of Studies on Alcohol*, 59(2), 163-73.

Pearsons, K.S., Barber, D.S. Tabachnick, B.G., and Fidell, S. (1995). Predicting noise-induced sleep disturbance, *Journal of the Acoustical Society of America,* 97, 331-338.

Pelham, T.W., White, H., Holt, L.E., and Lee, S.W. (2005). The etiology of low back pain in military helicopter aviators: Prevention and treatment, *Work: A Journal of Prevention, Assessment and Rehabilitation,* 24 (2), 101-110.

Pelmear, P., and Wasserman, D.E. (1998). *Hand-Arm Vibration: A Comprehensive Guide for Occupational Health Professionals.* (2^{nd} Edition). Beverly Farms, MA: OEM Medical Publishers.

Penetar, D.M., McCann, U., Thorne, D., Kamimori, G., Galinski, C., Sing, H., Thomas, M., and Belenky, G. (1993). Caffeine reversal of sleep deprivation effects on alertness and mood, *Psychopharmacology (Berlin),* 112, 359-365.

Penn, P.E., and Bootzin, R.R. (1990). Behavioral techniques for enhancing alertness and performance in shift work, *Work and Stress,* 4(3), 213-226

Perry, I.C. (Ed.) (1974). *Helicopter aircrew fatigue* (AGARD Advisory Report No. 69). Neuilly sur Seine, France: Advisory Group for Aerospace Research and Development.

Petrie, K., Conaglen, J.V., Thompson, L., and Chamberlain, K. (1998). Effect of melatonin on jet lag after long haul flights, *British Medical Journal*, 298, 705-7.

Petrie, K.J., Power, D., and Broadbent, E. (2004). Fatigue self-management strategies and reported fatigue in international pilots, *Ergonomics*, 47, 461-68.

Pevet, P., Bothorel, B., Slotten, H., and Saboureau, M. (2002). The chronobiotic properties of melatonin, *Cell Tissue Research,* 309, 183-91.

Physician's Desk Reference. (2003). *Dexedrine (brand of dextroamphetamine sulfate).* In: Physician's Desk Reference. New Jersey: Medical Economics Co., Inc.; p. 1500-1501.

Pigeau, R.A., Naitoh, P., Buguet, A., McCann, C., Baranski, J., Taylor, M., et al. (1995). Modafinil, d-amphetamine and placebo during 64 hours of sustained mental work. Effects on mood, fatigue, cognitive performance and body temperature, *Journal of Sleep Research,* 4, 212-228.

Pilcher, J.J., Nadler, E., and Busch, C. (2002). Effects of hot and cold temperature exposure on performance: A meta-analytic review, *Ergonomics*, 45, 682-698.

Pitts, D.G. (1993). Electromagnetic spectrum. Chapter Four. In: Pitts, D.G., and Kleinstein, R.N. (Eds.), *Environmental Vision: Interactions of the Eye, Vision, and the Environment.* Boston: Butterworth-Heinemann.

Polse, K. (Chair) (1990). *Contact lens use under adverse conditions: Applications in military aviation*. Washington, DC: National Academy Press.

Popke, E.J., Allen, S.R., and Paule, M.G. (2000). Effects of acute ethanol on indices of cognitive-behavioral performance in rats, *Alcohol*, 20(2), 187-192.

Porcu, S., Bellatreccia, A., Ferrara, M., and Casagrande, M. (1997). Performance, ability to stay awake, and tendency to fall asleep during the night after a diurnal sleep with temazepam or placebo, *Sleep,* 20(7), 535-541.

Pozos, R.S., and Danzl, D.F. (2001). Human physiological responses to cold stress and hypothermia. In: Pandolf, K.B. (Ed.), *Medical Aspects of Harsh Environments*. Washington, DC: Borden Institute.

Pulling, N.H., Wolf, E., Sturgis, S.P., Vaillancourt, D. R., and Dolliver, J. J. (1980). Headlight glare resistance and driver age, *Human Factors*, 22(1), 103-112.

Purkinje J. (1823). Beobachtungen und Versuche zur Physiologie der Sinnesorgane I, Ch.11 and 12. Berlin: Reiner.

Pyykko, I., Starck, J., Färkkila, M., Hoikkala, M., Korhonen, O., and Nurminen, M. (1981). Hand-arm vibration in the etiology of hearing loss in lumberjacks, *British Journal of Industrial Medicine*, 38(3), 281-9.

Quintyn, J.C., Massy, J., Quillard, M., and Brasseur, G, (1999), Effects of low alcohol consumption on visual evoked potential, visual field and contrast sensitivity, *Acta Ophthalmologica Scandinavica*, 77(1), 23-6.

Rash, C.E. (2006). Shutting out the noise. *Flight Safety Foundation*. November.

Rash, C.E., Kalich, M.E, van de Pol, C., and Reynolds, B.S. (2003). The Issue of Visual Correction Compatibility with Helmet-Mounted Displays. Fort Rucker, AL: U.S. army Aeromedical Research Laboratory. USAARL Report No. 2003-04.

Rash, C.E., and Manning, S. (2001). Laser-light displays, laser pointers disrupt crewmember vision, *Human Factors and Aviation Medicine*, 48(6) 1-8.

Rash, C.E., and McLean, W.E. (2000). Optical performance. In: Rash, C.E. (Ed.), *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft*. Bellingham, WA: SPIE Press.

Rash, C.E., Verona, R.W., and Crowley, J.S. 1990. Human factors and safety considerations of night vision systems flight using thermal imaging systems, *Helmet-Mounted Displays II, Proceedings of SPIE,* 1290, 142-164.

Recht, L.D., Lew, R.A., and Schwartz, W.J. (1995). Baseball teams beaten by jet lag, *Nature*, 377, 583.

Records, R.E. (1979). *Physiology of the Human Eye and Visual System.* New York: Harper and Row.

Refojo, M.F. (1991). Tear evaporation considerations and contact lens wear. In: Flattau, P.E. (Ed.) *Considerations in Contact Lens Use Under Adverse Conditions: Proceedings of a Symposium.* Washington, DC: National Academy Press.

Reinhart, R.O. (1992) *Basic Flight Physiology*. Blue Ridge Summit, PA: TAB Books

Reiter, R.J. (2003). Melatonin: Clinical relevance, *Best Practice and Research in Clinical Endocrinology Metabolism,* 17, 273-85.

Reppert, S.M., Weaver, D.R., Rivkees, S.A., and Stopa, E.G. (1988). Putative melatonin receptors in a human biological clock, *Science*, 242, 78-81.

Rezvani, A.H., and Levin, E.D. (2001). Cognitive effects of nicotine, *Biological Psychiatry*, 49(3), 258-67.

Roehrs, T., Burduvali, E., Bonahoom, A., Drake, C., and Roth, T. (2003). Ethanol and sleep loss: A ‘dose’ comparison of impairing effects, *Sleep,* 26(8), 981-985.

Rosa, R.R. (1995). Extended workshifts and excessive fatigue. *Journal of Sleep Research, 4*(Suppl. 2), 51-56.

Rosekind, M.R., Graeber, R.C., Dinges, D.F., Connell, L.J., Rountree, M.S., Spinweber, C.L., and Gillen, K.A. (1994). *Crew factors in flight operations IX: Effects of planned cockpit rest on crew performance and alertness in long-haul operations* (NASA Technical Memorandum no. 108839). Moffett Field, CA: National Aeronautics and Space Administration.

Rosekind, M.R., Gregory, K.B., Miller, D.L., Co, E.I., Lebacqz, J.V., and Brenner, M. (1996). Crew fatigue factors in the Guantanamo Bay aviation accident, *Sleep Research*, 25, 571.

Roth, T., Piccione, P., Salis, P., Kramer, M., and Kaffeman, M. (1979). Effects of temazepam, flurazepam and quinalbarbitone on sleep: Psychomotor and cognitive function, *British Journal of Clinical Pharmacology,* 8, 47S-54S.

Roth, T., and Roehrs, T. (1991). A review of the safety profiles of benzodiazepine hypnotics, *Journal of Clinical Psychiatry,* 52(9) (Suppl), 38-47.

Russell, R.M., Carney, E.A., Feiock, K., Garrett, M., and Karwoski, P. (1980), Acute ethanol administration causes transient impairment of blue-yellow color vision, *Alcoholism Clinical and Experimental Research*, 4(4), 396-9.

Rycroft, N., Rusted, J.M., and Hutton, S.B. (2005). Acute effects of nicotine on visual search tasks in young adult smokers, *Psychopharmacologia*, 181(1), 160-169.

Sacco, R.L., Elkind, M., Boden-Albala, B., Feng Lin, I., Kargman, D.E., Hauser, W.A., Shea, S., and Paik, M.C. (1999). The protective effect of moderate alcohol consumption on ischemic stroke, *Journal of the American Medicine Association*, 281, 53-60.

Sack, R.L., Lewy, A.J., Hughes, R.J., McArthur, A.J., and Blood, M.L. (1996). Melatonin as a chronobiotic drug, *Drug News Perspective* 9, 325-32.

Sagawa, K., and Takeichi, K. (1992). System of mesopic photometry for evaluating lights in terms of comparative brightness relationships, *Journal of the Optical Society of America*, 9(8), 1240-1246.

Salas, E., Driskell, and Hughes, S. (1996). Introduction: The study of stress and human performance. In: Driskell, J.E., and Salas, E. (Eds.), *Stress and Human Performance.* New Jersey: Lawrence Erlbaum.

Samel, A., Wegman, H.M., Vejvoda, M., Drescher, J., Gundel, A., Manzeu, D., and Wenzel, J. (1997). Two-crew operations: Stress and fatigue during long-haul night flights, *Aviation Space and Environmental Medicine,* 68(8), 679-687.

Sanders, M.S., and McCormick, E.J. (1993). *Human Factors in Engineering and Design*. New York: McGraw-Hill.

Sanger, D.J., Perrault, G., Morel, E., Joly, D., and Zivkovic, B. (1987). The behavioral profile of zolpidem, a novel hypnotic drug of imidazopyridine structure, *Physiology and Behavior*, 41, 235-240.

Sawka, M.N., and Pandolf, K.B. (2001). Physical exercise in hot climates: Physiology, performance, and biomedical issues. In: Pandolf, K.B. (Ed.), *Medical Aspects of Harsh Environments*. Washington, DC: Borden Institute.

Schallhorn, S.C. (1994). Photorefractive surgery in the Navy, *Navy Medicine*, 85, 27-32.

Schallhorn, S.C. (1999). PRK in Naval Aviation. Presented at Naval Aeromedical Medical Institute, Pensacola, FL. November 12, 1999.

Schallhorn, S.C., Blanton, B.L., Kaupp, S.E., Sutphin, J., Gordon, M., Goforth Jr, H., and Butler Jr, F.K. (1996). Preliminary results of photorefractive keratectomy in active-duty United States Navy personnel, *Ophthalmology*. 103, 5-22.

Schallhorn, S.C., Reid, J.L., Kaupp, S.E., Blanton, C.L., Zoback, L., Goforth Jr., H., Flowers, C.W., and McDonnell, P.J. (1998). Topographic detection of photorefractive keratectomy, *Ophthalmology*, 105, 507-16.

Schieber, F. (1994a.). Age and glare recovery time for low-contrast stimuli, *Proceedings of the Human Factors and Ergonomics Society*. Santa Monica, California: Human Factors and Ergonomics Society. p. 496-499.

Schieber, F. (1994b). *Recent Developments in Vision, Aging and Driving: 1988-1994*. Report No. UMTRI-94-26. Ann Arbor, Michigan: University of Michigan, Transportation Research Institute. Retrieved from: www.usd.edu/ ~schieber/pdf/umtri.pdf

Schieber, F. (1995). Effects of visual aging upon driving performance. In: Adrian, W. (Ed.), *Lighting for aging vision and health.* New York: Lighting Research Institute.

Schneck, M.E., Adams, A.J., Huie, K. and Lee, E. (1993). A filter for simulating color and spatial vision of the elderly. In: Drum, B. (Ed.), *Colour Vision Deficiencies XI*. Kluwer Academic Publishers: Netherlands.

Schmitz, F.H. (1995). Reduction of blade-vortex interaction (BVI) noise through x-force control. Moffett Field, CA: Ames Research Center. NASA Technical Memorandum 110371.

Schouten, J.F., and Ornstein, L.S. (1939). Measurements on direct and indirect adaptation by means of a binocular method, *Journal of the Optical Society of America*, 29, 168-182.

Schumer, D.J., Bains, H.S., and Brown, K.L. (2000). Dark-adapted pupil sizes in prospective evaluation of laser in situ keratomileusis patients, *Journal of Refractive Surgery*, 16, S239-S41.

Schwartz, S.H. (1994). *Visual Perception: A Clinical Orientation*. Norwalk, Connecticut: Appleton and Lange.

Seidel, H., and Heide, R. (1986). Long-term effects of whole-body vibration: A critical survey of the literature, *International Archives of Occupational Environment and Health*, 58, 1-26.

Selye, H. (1946). The general adaptation syndrome and the diseases of adaptation, *Journal of Clinical Endocrinology*, 6, 117-230.

Selye, H. (1952). *The Story of the Adaptation Syndrome*. Montreal, Quebec, Canada: Acta Inc.

Senechal, P.K. (1988). Flight surgeon support of combat operations at RAF Upper Heyford, *Aviation Space and Environmental Medicine*, 59, 776-777.

Serafin, W.E. (1996). *Drugs used in the treatment of asthma*. In: Hardman, J.G., Limbird, L.E., Molinoff, P.B., Ruddon, R.W., and Gilman, A.G. (Eds.), *Goodman and Gilman's The Pharmacological Basis of Therapeutics*. (9th Ed.). New York: McGraw Hill.

Shanahan, D.F., Shannon, S.G., and Bruckart, J.E. (1993). Projected effectiveness of airbag supplemental restraint systems in U.S. Army helicopters. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 93-31.

Shappell, S.A., Neri, D.F., and DeJohn, C.A. (1992). Simulated sustained flight operations and performance, Part 2: Effects of dextro-methamphetamine, *Military Psychology*, 4(4), 267-287.

Sheehy, J.B., and Wilkinson, M. 1989. Depth perception after prolonged usage of night vision goggles, *Aviation, Space. and Environmental Medicine*, 64(6), 573-579.

Shinsuke, S., Kamitani, Y., and Nishida, S. (2001). Afterimage of perceptually filled-in surface, *Science*, 293, 1677-1680.

Short, R.V., and Armstrong, S.M. (1984). Method for minimizing disturbances in circadian rhythms and bodily performance and function, *Australian Patent Application*. PG 4737, April 27, 1984.

Simmons, S.E., Saxby, B.K., McGlone, F.P., and Jones, D.A. (2008) The effect of passive heating and head cooling on perception, cardiovascular function and cognitive performance in heat, *European Journal of Applied Physiology*, 104, 271-280.

Simpson, H.W. (1980). Chronobiotics: Selected agents of potential value in jet-lag and other desynchronisms. In: Scheving, L.E., and Halberg, F. (Eds.), *Chronobiology: Principles and Applications to Shifts in Schedules*. The Netherlands: Nijthoff and Noordhoff.

Skene, D.J., Aldhous, M., and Arendt, J. (1988). Melatonin, jet-lag and the sleep-wake cycle. In: Horne, J. (Ed.), *Proceedings of the Ninth European Congress on Sleep Research*, 89, 39. Stuttgart, Germany: Gustav Fischer Verlag.

Smith, A.M. (2005). Hypoxia symptoms reported during helicopter operations below 10,000 feet: A retrospective survey, *Aviation Space and Environmental Medicine*, 76, 794-798.

Smith, A.M. (2006). The impact of short-duration light and moderate physical activity at 7,000 ft and 9,000 ft on the development of hypoxia, *Aviation Space and Environmental Medicine,* 77, 274.

Smith, A.P. (1989). A review of the effects of noise on human performance, *Scandinavian Journal of Psychology*, 30, 185-206.

Smith, A.P. (1992). Noise and performance. In: Smith, A.P., and Jones, D.M. (Eds.), *Handbook of Human Performance: The Physical Environment.* London: Academic Press.

Smith, D.G., Baranski, J.V., Thompson, M.M., and Abel, S.M. (2003). The effects of background noise on cognitive performance during a 70 hour simulation of conditions aboard the International Space Station. *Noise Health*, 6, 3-16.

Smith, G. (2002). Disability glare and its clinical significance, *Optometry Today-Optics Today*, 34-37. Retrieved 15 March 2009 from: http://www.optometry.co.uk/articles/20020419/smith20020418.pdf

Smith, S.D. (2002). Characterizing the effects of airborne vibration on human body vibration response, *Aviation, Space and Environmental Medicine*, 73(1), 36-45.

Smith, S.D. (2004). Cockpit seat and pilot helmet Vibration during flight operations on aircraft carriers, *Aviation, Space, and Environmental Medicine,* 75(1), 247-254.

Soames-Job, R.F., and Hatfield, J. (2000). Effective communication of health messages regarding noise-induced health effects, *Noise Health*, 2, 33-8.

Sobotta, J. (1990). *Atlas of Human Anatomy*. Baltimore: Urban and Schwarzenberg. Jochen. Vol. 1.

Solomon, Z. (1993). *Combat Stress Reaction*: *The Enduring Toll of War*. New York: Springer.

Stabell, B., and Stabell, U. (1998). Chromatic rod-cone interaction during dark adaptation, *Journal of the Optical Society of America*, 15(11), 2809-2815.

Stabell, B., and Stabell, U. (1979). Rod and cone contributions to change in hue with eccentricity, *Vision Research*, 19, 1121-1125.

Stanley, P.F., Tanzer, D.J., and Schallhorn, S.C. (2008). Laser refractive surgery in the United States Navy. *Current Opinions in Ophthalmology*, 19(4), 321-4.

Stansfeld, S.A., Berglund, B., Clark, C., Lopez-Barrio, I., Fischer, P., Ohrstrom, E., Haines, M.M., Head, J., Hygge, S., van Kamp, I., and Berry, B.F. (2005). Aircraft and road traffic noise and children's cognition and health: A cross-national study, *Lancet*, 365, 1942-1949.

Stevens, S.S. (1975). *Psychophysics: Introduction to its perceptual, neural and social prospects*. New York: Wiley.

Stiles, W.S. (1929). The effect of glare on the brightness difference threshold, *Proceedings of the Royal Society (London*) B 104, 322-350.

Stiles, W.S., and Crawford, B.H. (1937). The effect of a glaring light source on extrafoveal vision, *Proceedings of the Royal Society (London)* B, 122, 255-280.

Stocks, J.M., Taylor, N., Tipton, M.J., and Greenleaf, J.E. (2004). Human physiological responses to cold exposure, *Aviation, Space, and Environmental Medicine*, 75, 444-457.

Stone, B.M., and Turner, C. (1997). Promoting sleep in shiftworkers and intercontinental travelers, *Chronobiology International, 14*(2), 133-143.

Stone, B.M., Turner, C., Mills, S.L., and Nicholson, A.N. (2000). Hypnotic properties of melatonin, *Sleep,* 23, 663-69.

Suhner, A., Schlagenhauf, P., Hofer, I., Johnson, R., Tschopp, A., and Steffen, R. (2001). Effectiveness and tolerability of melatonin and zolpidem for the alleviation of jet lag, *Aviation Space and Environmental Medicine*, 72(7), 638-646.

Swanson, M.W. (1998). Changes in crystalline lens density with age, and fundus reflections. In: Benjamin, W.J. (Ed.), *Borish's Clinical Refraction*. Philadelphia: W.B Saunders Company.

Takahashi, J.S., DeCoursey, P.J., Bauman, L., and Menaker, M. (1984). Spectral sensitivity of a novel photoreceptive system mediating entrainment of mammalian circadian rhythms, *Nature*, 308, 186-88.

Talbott, E.O., Brink, L.L., Burks, C., Palmer, C., Engberg, R., Cioletti, M., and Ingman, C. (1996). Occupational noise exposure, use of hearing protectors over time and the risk of high blood pressure, *Proceedings of Internoise '96*. U.K.: Institute of Acoustics. p. 2131-2136.

Tanzer, D.J., Isfahani, A., Schallhorn, S.C., LaBree, L.D., and McDonnell, P.J. (1998). Photorefractive keratectomy in African Americans including those with known dermatologic keloid formation, *American Journal of Ophthalmology*, 126, 625-9.

Task, H. L, and Gleason, G. 1993. Eyepiece focus for night vision goggles (NVGS) and helmet-mounted displays: Fixed or adjustable? Wright-Patterson AFB, OH: Armstrong Laboratory. Preliminary report of two focus studies prepared for the Fall 1993 TARP Meeting.

Tennyson, E.A. (2001). How to buy a headset: A guide to get you started, *AOPA Flight Training*, 44(4).

Teo, R.K.C., and Ferguson, D.A. (1986). The acute effects of ethanol on auditory event-related potentials, *Psychopharmacology*, 90(2), 179-184.

Teschke, K., Nichol, A.M., Davies, H., and Ju, S. (1999). Whole body vibration and back disorders among motor vehicle drivers and heavy equipment operators: A review of the scientific evidence. Report to the Workers' Compensation Board of British Columbia.

The European Union. (2002). European Union Directive 2002/44/EC (for HAV and WBV), The minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (vibration), *The European Union.*

The University Health Center. (2008). Alcohol and athletic performance. Retrieved on 15 August 2008 from: http://www.uhs.uga.edu/aod/athletic-performance.html

Thomas, J.R., Ahlers, S.T., House, J.F., and Schrot, J. (1989). Repeated exposure to moderate cold impairs matching-to-sample performance, *Aviation, Space, and Environmental Medicine*, 1989, 1063-1067.

Thomas, M., Sing, H.C., and Belenky, G. (1993). Cerebral glucose utilization during task performance and prolonged sleep loss, *Journal of Cerebral Blood Flow and Metabolism*, 13(1), S351.

Thomson, D. (2001). Methods of assessing cataracts and the effect of opacities on vision, *Optometry Today-Optics Today*. 26-30. http://www.optometry.co.uk/articles/ 20010601/thomson.pdf

Thompson, M. (2008). America's Medicated Army, *Time Magazine*, June, 38-42.

Tikuisis, P., and Keefe, A.A. (2007). Effects of cold strain on simulated sentry duty and marksmanship, *Aviation, Space, and Environmental Medicine*, 78, 399-407.

Tilley, A.J., Wilkinson, R.T., Warren, P.S.G., Watson, B., and Drud, M. (1982). The sleep and performance of shift workers, *Human Factors*, 24, 629-641.

Toff, W.D., Jones, C.I., Ford, I., Pearse, R.J., Watson, H.G., Watt, S.J., Ross, J.A.S., Gradwell, D.P., Batchelor, A.J., Abrams, K.R., Meijers, J.C.M/. Goodall, A.H., and Greaves, M. (2006). Effect of hypobaric hypoxia, simulating conditions during long-haul air travel, on coagulation, fibrinolysis, platelet function, and endothelial activation, *Journal of the American Medical Association*, 295, 2251-2261.

Tredici, T.J., and Flyinn, W.J. (1987). The use of contact lenses by USAF aviators, *Aviation, Space, and Environmental Medicine*, 58(5), 438-443.

Troy, S.M., Lucki, I., Unruh, M.A., Cevallos, W.H., Leister, C.A., Martin, P.T., Furlan, P.M.B., and Mangano, R. (2000). Comparison of the effects of zaleplon, zolpidem, and triazolam on memory, learning, and psychomotor performance, *Journal of Clinical Psychopharmacology*, 20(3), 328-337.

Tyler, D.B. (1947). The effect of amphetamine sulfate and some barbiturates on the fatigue produced by prolonged wakefulness, *American Journal of Physiology*, 150, 253-262.

Upile, T., Sipaul, F., Jerjes, W., Singh, S., Nouraei, S.A.R., Maaytah, M.E., Andrews, P., Graham, J., Hopper, C., and Wright, A. (2007). The acute effects of alcohol on auditory thresholds, *BMC Ear Nose Throat Disorders*, 7, 4.

U.S. Air Force School of Aerospace Medicine. (2008). Official Air Force Approved Aircrew Medications. Retrieved 15 July 2008 from: https://kx.afms.mil/kxweb/dotmil/kj.do?functionalArea=WaiverGuide

U.S. Army Aeromedical Research Laboratory and the Army Safety Center. (1996). *Leader's Guide to Crew Endurance.* Fort Rucker, AL: U.S. Army Safety Center.

U.S. Bureau of Labor Statistics. (2004). Workers on flexible and sift schedules in May 2004. Retrieved 7 Nov 2008 from: https://data.bls.gov/cgi-bin/print.pl/news.release/flex.nr0.htm

U.S. National Library of Medicine and the National Institutes of Health. (2004). Drug information: Temazepam. In: *Medline Plus*.

U.S. Navy Aerospace Medical Research Laboratory. (2001). *Performance maintenance during continuous flight operations.* Pensacola, FL: U.S. Navy Aerospace Medical Research Laboratory.

U.S. Navy Safety Center. (2008). Acquisition safety – Noise control aboard ships. Retrieved 15 August 2008 from: http://safetycenter.navy.mil/acquisition/noise/index

van de Pol, C., Greig, J.L., Estrada, A., Bissette, G.M., and Bower, K.S. (2007). Visual and flight performance recovery after PRK or LASIK in helicopter pilots, *Aviatian Space and Environmebntal Medicene*, 78(6), 547-53.

van de Pol, C., Schallhorn, S.C. , Brown, M.C., Tanzer, D.J., Kaupp, S.E., and Lattimore, M.R.J. (2000). Effect of altitude on anterior and posterior corneal curvature: LASIK, PRK, and RK compared. *Association of Research in Vision and Ophthalmology,* Ft. Lauderdale, FL.

van den Berg, T.J.T.P. (1991). On the relation between glare and stray light, *Documenta Ophthalmologica*, 78, 177-181.

van den Berg, T.J.T.P., Ijspeert, J.K., and de Waard, P.W.T. (1991). Dependence of intraocular stray light on pigmentation and light transmission through the ocular wall, *Vision Research*, 31(7/8), 1361-1367.

van Dongen, H.P.A., Maislin, G., Mullington, J.M., and Dinges, D.F. (2003). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation, *Sleep,* 26(2), 117-126.

van Rijn, L.J., Nischler, C., Gamer, D., Franssen, L., de Wit, G., Koper, R., Vanhoff, D., Grabner, G., Wilhelm, H., Volker-Dieben, H.J., and van den Berg, T.J.T.P. (2005a). Measurement of stray light and glare: Comparison of Nyktotest, Mesotest, stray light meter, and computer implemented stray light meter, *British Journal of Ophthalmology*. 89, 345-351. http://bjo.bmjjournals.com

van Rijn, L.J., Wilhelm, H., Emesz, M., Kaper, R., Heine, S., Nitsch, S., Grabner, G., and Volker-Dieben, H.J. (2005b). Relation between perceived driving disability and score of vision screening tests, *British Journal of Ophthalmology,* 86, 1262-1264. http://bjo.bmjjournals.com

Vermeeren, A., Danjou, P.E., and O'Hanlon, J.F. (1998). Residual effects of evening and middle-of-the-night administration of zaleplon 10 and 20 mg on memory and actual driving performance, *Human Psychopharmacology and Clinical Experience*, 13, S98-S107.

Vidacek, S., Kaliterna, L., Radosevic-Vidacek, B, and Folkard, S. (1986). Productivity on a weekly rotating shift system: Circadian adjustment and sleep deprivation effects? *Ergonomics,* 29(12), 1583-1590.

Vos, J.J. (2003). On the cause of disability glare and its dependence on glare angle, age and ocular pigmentation, *Clinical and Experimental Optometry (Journal of the Australian Optometric Association),* 86(6), 363-370.

Vos, J.J. (1963). Contribution of the fundus oculi to entoptic scatter, *Journal of the Optical Society of America*, 53, 1449-1451.

Vos, J.J., and Boogaard, J. (1963). Contribution of the cornea to entoptic scatter, *Journal of the Optical Society of America*, 53, 869-873.

Vos, J.J., and Bouman, M.A. (1964). Contribution of the retina to entoptic scatter, *Journal of the Optical Society of America*, 54, 95-101.

Vos, J.J., and van den Berg, T.J.T.P. (1999). Report on disability glare. In: *CIE Collection 1999: Vision and Colour. Physical Measurement of Light and Radiation. CIE* 135-1999, 1-9.

Wangsa-Wirawan, N.D., and Linsenmeier, R.A. (2003). Retinal oxygen: Fundamental and clinical aspects, *Archives of Ophthalmology*, 121(203), 547-557.

Ward, D. (1993). Developments in noise induced hearing loss during the last 25 years. In: Vallet, M. (Ed.), *Noise as a Public Health Problem*. Arcuil Cedex, France: INRETS.

Wasserman, D.E. (2003). Occupational vibration: Are you at risk? Retrieved 26 July 2008 from Quest Technologies website: http://www.quest-technologies.com/ APP_Notes/Vibration_Paper.pdf

Wasserman, D., Taylor, W., and Behrens, V. (1982). Vibration White Finger Disease In: U.S. Workers Using Chipping and Grinding Hand Tools. Vol.1, Epidemiology. NIOSH Pub. 82-101.

Waterhouse, J., Reilly, T., and Atkinson, G. (1997). Jet-lag, *The Lancet*, 350, 1611-1616.

Waterhouse, J., Nevill, A., Edwards, B., Godfrey, R., and Reilly, T. (2003). The relationship between assessment of jet lag and its symptoms, *Chronobiology International,* 20, 1061-73.

Weaver, D.R., and Reppert, S.M. (1996). The Mel 1a melatonin receptor gene is expressed in human suprachiasmatic muclei, *Neuroreport*, 8(1), 109-112.

Wechsler, H., Lee, J.E., Kuo, M., Seibring, M., Nelson, T.F., and Lee, H. (2002). Trends in college binge drinking during a period of increased prevention efforts, *Journal of American College Health*, 50, 203-217.

Weiss, B., and Laties, V.G. (1962). Enhancement of human performance by caffeine and the amphetamines, *Pharmacology Review*, 14, 1-36.

Wells, M.H., and Griffin, M.J. (1984). Benefits of helmet-mounted display image stabilization under whole-body vibration, *Aviation, Space, and Environmental Medicine*, 55, 13-18.

Wells, M.H., and Griffin, M.J. (1987a). Flight trial of a helmet-mounted display image stabilization system, *Aviation, Space, and Environmental Medicine*, 58, 319-322.

Wells, M.H., and Griffin, M.J. (1987b). Performance with helmet-mounted sights. Institute of Sound and Vibration Research, Technical Report 152.

Wells, M.H., and Griffin, M.J. (1987c). A review and investigation of aiming and tracking performance with head-mounted sights, *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-17, 210-221.

Wells, M.J., and Haas, M. (1992). The Human factors of helmet-mounted displays. *Electro-Optical Displays*. Karim, M.A., (Ed.). Boca Raton, FL: CRC Press.

Wesensten, N., Balkin, T., Thorne, D., Killgore, R., Reichardt, R.M., and Belenky, G. (2004). Caffeine, dextroamphetamine, and modafinil during 85 hours of sleep deprivation. I. Performance and alertness effects, *Aviation Space and Environmental Medicine*, 75(4 [Section II]), B108.

Wesensten, N., Belenky, G., Kautz, M.A., Thorne, D., Reichardt, R.M., and Balkin, T. (2002). Maintaining alertness and performance during sleep deprivation: Modafinil versus caffeine, *Psychopharmacology (Berlin)*, 159(3), 238-247.

Wesnes, K., and Warburton, D.M. (1984). A comparison of temazepam and flurazepam in terms of sleep quality and residual changes in performance, *Neuropsychobiology*, 11, 255-259.

Wesnes, K., and Warburton, D.M. (1986). Effects of temazepam on sleep quality and subsequent mental efficiency under normal sleeping conditions and following delayed sleep onset, *Neuropsychobiology*, 15, 187-191.

Westerman, R.A. (2004). Hypoxia familiarisation training by the reduced oxygen breathing method, *ADF Health*, 5, 11-15

Wickens, C.D., and Hollands, J.G. (2000). *Engineering Psychology and Human Performance (3rd Ed.)*. New Jersey: Prentice Hall Inc.

Widnall, S. (1971). Helicopter noise due to blade-vortex interaction. *The Journal of the Acoustical Society of America*, 50(1B), 354-365.

Wiegmann, D.A., and Shappell, S.A. (2003). *A Human Error Approach to Aviation Accident Analysis: The Human Factors Analysis and Classification System*. Farnham, UK: Ashgate Publishing.

Wiley, R.W. (1993). Military research with contact lenses. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 93-19.

Winfield, R.H. (1941). *The use of Benzedrine to overcome fatigue on operational flights in coastal command.* Flight Personnel Research Committee Report. London: Flying Personnel Research Committee; Report No. FPRC-361.

Winfree, A.T., (1987). Circadian rhythms: Hamsters without jet lag, *Nature,* 330(6146), 311-312.

Wisor, J.P. (2002). Disorders of the circadian clock: Etiology and possible therapeutic targets, *Current Drug Targets CNS Neurological Disorders,* 1, 555-566.

Woodruff, M. (1988). How to avoid jet lab, *Nature*, 331(6153), 217.

Wooten, B.R., and Geri, G.A. (1987). Psychophysical determination of intraocular light scatter as a function of wavelength, *Vision Research*, 27(8), 1291-1298.

Wright, J.E., Vogel, J.A., Sampson, J.B., Knapik, J.J., and Patton, J.F. (1983). Effects of travel zones (jet lag) on exercise capacity and performance, *Aviation Space and Environmental Medicine*, 54, 132-37.

Wright, K.P., Hull, J.T., and Czeisler, C.A. (2002). Relationship between alertness, performance, and body temperature in humans, *American Journal of Physiology—Regulatory, Integrative, and Comparative Physiology*, 283, R1370-R1377.

Wright, N., and McGown, A. (2001). Vigilance on the civil flight deck: Incidence of sleepiness and sleep during long-haul flights and associated changes in physiological parameters, *Ergonomics,* 44(1), 82-106.

Wyatt, J.K., Cajochen, C., Ritz-De Cecco, A., Czeisler, C.A., and Dijk, D. (2004). Low-dose repeated caffeine administration for circadian-phase-dependent performance degradation during extended wakefulness, *Sleep*, 27(3), 374-381.

Yesavage, J.A., and Leirer, V.O. (1986). Hangover effects on aircraft pilots 14 hours after alcohol ingestion: A preliminary report, *The American Journal of Psychiatry*, 143,1546-1550.

Zingler, V.C., Denecke, K., Jahn, K., von Meyer, L., Krafczyk, S., Krams, M., Elfont, R., Brandt, T., Strupp, M., and Glasauer, S. (2007). The effect of nicotine on perceptual, ocular motor, postural, and vegetative functions at rest and in motion, *Journal of Neurology*, 254(12), 1689-1697.

Part Five

Meeting the HMD Design Challenge

The goal of any designer for any system, helmet-mounted displays (HMDs) included, is to develop a system that provides optimized performance for the intended user in the intended environment. For military HMDS, the intended user is the Warfighter, and the operational environment is the battlespace. This is a truly difficult task for the HMD designer. The innumerable factors that must be considered in a design are diverse and frequently in contradiction. These factors obviously include optical and acoustical engineering parameters. Next are the human-related issues of vision and audition. These are joined by ergonomic, biodynamic, and human factor considerations. In the end, there may be no “optimal” HMD design, but instead, a variety of designs that are task and user specific.

17 GUIDELINES FOR HMD DESIGN

James E. Melzer
Frederick T. Brozoski
Tomasz R. Letowski
Thomas H. Harding
Clarence E. Rash

Helmet-mounted displays (HMDs) have been in development since the 1960s. Now, almost five decades later, the technology has improved significantly; HMDs have made some inroads into commercial applications (Ellis, 1995; Kalawsky, 1993; Pankratov, 1995); their use has become standard within the military community for flight applications, training and simulation (Simons and Melzer, 2003); and they are rapidly expanding into military applications for the dismounted and vehicular-mounted Warfighter (see Chapter 3, *Introduction to Helmet-Mounted Displays*). Unfortunately, design guidance for HMDs has not kept pace. This is due in part to the rapid advances in enabling technologies (e.g., micro-electromechanical devices, microprocessors, emissive image sources, microdisplays).[1] However, it is mostly because HMDs are both engineering- and human-centered systems; whenever humans are a key system component, their complex sensory, neural mechanisms and their variability across the population makes the design of HMDs and the human-machine interface extremely challenging. This, in turn, makes universal design guidelines equally challenging.

This is not to imply that the design community has been negligent in the development of guidelines. In a 1972 symposium on visually-coupled systems sponsored by the U.S. Air Force System Command's Aerospace Medical Division held at Brooks Air Force Base, Texas, participates attempted to address many of the fundamental design issues for HMDs (Birt and Task, 1973). Hughes, Chason and Schwank (1973) provided an overview of the history and the known and potential psychological problems of HMDs and included an extensive annotated bibliography of relevant material on such issues as eye dominance, brightness disparity, helmet-mounted displays/helmet-mounted sights, retinal rivalry, and others identified during the 1972 symposium. Chisum (1975) expanded this discussion by presenting visual considerations associated with the head-coupled aspects of HMDs.

As a special subset of displays, HMDs are subject to the practices for display development in general, many of which are based on decades of human performance research. Two of the most comprehensive volumes are Farrell and Booth's (1984) *Design handbook for Imagery Interpretation Equipment* and Boff and Lincoln's (1988) *Engineering Data Compendium: Human Perception and Performance*.

HMDs are also a specialized class of displays called head-up displays (HUDs), defined as transparent, fixed location displays that present data without obstructing the user's view (Figure 17-1). Developed originally as gun sights for military aircraft, they have expanded into commercial aircraft (Steenblik, 1989) and recently have become an option in some automobiles (Oldsmobile Club of America, 2006). HUD guidelines concentrate mostly on symbology and related display criteria such as clutter, dynamic response and viewing comfort issues, and many of these criteria have a firm foundation in human factors and human perception (Prinzel and Risser, 2004; Ververs and Wickens, 1998; Weintraub, 1992; Wickens, 1997; Wickens, Fadden, Merwin, and Ververs, 1998**)**. Two important reference books on HUDs are Wood and Howells' (2001) *Head-Up Displays* and Newman's (1995) *Head-Up Displays, Designing the Way Ahead*.

However, of the vast amount of research conducted over the last half-century, only four reference books have been written specifically for HMDs; and the first three of these focus on aviation applications only. The first

---

[1] Suggested reading on these enabling technologies is Brennesholtz, M.S., and Stupp, E.H. (2008), *Projection Displays*, John West Sussex, UK: Wiley and Sons.

book, *Head-Mounted Displays: Designing for the User* (Melzer and Moffitt, 1997), addresses HMD development for fixed-wing aircraft. It could be considered an engineering guide with its coverage of the traditional engineering design approach, but it also places a significant emphasis on the end user, addressing a wide array of human-centered disciplines required for the design of head-mounted virtual reality, industrial and military displays. Topics include optical requirements, lens designs, cybersickness, eye strain, head-supported weight,[2] stereoscopic imagery, anthropometry, and user acceptance. The book also introduces the potential of HMDs to serve as an interface for brain-actuated control functions, a concept explored in this volume (see Chapter 19, *The Potential of an Interactive HMD*).

Figure 17-1. Examples of head-up displays (HUDs): (left) a HUD in a fighter cockpit and (right) a HUD designed for aircraft simulation (Rockwell Collins).

Melzer and Moffitt's book was quickly followed by *Helmet-Mounted Displays and Sights* (Velger, 1998) that is described by its author as "an in-depth, design practitioner's study of helmet-mounted display and sight technology (HMD/HMS)." Velger's book discusses human factors associated with the use of HMDs and details image source and display technologies. It offers practical recommendations for evaluating various optical designs and technologies, selecting appropriate image sources and displays, and applying the human-centered design concepts to helmet display systems. The book also provides insight into head-tracking systems and techniques for stabilizing display images, examining the effects of aircraft vibration on HMD effectiveness.

The third aviation-oriented HMD book was *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft* (Rash, 2001). This book differs from the first two reference books in that it focuses on the use of HMDs in U.S. Army rotary-wing aircraft, emphasizing the important issues associated with interfacing HMDs with the U.S. Army aviator. Topics include optics, vision, acoustics, audition, biodynamics, safety, ergonomics, and visual human factors.

While these three books are aviation-focused, the last of the HMD-specific books (National Research Council, 1997) is the end product of a special panel, *Human Factors in the Design of Tactical Display Systems for the Individual Soldier*, conducted by the Committee on Human Factors, National Research Council Washington, DC (National Research Council, 1995). This panel was established at the request of the U.S. Army Natick Soldier Research, Development, and Engineering Center (NSRDEC),[3] Natick, MA, for the purpose "of explicating the human factors issues and approaches associated with the development, testing, and implementation of HMD technology in the (U.S. Army's) Land Warrior System." More specifically, the panel was charged with examining the relationship among tactical information needs of individual Warfighters; the possible devices available at that

[2] The terms *head-supported weight* (HSW) and *head-supported mass* (HSM) are used interchangeably.
[3] Also known as the U.S. Army Natick Soldier Systems Center, Natick, MA.

time and in the near future for processing, transmitting, and displaying such information, and the human performance implications of the use of such devices.

In this chapter we summarize general guidelines and recommendations drawn from the references cited above (and others) that are useful in developing the "optimal" HMD design. In addition, we summarize the perceptual and cognitive principles presented in the previous chapters of this volume, and discuss design tradeoffs and their impact on system and human performance. However, the reader is cautioned that each HMD design is application and user specific. There is no single set of design criteria or guidelines that can be blindly followed. The designer must apply those guidelines that are best fitted to the desired design application and user population.

User-Centered Design Focus

The idea of a single optimal HMD design is an unobtainable goal. Simply stated, the concept of a one-size-fits-all HMD is a false one, primarily because of the wide variation in user tasks and users themselves; an HMD designed for the pilot of a fast fighter jet flying at 10,000 feet (3048 meters) will not meet the needs of a helicopter pilot flying close to the ground, and neither of these will meet needs of a dismounted Warfighter negotiating a brush-laden forest. Although there are some common design features – primarily driven by human perceptual and anthropometric limitations – across the various HMD configurations, many of the performance requirements and tradeoffs are user-, application- and environment-driven.

The following sections address the guidelines for HMD design while attempting to frame them in the context of a user-centric design process. It will become clear that there are a number of different configurations and tradeoffs available to the designer, each of which has advantages and disadvantages depending on the user's characteristics and application. While significant breakthroughs have been realized in lightweight protective materials, optical design and fabrication, head/eye-orientation tracking, and miniature flat-panel image sources since Ivan Sutherland's true HMD[4], the following important questions must still be asked: "Who is the user?" and "What will he/she be doing with the HMD?" Once these questions are answered and key performance requirements identify, *suboptimizing* can begin. In doing so for the Warfighter, the minimum set of HMD features – separate from those of the protective helmet/platform – that are sufficient to allow the Warfighter to accomplish his mission without affecting his safety are indentified. The beginning of the chapter emphasizes these minimum HMD features because the Warfighters' mission and environment demands this suboptimizing process to simplify the design, reduce head-supported weight and drive down the cost. Thus, the recommendations in the following sections are best considered as a shopping cart of (often conflicting) advice, where the designer must make tradeoffs based on the specific environment, user population and application.

Design criteria, guidelines and recommendations are grouped and presented in the following sections for optical/visual, biodynamic, acoustic/auditory, perceptual/cognitive and user adjustment parameters.

---

[4] Ivan Sutherland (1968) is credited with implementing the first virtual reality system. Using wire-frame graphics and a head-mounted display (HMD), it allowed users to visually occupy the same space as virtual objects. An even earlier head-mounted sighting system is described within the context of the historical pursuit for an accurate measure of longitude at sea (Sobel, 1996). In 1610, Galileo Galilei discovered the moons of Jupiter and soon concluded that knowledge of their precise position could give navigators at sea an accurate measure of Greenwich Mean Time (GMT) and therefore help determine their east/west location. In 1618, he designed a sighting system to aid in observing what are now referred to as the Galilean satellites. The navigator sat in a special gimbaled chair designed to compensate for the motion of the ship and observed the moons with a *celatone*, a face-mounted device somewhat like a gas mask that held one, or possibly two telescopes. Though no drawings or models exist, it is interesting to be able to push back the invention of the earliest head-mounted display sighting system to one of the most brilliant inventors in history.

Optical/Visual Guidelines and Recommendations

The following optical/visual parameters and issues are addressed:

- Ocularity (monocular, biocular or binocular)
- Field-of-view (FOV)
- Resolution
- Pupil-forming versus non-pupil-forming optics
- Exit pupil and eye relief
- Optical distortion
- Luminance and contrast
- See-through versus non-see-through considerations
- Considerations for helmet-mounted sensors

Ocularity

Ocularity refers to whether the HMD provides monocular, biocular and binocular imagery as defined by:

- *Monocular* – a single image source is viewed by a single eye
- *Biocular* – a single image source viewed by both eyes
- *Binocular* – each eye views an independent image source

Moffitt (2008) further subdivides monocular and binocular HMD configurations[5] into categories that focus on their respective applications (Table 17-1). The variety of monocular configurations is discussed first. If the HMD will be used to provide moving map or text information for a dismounted Warfighter, or to allow a pilot to view imagery with the simplest, lightest and least costly system, a monocular design is best. Variations include the Joint Helmet-Mounted Cueing System (JHMCS) for fixed-wing (F/A-18, F-16 and F-15) aircraft and the Integrated Helmet Display Sighting System (IHADSS) for the Army's AH-64 Apache helicopter (see Chapter 3, *Introduction of Helmet-Mounted Displays*). The AN/PVS-14 night vision goggle (NVG) and the SO35A for the Land Warrior program are also monocular. The single image source configuration reduces cost, head-supported weight/mass, power consumption and simplifies the opto-mechanical alignment requirements. There is the potential for a laterally asymmetric center-of-mass (CM)[6] and issues associated with focus, eye dominance, binocular rivalry and ocular-motor instability (Moffitt, 1989; Rash and Verona, 1992), although these issues have not been shown to have an insurmountable impact on performance if the user is properly trained. Table 17-2 presents a summary of performance and ergonomic benefits and disadvantages of monocular optical designs.

Questions of how monocular HMDs interact with the dominant eye continue. Moffitt (2008) cites research (Mapp, Ono and Barbeito, 2003) that defines the dominant eye as the one that individuals use for monocular sighting tasks. This is the situation for Warfighters wearing the AN/PVS-14 NVG or the Land Warrior HMD, where they place the HMD or NVG over their *non-dominant* eye, leaving their dominant eye clear for weapon aiming.

To achieve the widest field-of-view (FOV) possible, the HMD must be either *biocular* or *binocular*. The biocular/binocular approach is more complex than the monocular design, but because it stimulates both eyes, it eliminates some of the visual rivalry issues associated with monocular displays and reduces cost because there is

---

[5] Moffitt (2008) considers the biocular design to be a subset of binocular HMDs as "binocular HMD using a single display," although they will be separated in this chapter for clarity.

[6] The terms center-of-mass (CM) and center of gravity are used interchangeably.

only a single image source as in the case of the AN/PVS-7D NVG. Taxonomy of binocular HMD configurations with design considerations and examples is presented in Table 17-3.

Table 17-1.
Taxonomy of monocular HMD configurations, considerations and examples.
(adapted from Moffitt, 2008)

Designation	Description	Considerations	Example
Monocular HMD Configurations			
Compact offset	The compact offset monocular is positioned for brief viewing away from the user's forward line-of-sight. It is used for applications not requiring extended viewing.	Imagery is not in spatial correspondence with the outside world. May distract user from primary viewing tasks. Does not block forward line-of-sight, preserving binocular vision for motor and navigation tasks.	• Rockwell Collins *SO35A*[7] • Sportvue® *MC2*[8]
Opaque	Similar to the compact offset, except the HMD is positioned directly in the user's line-of-sight. Provides only graphic or text information, it is not intended to present imagery that is overlayed on the outside world.	Better configuration for longer duration viewing, though with potential issues with binocular rivalry between viewing and non-viewing eye.	• Eyetop™ Centra[9] • Liteye *Le-700A*[10]
Video see-through	A head-mounted camera or sensor is used as one source of imagery in-line with the viewing eye.	Could distract user from other tasks with FOV that is smaller than the non-viewing eye. Camera and HMD need to be in corresponding alignment. If offset from eye line-of-sight, may create viewing artifacts.	
Optical see-through	Imagery superimposed over see-through view.	Reduced see-through in viewing eye may induce perceptual artifacts. Monocular viewing of imagery may rival outside world view.	• IHADSS • JHMCS

Viewing imagery with two eyes vs. one has been shown to yield improvements in detection as well as providing a more comfortable viewing experience (Boff and Lincoln, 1988; Moffitt, 1997). Table 17-4 presents a summary of performance and ergonomic benefits and disadvantages of biocular and binocular optical designs.

Since it is a two-eyed viewing system, the biocular design is subject to a much more stringent set of alignment, focus and adjustment requirements.[11] This generally has been deemed a difficult design for flight applications,

---

[7] The Rockwell Collins' SO35A is the HMD used on the U.S. Army's Land Warrior program.
[8] Motion Research Corporation, Seattle, WA.
[9] Ingineo, SAS, Villiers Le Bel, France.
[10] Liteye Systems, Centennial, CO.
[11] For *absolute* horizontal alignment, in *non-see-through* HMDs, the binocular alignment is not critical as long as it agrees with the focus to within ±1/4 diopter. For *relative* horizontal alignment, in *see-through* HMDs, the horizontal binocular alignment must be within 5 to 10 arcminutes of the desired vergence distance and the focus must agree to within ±1/4 diopter.

For *absolute* vertical alignment, in *non-see-through* HMDs, the binocular alignment must be within 10 arcminutes. For *relative* vertical alignment, in *see-through* HMDs, the binocular alignment must be within 3 to 6 arcminutes (Boff and Lincoln, 1988; Moffitt, 1997; Self, 1986).

Table 17-2.
Human performance considerations of monocular optical design approaches.
(adapted from National Research Council, 1997; Melzer, 2006)

Ocularity	Human Performance, Sensory and Ergonomic Considerations	
	Benefits	Disadvantages
Monocular (one image source viewed by one eye)	• Minimum weight • Simplest HMD; less stringent alignment • Eye with no display remains dark adapted and continues to sample real world • Least expensive	• Possible visual rivalry problems, such as target suppression (involuntary), "cognitive switching," ocular-motor instability, eye dominance and focus issues • Asymmetric CM • Smallest FOV; least information capability; more and larger head movements required • No stereoscopic depth information • Difficult to navigate on uneven terrain

Table 17-3.
Taxonomy of binocular HMD configurations, design considerations and examples.
(after Moffitt, 2008)

Designation	Description	Considerations	Examples
Binocular Viewing Configurations			
Opaque	Provides completely occluded view of immersive synthetically-generated monoscopic or stereoscopic imagery.	Users are visually isolated from the real world. Potentially a safety hazard if users rely on this imagery for navigation.	MyVu *Personal Media Viewer* Daeyang *i-Vision*
Video see-through	Outside world imagery provided only through video camera(s)	Single camera provides monoscopic imagery. Separation of two cameras of greater than 2.5 inches can create hyperstereo imagery	Mirage *Augmented Reality System* Trivision *Scout 2*
Optical see-through	Video imagery is projected on a see-through combiner. Can provide geo-spatially registered imagery	Video overlay may confuse users as in the monocular see-through.	Rockwell Collins *SR100A* Rockwell Collins *JSF RSV HMD*
Extended Binocular Configurations			
Partial binocular overlap	Optical channels are canted inward (convergent) or outward (divergent) to increase the horizontal FOV	Critical binocular alignment requirements in the overlap region. Binocular rivalry possible in the unpaired binocular-monocular boarder region.	Rockwell Collins *SR100A*
Paneled or optically tiled	Individual display modules are optically "tiled" next to each other to enlarge the FOV	Difficult design in see-through. Overlap regions must be corrected for content, focus, distortion, alignment, color, and contrast.	Sensics *piSight*
Mixed resolution or dichoptic	Small FOV, high resolution image in one eye. Larger FOV, lower resolution image in other eye.	Image fusion using blur suppression is assumed. Ability of a wide range of users to fuse the images is not known.	DARPA's *MANTIS*

because it requires two optical paths of equal length, which puts the image source in the middle of the head, generally high and forward. In addition, the display luminance is cut in half because the single image source is split in order to be presented to both eyes.

Table 17-4.
Human performance considerations of biocular/binocular optical design approaches.
(adapted from National Research Council, 1997; Melzer, 2006)

Ocularity	Human Performance, Sensory and Ergonomic Considerations	
	Benefits	Disadvantages
Biocular (one image source viewed by both eyes)	• Wider FOV, more information, easier to navigate • No interocular rivalry, as with monocular • Less complex to adjust than binocular • Simple electrical interface • Lighter weight than binocular • Less expensive than binocular	• Heavier than monocular • Reduced luminance over monocular • No stereoscopic depth information • More complex alignment than monocular • Difficult to package, generally requires center of the forehead
Binocular (two image sources viewed by both eyes)	• Can provide stereo viewing • Better depth information for mobility • Improved target recognition over monocular • Partial binocular overlap • Symmetrical CM	• Heaviest optics • Alignment and adjustments are more complex and critical than monocular • Most expensive

A binocular HMD is subject to the same stringent alignment, focus and adjustment requirements as the biocular design but with more packaging design freedom, because the designer is able to move both the optics and the image sources *away* from the face. This also means that the CM can be moved back towards the *tragion notch* to reduce biodynamic fatigue and improve safety (see Biodynamics section below). This is the most complex, most expensive and heaviest of all three optical design approaches, but one which has the advantage of providing the widest FOV possible and stereoscopic imagery from the two independent video channels. Examples are the Helmet Integrated Display Sighting System (HIDSS) (for the since-cancelled U.S. Army RAH-66 Comanche helicopter), the Joint Strike Fighter (JSF), Rotationally Symmetric Visor (RSV) HMD and the SR-100A HMD for simulation and training applications.[12] A binocular system can also take advantage of some techniques for extending the horizontal FOV without compromising resolution (see section below on Resolution Tradeoff with FOV).

Field-of-view

Field-of-view (FOV) describes how extensive the image appears to the user,[13] measured in degrees as observed by one eye (for a monocular HMD) or both eyes (for either biocular or binocular HMDs).[14] The human visual system has a total binocular FOV of 200° horizontal (H) by 130° vertical (V) (Smith and Atchison, 1997). While it is desirable to replicate this in an HMD, optical design and image source considerations limit our ability to do so.

[12] These systems are products of Rockwell Collins, Cedar Rapids, IA.
[13] Field-of-view can be defined more formally as the maximum image angle of view that can be seen through an optical device.
[14] It also may be measured as the diagonal FOV across the entire monocular or binocular field.

We must think about the FOV of an HMD differently than we do for a HUD,[15] which is located in a fixed position in front of the pilot.[16] Since HMD imagery moves with the pilot's head, and is always in a fixed location with respect to the head position, the displayed information is readily available anywhere the pilot is looking. This "unlocks" the pilot's view from the eye-to-HUD line-of-sight, contributing to the pilot's sense of self-stabilization, and lowering workload by reducing the amount and range of head movements necessary to capture the displayed symbology (Kasper et al., 1997; Szoboszlay et al., 1995; Wells, Venturino and Osgood, 1989).

Early helmet-mounted sights such as the Visual Target Acquisition System (VTAS – Belt, Kelley and Lewandowski, 1998; Dornheim, 1995) and the ODEN[17] (Friberg, 1997; Waldelöf and Friberg, 1996) projected only a targeting reticle with a FOV of 6°. These, and the early Russian helmet-mounted sights, were elegantly simple, intended only for the task of aiming missiles away from the boresight of the aircraft, and they proved to be significant force multiplier for those pilots (Arbak, 1989; Merryman, 1994).

Early experiments with a more complete HUD-like symbol suite demonstrated that most fixed-wing pilots preferred the larger 20° FOV over a smaller 12° (Melzer and Larkin, 1987). Bahill, Adler and Stark (1975) found that most saccadic eye movements were in the ±10° to ±15° range. Any stimulus outside that range typically elicits a head movement to bring the object into a more "eyes forward" viewing position. If a designer decides to locate flight symbology such as altitude or pitch ladder along the outer vertical edges of the HMD the pilot should not be required to repeatedly rotate his eyes past the 10° or 15° point, as doing so will cause eye strain and probably reduce performance.

If our goal is to create an opaque, fully-immersive visual environment for gaming or simulation and training, a large FOV is desirable in order to stimulate the ambient visual mode[18] and provide a more compelling sense of immersion. This is similar to the feeling encountered when watching a large screen IMAX® film, that of "being in" rather than "looking at" it. Patterson, Winterbottom and Pierce (2006) reviewed several perceptual studies on FOV. Allison, Howard and Zacher (1999 – cited in Patterson, et al., 2006) showed that limiting the FOV to 50° reduced the perception of self motion. Osgood and Wells (1991 – cited in Patterson, et al, 2006) showed that target acquisition in a simulated environment improved with increasing FOV, approaching a performance asymptote at 40°. Another study (Lin et al., 2002 – cited in Patterson, et al, 2006) showed increased levels of simulator sickness and "presence" up to 140° FOV. Based upon their findings, Patterson and his colleagues recommend a minimum 60°-FOV to achieve a full sense of immersion for simulator applications. One example of a wide FOV HMD is in the US Army's Aviation Combined Arms Tactics Trainer (AVCATT), a mobile, re-configurable training system for helicopter pilots that relies on the HMD for all the out-the-window visuals (Simons and Melzer, 2003) (see Chapter 3, *Introduction to Helmet-Mounted Displays*). This system uses a Rockwell Collins' HMD that provides a 100° (H) by 52° (V) FOV (recently upgraded to SXGA resolution). The price for this larger FOV is more head-supported weight. Although for these non-flight applications, it is tolerable over the training period and does not constitute a safety hazard to the user.

If the goal is a safety-of-flight-qualified HMD, then head-supported weight and CM become critically important, and a more moderate FOV of 40° horizontal by 30° vertical is acceptable. Reducing the FOV reduces head-supported weight/mass, which improves safety and reduces pilot fatigue, and the 40° horizontal FOV is within the threshold region of providing the "being in" rather than the "looking at" sensation. The IHADSS is an example of a 40° horizontal by 30° vertical FOV that has been successfully used on the US Army's AH-64 Apache helicopter since the early 1980's. The new RSV HMD for the Joint Strike Fighter also provides a 40° (H)

---

[15] Typical HUD FOVs range from 10° to 18° for conventional non-pupil-forming designs, and up to 30° for the more complex, holographic, pupil-forming curved combiner designs.

[16] Because the HUD does not move, designers must specify a relatively large "viewing eye box" within which the pilot can move his head and still see all the imagery. This drives the size of the combiners and the projection optics, which are competing for space on the very crowded forward cockpit panel.

[17] ODEN is a 1990s HMS developed by FFV Aerotech, Sweden.

[18] See Chapter 19, *The Potential for an Interactive HMD*, for an in-depth discussion of how an HMD stimulates the focal and ambient modes of vision.

by 30° (V) FOV to display symbology and real-time imagery. The AN/AVS-6 (Aviator's Night Vision Imaging System - ANVIS) NVG provides the helicopter pilot a fully overlapped 40° circular binocular FOV and also has been in use since the 1970s. For the dismounted Warfighter, the monocular AN/PVS-14 NVG also provides a 40° circular FOV. One notable exception is the QuadEye™ NVG currently in limited deployment for helicopter applications (see Chapter 3, *Introduction to Helmet-Mounted Displays*). While the ANVIS NVG has one image intensifier (I^2) tube per eye, the QuadEye™ has two per eye and provides a panoramic 100° (H) by 40° (V) FOV (see Figure 3-25 in Chapter 3, *Introduction to Helmet-Mounted Displays*). The additional outboard I^2 tubes on each eye provide the pilot with more peripheral field imagery. The problem is the additional head-supported weight and more forward CM caused by the added I^2 tubes. Even though the design is based upon a lighter 16-millimeter (mm) tube versus the standard 18-mm tube, the mass is 700 grams versus 525 grams for the standard ANVIS NVG.

In most cases, it is necessary to match the FOV of the HMD with that of the sensor to achieve a 1:1 correspondence between sensor and display FOVs to ensure an optimum task configuration. Similarly, NVGs are specified to be a unity magnification. That is, the input FOV must match the output FOV to within a few percent.

Resolution

Resolution refers to the apparent angular size of a displayed pixel or image element and the ability for the user to view and correctly interpret an object as imaged by that pixel (and others). Resolution contributes to overall image quality, but there is also a direct relationship with performance. Increased resolution means there are more pixels or image elements available to let a Warfighter see the target. Depending on the user's task, the Johnson criteria (Lloyd, 1975) determines the resolution required to *detect* ("something is there"), *recognize* ("it's a tank"), or *identify* ("it's a T-72 tank") an object of a specific size at a given distance with an increasing number of pixels per target area required.[19]

Often the resolution of an HMD is given as the number of pixels on the image source for a given FOV value, and the user is left to determine the corresponding resolution. As sensor and image source technologies have improved, so has the user demand for better resolution, approaching the one-arc minute value associated with Snellen 20/20 human visual acuity.

If the user's task does not require identifying an object at great distance or the displayed imagery will only be HUD-like symbology, it may be preferable to reduce the resolution. It has been observed that an acceptable line width for HUD-like symbology should subtend on the order of 1 milliradian (3.4 arc minutes) as observed by the user. Anything smaller tends not to be visible (Boff and Lincoln, 1988).

But, even with a high quality flat panel image source, resolution is not simply a function of the number of pixels on a given target. As discussed in Chapter 4, *Visual Helmet-Mounted Displays*, the Modulation Transfer Function (MTF) is the measure of a display system's ability to transfer modulation from target to display as a function of spatial frequency. For a system such as the IHADSS, simply calculating the MTF of the HMD is not sufficient. This is because the HMD performance must be convolved with that of the imaging sensor and the transfer of video data to the HMD. Thus, while an HMD with very high resolution may provide a high quality image, visual performance of the user's overall visual system may still be limited by the resolution (and MTF) of the imaging sensor such as the forward-looking infrared (FLIR) or video camera (Velger, 1998). For NVGs, the resolution is a function of the performance of the objective lens, the I^2 tube and the eyepiece lens. For an aircraft

---

[19] The Johnson Criteria says that for a 50% probability: 1) *Detection* requires 1.0 ± 0.25 cycles across the minimum dimension of the target, 2) *Recognition* requires 4.0 ± 0.8 cycles across the minimum dimension and, 3) *Identification* requires 6.4 ± 1.5 cycles across the minimum dimension. Increasing the probability to 90% requires an increase of 1.75X in the number of cycles across the minimum dimension (Leachtenauer, 2003). A new metric, Targeting Task Performance (TTP), that shows improvement over the Johnson Criteria has been recommended (Hixon, Jacobs and Vollmerhausen, 2004; Vollmerhausen, Jacobs and Driggers, 2003).

sensor system, such as the IHADSS, the resolution is a function of the performance of the sensor objective lens, sensor focal plane, image stabilization, image processing, video bus, HMD image source and the display optics.

Calculations of MTF for CRT-based displays are well known (Velger, 1998). MTFs of pixilated or "sampled data" displays such as Liquid Crystal Displays (LCDs) or Organic Light-Emitting Diodes (OLEDs) differ because of dependence on the phase of the input signal, as phase is shifted, there occurs a drop in modulation. Balram and Olsen (1996) and Olsen and Balram (1996) define the Multi-valued Modulation Transfer Function (MMTF), which includes the effects of frequency and phase and has a sync function-like appearance. More complex still is the effect of pixilated sensor data displayed by a pixilated image source, where changes in phase due to differences in residual distortion – barrel for the input sensor optics and pincushion for the HMD optics – can present a Moiré pattern.[20]

Resolution tradeoff with FOV

Users typically want more of both FOV *and* resolution. This is not always possible because resolution often is a direct tradeoff with FOV, as a result of the "resolution/FOV invariant" (Melzer, 1998), and are related by the equation (Figure 17-2):

$$H = F * \tan \Theta \qquad \text{Equation 17-1}$$

where F is the focal length of the collimating lens and:

- If H is the size of the image source, then Θ is the FOV, or the apparent size of the virtual image in space (which is desired to be large).
- If H is the pixel size, then Θ is the resolution or apparent size of the pixel in image space (which is desired to be small).

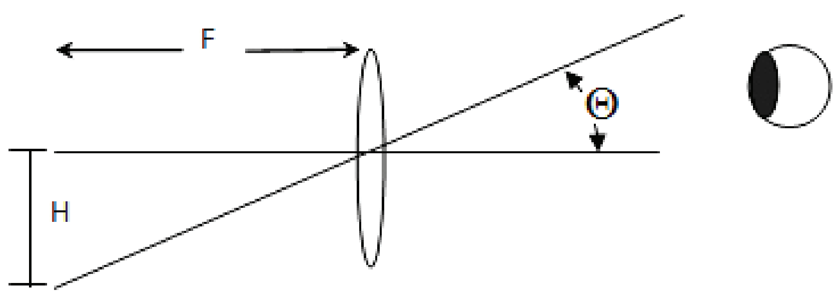


Figure 17-2. The relationship between the size of the image source and the resulting FOV or the size of the individual pixel and the resulting resolution.

Thus, the focal length of the collimating optics *simultaneously* governs the FOV *and* the resolution. For a display with a single image source, the result is either wide FOV *or* high resolution, but not both at the same time. Generally, a larger FOV is preferred in order to provide a more immersive experience. But, also, high resolution (small pixels) is desired: how high depends on the user's task. If the task is nothing more than watching simple

[20] A Moiré pattern is an undesired image artifact; a geometrical design resulting from interference when one set of straight or curved lines is superposed onto another set.

video imagery, lower resolution may be acceptable. If the task is flying a helicopter at night, close to the ground, however, the best (highest) resolution possible is required to allow the pilot to see objects as small as power lines viewed at a distance or judge altitude using ground texture. If the human eye acuity of 20/20 it has a limiting resolution of 1 arc minute and this should be one resolution goal.

Given the H = F*Tan Θ invariant, there are at least four ways to increase the FOV and still maintain resolution. These are: 1) partial binocular overlap, 2) optical tiling (which Moffitt, 2008, refers to as “paneled”), 3) high-resolution for a limited area of interest, and 4) dichoptic area of interest (which Moffitt, 2008 refers to as “mixed resolution”) (Hoppe and Melzer, 1999; Melzer, 1998). Of these, the first two will be discussed in detail, as they have been implemented in more than just a laboratory environment.

Partial binocular overlap results when the two HMD optical channels are canted either inward (*convergent* overlap) or outward (*divergent* overlap – see also Figure 3-3, Chapter 3, *Introduction to Helmet-Mounted Displays*). This latter configuration is similar to human vision with two monocular channels viewing the outward portion of the visual field and a central binocular region (Melzer and Moffitt, 1989; Melzer and Moffitt, 1991). Partial binocular overlap requires two image sources and two video channels with the optics and imagery properly configured to compensate for any residual optical aberrations.

Luning[21] is a psychophysical binocular rivalry phenomenon observed in partial overlap displays from viewing dissimilar imagery with the two eyes. Concerns have been expressed about the minimum binocular overlap needed as well as the possibility that perceptual artifacts may have an adverse impact on pilot performance. Although the studies that found image fragmentation did place some additional workload on the pilot/test subjects (Klymenko et al., 1994; 2000), the research was conducted using static imagery. Although not substantiated by rigorous studies, anecdotal evidence indicates that users viewing dynamic imagery under some degree of workload – such as flying a helicopter simulator – do not experience the detrimental effects of luning. This agrees with earlier reports, which stated that users adapt to partial overlap after 30 minutes of use (McLean and Smith, 1987). Early efforts attempted to explain the difference in the *degree* of luning observed between convergent and divergent displays with an ecological vision model. Here, convergent overlap was theorized to induce less luning because it was more “ecologically valid” than the divergent case (Melzer and Moffitt, 1991). Several techniques have been shown to be effective in reducing the rivalry effects and their associated perceptual artifacts (Melzer and Moffitt, 1991; Moffitt and Melzer, 1993).

Good HMD design practice (similar to HUD design) dictates that the binocular alignment requirements for horizontal and vertical vergence be met within the central binocular overlap region (see Footnote 4 of this chapter), regardless of whether the HMD uses an extended configuration or not. The importance of ensuring good optical quality was shown in a series of experiments conducted using canted displays without sufficient optical compensation, resulting in subjects’ reports of eyestrain (Landau, 1990).

Another method of enlarging the FOV without compromising resolution is optical tiling. In this method, a series of small-FOV high-resolution displays are arranged in a mosaic pattern, similar to a video wall. Optically overlapping the display fields minimizes the seams between the adjacent tiles (Hoppe and Melzer, 1999). The overall FOV is the equivalent of all the tiles butted together, while the resolution remains that of the individual tiles or display modules. One example is the piSight™, developed by Sensics, Inc.[22] The major difficulty with optical tiling is in positioning the image generator windows to provide good alignment and a smooth image across the tiles. Optical tiling also has been used with NVGs to enlarge the horizontal FOV (Jackson and Craig, 1999) – the QuadEye™.

A third method of enlarging the FOV involves providing mixed resolution (e.g., different resolutions for different FOVs and high-resolution insets) (Melzer, 1998; Moffitt, 2008). A low resolution, wide FOV channel is displayed to one of the user's eyes while a much higher resolution, but smaller FOV channel is displayed to the

---

[21] The term “luning” originated from the crescent-shaped edges of the circular image sources (e.g., CRT or fiber optic image bundle) (CAE Electronics, 1989).

[22] Sensics, Inc., 810 Landmark Drive, Suite 128, Baltimore, MD 21061

user's other eye (Kooi, 1993). The user fuses the two images and suppresses the low-resolution central portion in favor of the higher resolution information, while retaining the wide FOV, low-resolution portion around it. The result is a high-resolution area of interest (AOI) that is fixed in the center of a wide FOV, but lower resolution, display.[23] This concept has been implemented on the Defense Advanced Research Projects Agency's (DARPA's) Multispectral Adaptive Networked Tactical Imaging System (MANTIS) prototype HMD to provide high resolution, wide FOV, multi-spectral imagery to the dismounted Warfighter. While creative in design, Curry, Harrington and Hopper (2006) have expressed concerns with this system on perceptual grounds, which have not yet been confirmed with laboratory or field testing.

Pupil-forming and non-pupil-forming optical designs

In an HMD, the optics serve to: 1) collimate the image source (creating a virtual image, which appears to be farther away than just a few inches from the face), 2) magnify the image source (making the imagery appear larger than the actual size of the image source), and 3) relay the image source (creating the virtual image away from the image source, away from the front of the face).

There are two optical design approaches common in HMDs. The first is the *non-pupil-forming* design or simple magnifier (Cakmakci and Rolland, 2006; Fischer, 1997; Task, 1997) (Figure 17-3). It is the easiest to design, the least expensive to fabricate, the lightest and the smallest, though it does have only a short throw distance between the image source and the virtual image, forcing the designer to locate the whole assembly on the front of the head, close to the eyes. It is typically used for simple viewing applications such as the Rockwell Collins' SO35A HMD for the Land Warrior program. (See Figures 3-30, 3-31, Mounted Warrior Soldier System HMD, and 3-32, Microvision, Inc's NOMAD in Chapter 3, *Introduction to helmet-Mounted Displays*.)

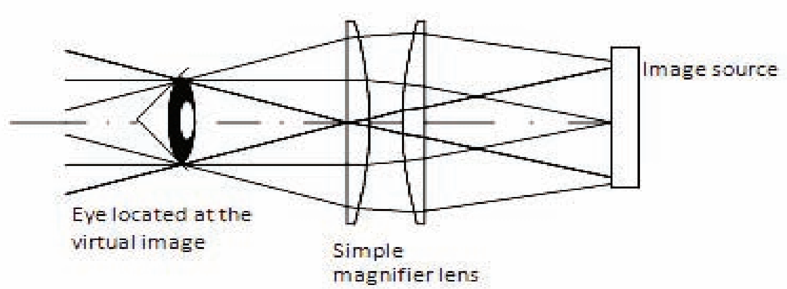


Figure 17-3. A diagram of a simple magnifier, a non-pupil-forming lens.

The second design form is the *pupil-forming* design (Figure 17-4). This is similar to the compound microscope, or a submarine periscope in which a first set of lenses creates an intermediate image of the image source. This intermediate image is *relayed* by another set of lenses to where it creates a pupil, or a hard image of the aperture stop.

---

[23] This dichoptic AOI approach is based on an eyeglass prescription technique used by optometrists known as *monovision*. A person whose eyes have limited ability to focus (a presbyope) is typically prescribed a bi- or trifocal correction. If this same person wants contact lenses, they are sometimes given a prescription in which one eye is corrected for near focus while the other is corrected for a distant focus. When the person attends to an object close by, the eye corrected for distance viewing is blurred. Similarly, when the person attends to an object at a distance, the eye corrected for viewing close up is blurred. The visual system suppresses the blurred image in favor of the non-blurred image (Schor, Landsman and Ericson, 1987). The dichoptic area of interest presents the wide field of view background image to one eye with the smaller image inset in the center. The user blurs that portion of the low-resolution image in favor of the higher resolution image.

The advantage is that the pupil-forming design provides more path length from the image plane to the eye. This gives the designer freedom to insert mirrors as required to fold the optical train away from the face and to a more advantageous head-supported weight and CM location. The disadvantages are that the additional lenses increase the weight and cost of the HMD and, outside the exit pupil, there is no imagery. The IHADSS and HIDSS HMDs are examples of pupil-forming HMDs (see Figure 3-22, Chapter 3, *Introduction to Helmet-Mounted Displays*).

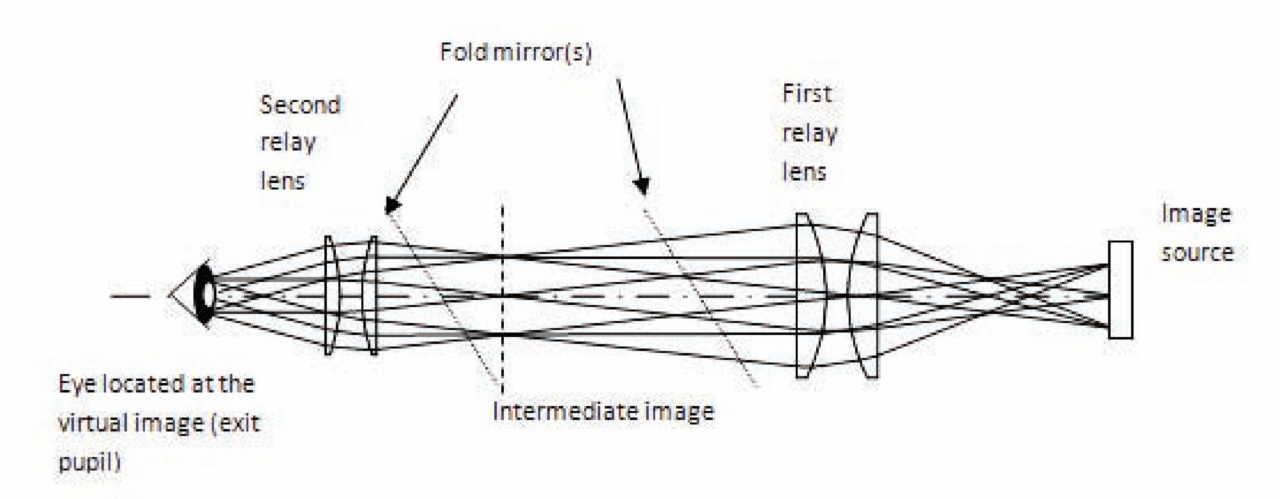


Figure 17-4. A pupil-forming optical design is similar to a compound microscope, binoculars or a periscope. Note that the increased length from image source to exit pupil provides the opportunity to insert mirrors to fold the optical path around the head.

Table 17-5 provides a summary of some of the advantages and disadvantages of pupil-forming and non-pupil-forming optical designs for HMDs.

Table 17-5.
Summary of some of the advantages and disadvantages
of pupil-forming and non-pupil-forming optical designs for HMDs.

	Non-pupil-forming (simple magnifier)	Pupil-forming (relayed lens design)
Advantages	• Simplest optical design • Fewer lenses and lighter weight • Doesn't "wipe" imagery outside of eye box • Less eyebox fit problems • Mechanically the simplest and least expensive	• Longer path length means more packaging freedom. Can move away from front of face. • More lenses provide better optical correction
Disadvantages	• Short path-length puts the entire display near the eyes/face • Short path-length means less packaging design freedom	• More complicated optical design • More lenses mean heavier design • Loss of imagery outside of pupil • Needs precision fitting, more and finer adjustments

Exit pupil and eye relief

In all cases, the optical design must provide a sufficiently large exit pupil or viewing eye box.[24] This is the area shown in Figure 17-5 for the non-pupil-forming and Figure 17-6 for the pupil-forming system. A large exit pupil is important for a flight HMD, so the user doesn't loose the image if the HMD shifts on his head. A value of 12 to 15 mm has been deemed an acceptable value for these applications.

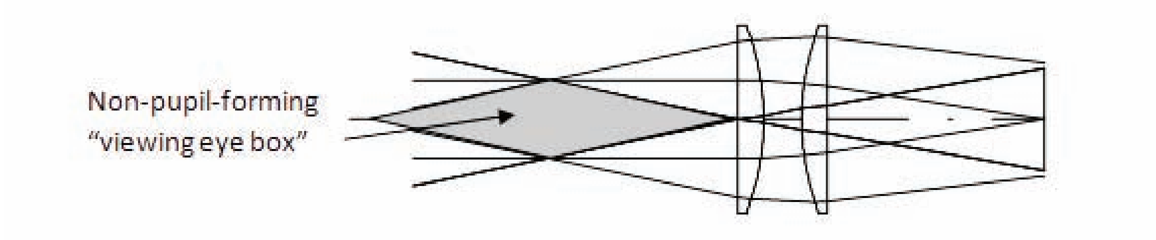


Figure 17-5. The viewing eyebox within which there will be unvignetted viewing of the HMD image source (shown in gray). Outside of that area, the image will vignette or be clipped, but still visible.

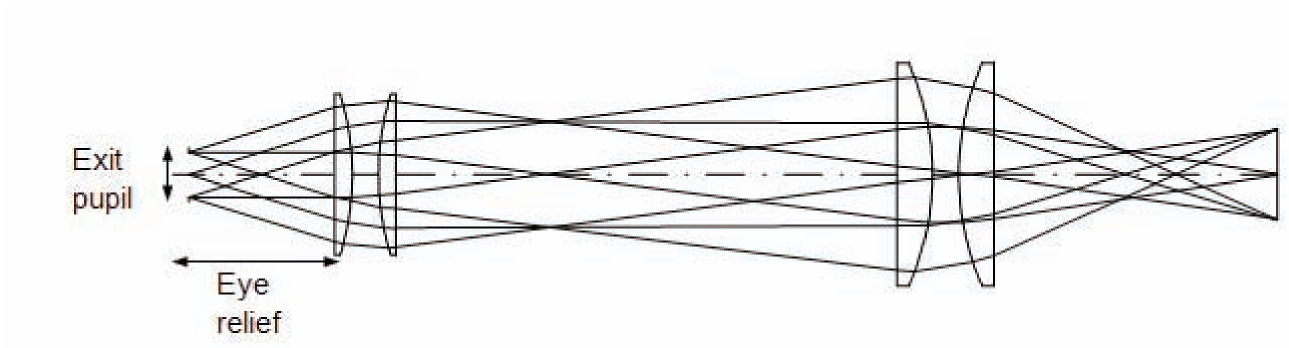


Figure 17-6. The exit pupil and eye relief of a pupil-forming optical design. Note that outside of the pupil area there is no imagery.

In Figure 17-7, it can be seen that the size of the off-axis exit pupil plays a disproportionately large role in determining the size and weight of the optics. The off-axis exit pupil is important for a partially overlapped HMD, where the on-axis ray from the image source actually traverses the off-axis portion of the optics. It is also important so the user does not loose the image when rotating their eyes to view imagery on the edge of the FOV, though most eye movements tend to be less than ±10° to ±15° (Bahill, Adler and Stark, 1975). Depending on the application, it is possible to trim the off-axis exit pupil so it is only 50% of the on-axis exit pupil diameter, reducing the size and weight, but not significantly reducing performance. By trimming the size of the off-axis exit pupil, we can reduce the size of the optics.

The HMD needs sufficient eye relief to allow the user to wear spectacles,[25] with a generally accepted minimum value of 25 mm. However, care must be taken with this terminology, because in classical optical design, the eye relief is measured as the distance along the optical axis from the last optical surface to the actual exit pupil. In most HMDs, the final optical surface in front of the eye may be an angled combiner which will fold the optical

---

[24] The exit pupil is found only in pupil-forming designs. In non-pupil-forming designs, it is more correct to refer to a *viewing eyebox*, because there is a finite unvignetted viewing area.

[25] Approximately one-third of U.S. Army aviators are required to wear vision correction, which increases as the population of qualified pilots ages. Though spectacles are the typical choice for visual correction, the U.S. Army has also investigated the use of contact lenses as well as surgical correction methods (Rash et al., 2002).

path to get the rest of the optics away from the front of the face, so the actual *eye clearance distance* (ECD) (measured from the face to the closest point of the combiner) may be considerably less. Thus, it is important that the useable distance from the eye to the first contact point of the HMD optics – the ECD – provide the minimum 25 mm.

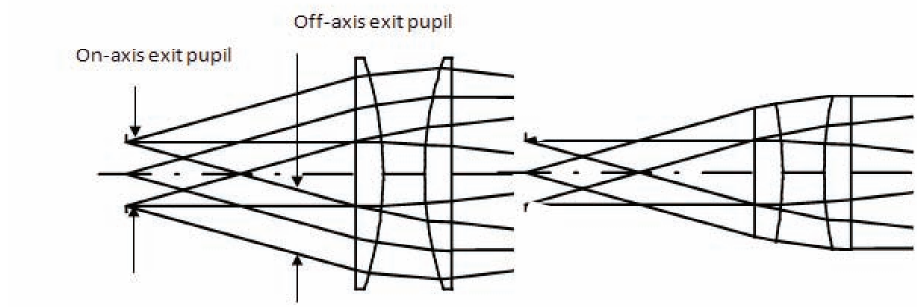


Figure 17-7. A comparison of the size of the collimating lens with full off-axis exit pupil (left) and the off axis exit pupil trimmed to 50% vignetting (right). Doing so significantly reduces the size and weight of the lens assembly.

Optical distortion

One of the important issues in an optical design is the control of residual optical aberrations such as focus, field curvature, and astigmatism. While this can be done with careful attention to the optical design, adding perhaps an additional lens or aspheric surface, distortion (defined as an off-axis image located at a different height than that expressed by paraxial equation) is more difficult to control, usually taking on a pincushion form in an imaging system. HMDs with off-axis optical designs (such as the JHMCS and JSF RSV) can have more complex asymmetric distortion. The result of residual distortion is:

- Non-linear motion across the FOV – an object moving across the visual field will appear to move at a different velocity at the edges of the field rather than in the center of the field.
- Non-linearity of horizontal and vertical lines – a horizon line that is supposed to be flat will be curved at the edges of the FOV.
- Binocular images don’t line up – this is especially important in a partially overlapped binocular system where the edges of the field are in the center of the binocular field (Figure 17-8).

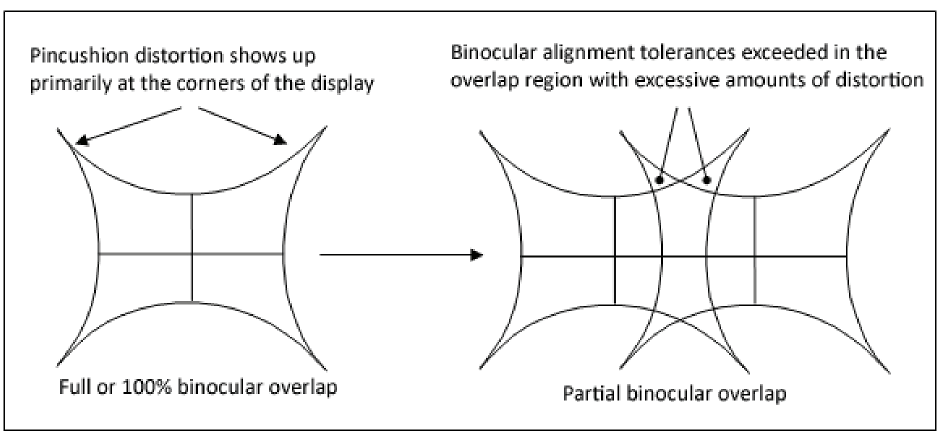


Figure 17-8. The effect of residual pincushion distortion on the binocular alignment in the overlap region. If the residual distortion is not properly controlled, it can induce eyestrain in the user.

If the HMD will be viewing imagery from a sensor, which generally has the distortion of the opposite sign, or "barrel," it is possible that the pincushion distortion on the HMD may compensate, though not completely. In CRT-based HUD design, this can be corrected by pre-distorting the image plane. With a pixilated or finite-addressable display such as an LCD, however, the pixels cannot be moved, though it is still possible to *pre-distort* the imagery. Watson and Hodges (1995) reported a method of applying a geometric pre-distortion in the *texture* memory on high end image generators prior to final image rendering. Similarly, image warping engines are now available that accept imagery from a sensor, apply a polynomial correction to the imagery, and pass it on to the image source.

See through versus non-see-through HMDs

The decision to use a *see-through* or *non-see-through* HMD depends on the particular application, environment and the imagery desired for viewing. As with almost all HMD requirements, there are several key tradeoffs that must be made.

A see-through design is desired for aviation applications. Completely occluding one or both eyes is generally not acceptable. In particular, a see-through HMD design allows the superposition of imagery over the outside world, sometimes referred to as *augmented reality* (Azuma, 1997). As discussed in Chapter 19, *The Potential of an Interactive HMD*, see-through HMD imagery can be displayed in three frames of reference (Procter, 1999; Yeh, Wickens and Seagull, 1998): Aircraft-, earth- and screen-referenced.[26] With the HMD, navigational guidance and targeting data, as well as head-tracked sensor imagery, can be displayed. This allows a Warfighter to remain in contact with the real world and have the information aid in accomplishing the mission.

While the see-through design provides distinct advantages for an aviation application, it is a more difficult optical design because the see-through combiner must be large enough to provide sufficient FOV, exit pupil and eye relief without excess weight or adversely impacting pilot safety. Examples of see-through designs that use a separate optical combiner are shown in Figure 3-22 of Chapter 3, *Introduction to Helmet-Mounted Displays*; these include the IHADSS, HIDSS, Knighthelm, TopOwl®, VCOP and Q-Sight HMDs. For many fighter aircraft applications, the protective visor also serves as the HMD combiner, such as in the DASH-3, JHMCS, and JSF RSV shown in Figure 3-20. In this case, it is necessary to stabilize the visor to ensure that it can still maintain the proper focus and binocular alignment tolerances.

Most aviation applications use only monochromatic imagery, typically centered at 555 nanometers (nm), because this is the peak daylight (photopic) visual sensitivity (Boff and Lincoln, 1988). One of the ways to improve both see-through transmission and reflectance is to take advantage of high reflectance holographic notch filters and V-coats.[27] The drawback is that while these special coatings reflect more of a specific display color, they transmit less of that *same* color, which makes the world look pink. With the advent of more use of color in the cockpit, selectively reflecting green over another color may miscue the pilot. For these reasons, many aircraft HMD combiners have spectrally neutral reflective coatings.

---

[26] *Aircraft-referenced* – An example would be in the RAH-66 Comanche helicopter program where a wire-grid frame was drawn to represent the front of the aircraft, giving the pilot an intuitive understanding of the direction of "aircraft-forward," regardless of where his head was pointing. *Earth-referenced* – Here the pilot sees either real objects such as runways or horizon lines or virtual objects such as safe pathway in the sky, threat/friendly aircraft locations, engagement areas, waypoints, and adverse weather. Similarly, the pilot can be provided with head-tracked imagery as in the case of the AH-64 Apache helicopter or the JSF aircraft. *Screen-referenced* – This is information that does not require any reference to be seen, such as altitude, airspeed, or fuel status, similar to how a HUD displays this information. Also consider the case of a dismounted Warfighter viewing moving map information or text information. In this latter case, it may not even be necessary to have a see-through design. The first two require accurate head-tracking and image registration.

[27] V-coating refers to an antireflective boundary layer coating technique designed to reduce reflections at a single wavelength.

Spectral security is not generally as important in an aviation application because most aviation cockpit designs are governed by MIL-STD-3009, Department of Defense Standard for Lighting, Aircraft, Night Vision Imaging System (NVIS) Compatible (Department of Defense, 2001) which specifies different configurations for fixed- and rotary-wing applications.[28] The minus-blue filter on the NVGs filters out light below the 625 nm or 665 nm regions, primarily the red and orange spectrum. NVGs used by dismounted Warfighters, however, do not have this filter. Since the I^2 sensors are sensitive into the green spectral region, it means a dismounted Warfighter may give away his position at night if he is viewing imagery on a see-through HMD. In this case, a non-see-through HMD with an eyecup may be preferable. Without the need for a see-through combiner or the requirement for high luminance against a bright background, the end-to-end transmission efficiency is improved reducing the power for the image source.

Luminance and contrast

With the exception of flying with NVGs at night, every aviation task (both fixed- and rotary-wing) requires a see-through HMD to direct imagery to the user's eyes in much the same way that aircraft HUDs present imagery that is superimposed on the outside world. But the ability to see imagery in the high ambient luminance environment of an aircraft cockpit is counterbalanced by the need for high see-through transmission combiners on the HMD. To view the imagery against a bright background such as sun-lit clouds or snow, this less-than-perfect reflection efficiency means that the image source must be that much brighter. The challenge is to provide a combiner with good see-through transmission and still provide an image with sufficiently high contrast against the high luminance background. Figure 17-9 below, shows a diagram of a simple HMD optical design (see also Chapter 4, *Visual Helmet-Mounted Displays*). There are limitations, though, because most image sources have a luminance maximum governed by the physics of the device as well as by size, weight and power of any ancillary illumination. Other factors such as the transmission of the aircraft canopy and pilot's visor must be considered when determining the required image source luminance as shown.

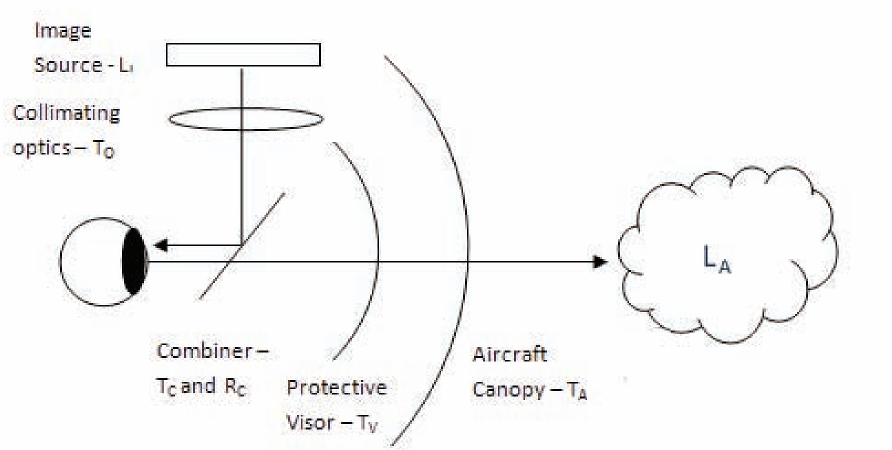


Figure 17-9. The contributions for determining image source luminance requirements for an HMD in an aircraft cockpit.

---

[28] MIL-STD-3009 (which replaced MIL-L-85762A) Lighting, Aircraft, Interior, Night Vision Imaging System (NVIS) Compatible, was written to guide aircraft cockpit designers (Breitmeyer and Reetz, 1985) in cockpit design that is compatible with night vision goggles by 1) limiting the spectral output of all interior display and cockpit lighting and 2) filtering the spectral response of the pilot's NVG so that when operational, no lighting would affect the gain of the NVGs. Helicopters with ANVIS goggles are Type I (direct view), Class A (625 nm minus blue filter). Fixed wing aircraft with a Cats-Eyes NVG are Type II (projected view), Class B (665 nm minus blue filter).

For the HMD design in Figure 17-9, the HMD luminance is given by:

$$L_{HMD} = L_I * T_O * R_C \qquad \text{Equation 17-2}$$

where L_I is the image source luminance, T_O is the transmission of the collimating optics and R_C is the reflectance of the optical combiner. The pilot views the outside world through the combiner (transmission of T_C or $1 - R_C$[29]), the protective visor (T_V, either Class 1 or Class 2[30]), and the aircraft canopy transparency (T_A) against the background luminance (L_A). Thus, the background luminance observed by the pilot, L_o is given by:

$$L_O = L_A * T_A * T_V * T_C \qquad \text{Equation 17-3}$$

For a see-through configuration, CR is given by the expression:

$$CR = \frac{(L_O + L_{HMD})}{L_O} \qquad \text{Equation 17-4}$$

Combining Equations 17-2, 17-3 and 17-4, CR can be expressed as:

$$CR = 1 + \frac{(L_I * T_O * R_C)}{(T_C * T_V * T_A * L_A)} \qquad \text{Equation 17-5}$$

For a nominal CR value of 1.2 against a worst case 10,000-foot-Lamberts (fL) background ambient (Foote, 1998), substituting values for the additional factors produces the required image source luminance values (L_I) presented in Table 17-6.

Table 17-6.
Required image source luminance is shown for four different HMD configurations.

		Case 1 Clear visor, 50% combiner transmission	Case 2 Dark visor, 50% combiner transmission	Case 3 Clear visor, 80% combiner transmission	Case 4 Dark visor, 80% combiner transmission
Collimating optics transmission	T_O	85%	85%	85%	85%
Combiner reflectance	R_C	50%	50%	20%	20%
Combiner transmission	T_C	50%	50%	80%	80%
Visor transmission	T_V	85%	15%	85%	15%
Aircraft canopy transmission	T_A	80%	80%	80%	80%
Ambient background luminance	L_A	10,000 fL	10,000 fL	10,000 fL	10,000 fL
Required image source luminance	L_I	1600 fL	282 fL	6400 fL	1129 fL

[29] To simplify our calculations, we will ignore Fresnel losses at each of the optical surfaces and assume that $T_C + R_C = 1$ although the small losses must be included in any rigorous analysis.
[30] Aviator visor configurations are given in MIL-V-43511. The Class 1 visor transmission is specified as ≥85% and the Class 2 visor transmission is specified as between 12% and 18%. For our calculations the latter is assumed to be 15% transmission.

In the first two cases, we can see the impact of wearing a Class 1 (clear) versus a Class 2 (dark) visor on the required image source luminance. The dark visor reduces the ambient background luminance, improving HMD image contrast against the bright clouds or snow and reducing the required image source luminance. These first two cases are relatively simple because they assume a combiner with 50% transmission and 50% reflectance. Most pilots want higher see-through transmission because it allows them to see low-luminance targets at longer distances. This dictates a reduced combiner reflectance, demanding higher image source luminance. Cases 3 and 4 in Table 17-6 assume this higher transmission with both clear and dark visors, demonstrating how this requires higher image source luminance, which in-turn requires more power. While more power may not be an issue for the aircraft pilot, for the dismounted Warfighter – who carries his own batteries - HMDs power consumption is critical.

Helmet-mounted imaging sensors

The ANVIS and AN/PVS-14 NVGs have been implemented successfully to augment the Warfighter's vision in low light conditions. More recently long-wave thermal imaging has been added for dismounted Warfighters, under a program called Enhanced Night Vision Goggle (ENVG) and designated the AN/PSQ-20. These sensor systems typically mount to the front of the Warfighter's helmet, with the entrance aperture in line with the user's eye. While aiding the Warfighter in low light or night time conditions, these sensors have an adverse effect on head-supported weight and CM because the components protrude out in front of the user's face. A better approach would be to integrate the sensor hardware into the helmet so as to minimize bulk and protrusions and to better optimize weight and balance. Unfortunately, this can create an offset of the sensor aperture with respect to the wearer's normal line-of-sight.

Melzer and Moffitt (2007) investigated the perceptual and performance effects of viewing offset (forward, high and centered and side) monocular sensor video, replicating potential integrated design solutions for dismounted Warfighter applications. The results indicated little or no eye dominance issues but demonstrated that the sensor and display must be aligned to within 0.5°. If the alignment error was in the horizontal plane, subjects walked in an arc, rather than in a straight line.

Melzer and Moffitt (2007) also found that when aligned, the high-mounted sensor gave an indication of a slated floor, and the side mounted sensor produced a blind pointing error that was opposite of the sensor location. However, as long as the test subjects were able to view their feet and hands, they were able *re-calibrate* their hand-eye coordination to perform close-in tasks, albeit with some temporary after effects (Bertelson and de Gelder, 2004).

The offset sensors also may have implications in the cockpit, because in designing an aircraft cockpit, the starting point is the design-eye location[31], the assumed origin from which the pilot will view out the windows and all cockpit displays. When helmet-mounted sensors are spaced further out than the assumed 2.5-inch nominal spacing, these assumptions may no longer be valid as the sensors may be staring directly into a canopy bow or strut. The result is that when looking at a see-through HMD, pilots may see one pair of struts with their normal vision and a second set of struts with their sensors.

The most common approach is to place night vision sensors on either side of the helmet, creating a perceptual condition referred to as hyperstereopsis. While this has been purposefully implemented to exaggerate stereo depth cues for enhancing detection of terrain drop off (Mohananchettiar et al., 2007), displacement of sensors in an aviator's HMD presents additional perceptual issues. The Thales TopOwl® has sensors located on either side of the helmet with a separation of approximately 10 inches (Priot et al., 2006).

Humans perceive depth visually several ways, based on both monocular cues (optical flow and optical expansion) and binocular (stereopsis). One of the key perceptual conflicts created in hyperstereopsis is the

---

[31] In the design of human-machine interfaces (HMIs) (to include HMDs), the *design-eye position* is the position from which the user's eye is expected to view.

exaggeration of depth due to disparity magnification, which results in objects appearing closer than they are (object size is not affected, as there is no effect in the vertical direction). Helicopter pilots report feeling as though they have already landed, though they may still be a few feet off the ground. In addition, there are anecdotal reports of fixed-wing pilots “landing long” because they thought they were already on the ground. Kalich et al. (2007) describe the illusion as being in a “mountain top crater.”[32] Flanagan, Stuart and Gibbs (2007a) found no effect on *absolute* distance perception, though they do state that this is unrelated to the impact of hyperstereopsis on *relative* distance perception. In another study, these same researchers (Flanagan, Stuart and Gibbs, 2007b) found that hyperstereopsis affected a pilot’s estimate of “time to contact”, making the approach seem faster than what would be expected with normal vision. In a third study, Stuart, Flanagan and Gibbs (2007), report that slope estimation is also affected. These researchers point out that although the effects were noted, the magnitude was not as much as would be expected from the 4X interocular separation. They speculate that even though the monocular and binocular cues are in conflict, the monocular cues are perceptually weighted more heavily than the incorrect binocular cues. They further speculate that because they found strong individual differences between subjects, it may indicate differences in the levels of suppression of these binocular cues and that this may be from individual adaptation strategies the subjects were using to overcome the perceptual errors. Given the reports of the pilots who have qualified with the TopOwl® (Mace, Van Zyl and Cross, 2001; Priot et al., 2006), it is tempting to speculate that subjects have the ability to adapt to the new perceptual condition. Bertelson and de Gelder (2004), use the term “re-calibrate” to describe such a condition where subjects adjust their hand-eye coordination to a new-found perceptual construct. Mohler et al. (2007) found that visual perception of the speed of self-movement will cause subjects to re-calibrate their visually directed tasks. There are, as yet, unanswered questions of how effective this re-calibration is and whether there are any residual aftereffects as pilots switch between the hyperstereo and normal vision.

Visual vs. Auditory Mode for HMDs

The visual channel is the mode of choice for providing information at high rates. However, for certain tasks and situations an auditory display may be more effective. Auditory displays are best used for alerting, warnings, and alarms situations in which the information occurs randomly and requires immediate attention. The near omnidirectional character of auditory displays is a major advantage over other types of HMD modes.

Table 17-7 (Deatherage, 1972, cited in National Research Council, 1997) summarizes some key factors when making a choice between an auditory and a visual display. Care must be taken that the auditory signal be consistent in its message or meaning (i.e., represent the same information in all situations) and not be too intense as to produce a startle response. Effective auditory displays should employ frequencies different from environmental background noise (to avoid masking); for choice situations, use moderate intensity, easily discernible frequency or amplitude signals; and use separate auditory warnings, different from other auditory signals (Hedge, 2007). Further discussion on the advantages of auditory cueing can be found in Chapter 14, *Auditory-Visual Interactions*, and Chapter 19, *The Potential of an Interactive HMD*.

Glumm et al. (1999) conducted a field study to investigate the effects of an auditory versus a visual presentation of position information on Warfighter performance of land navigation and target acquisition tasks. Additional measures of situational awareness, stress, cognitive performance, and workload were obtained. In the auditory mode, position information was presented in verbal messages. In the visual mode, the same information was provided in text and graphic form on a map of the area of operation presented on an HMD. During the study, 12 military volunteers navigated densely wooded unmarked paths that were 3 kilometers (1.9 miles) long.

[32] “The observer describes the ground nearest to him as appearing closer (higher), with this exaggerated depth effect (the closer than effect) decreasing with distance away from the observer. When the helicopter is on the ground, the pilot perceives the near ground as being at chest level, while distant objects may look natural, a result of the non-linearity of the exaggerated depth perception with increasing distance from the observer.” (Kalich et al., 2007, page 3).

Although no differences were found between the two display modes in the frequency at which navigational and other tactical information was accessed, study participants reported that they maintained a greater awareness of position with respect to waypoints, targets, and other units when information was presented visually than when information was presented in an auditory mode via verbal messages. Although visual presentation of information appeared to enhance position awareness, differences between the two display modes in navigation and target acquisition performance were not found to be statistically significant. The findings of the investigation suggest differences in cognitive processing requirements between the two display modes and the impact of attentional focus and practice on cognitive performance.

Table 17-7.
Auditory vs. visual form of presentation.
(Deathridge, 1972; National Research Council, 1997)

Use auditory presentation if:	Use visual presentation if:
The message is simple (e.g., fire alarm).	The message is complex.
The message is short.	The message is long.
The message will not be referred to later (e.g., ambulance siren).	The message will be referred to later.
The message deals with events in time (e.g., telephone ring).	The message deals with location in space.
The message calls for immediate action (i.e., engine fire warning).	The message does not call for immediate action.
The visual system of the person is overburdened.	The auditory system of the person is overburdened.
The visual system is unavailable (e.g., receiving location is too bright or too dark; individual is asleep).	The receiving location is too noisy.
The person's tasks require him or her to move about continually.	The person's job allows him or her to remain in one position.
The origin of the signal itself is a sound (e.g., automobile horn).	

Acoustic/Auditory Guidelines and Recommendations

Human senses form a reception suite that provides orientation and security to human beings in a variety of conditions. *Vision* allows precise discerning of self-contained patterns and motion from a background covered by a FOV. *Smell* informs about a global invisible change in the environment that may affect human well-being and may crudely guide a person toward or away from the source. *Taste* confirms that what looks like a specific substance is really that substance and warns against the wrong kind of nutrient. *Touch* provides another type of feedback for our actions and often provides a last line of defense against immediate danger. *Balance* allows us to move and understand our relationship with the local environment. Finally, *audition* allows us to localize activities in a full 360° spherical angle that vastly extends our global FOV. While both the senses of smell and audition are comprised of telereceptors that inform humans about changes in global environment, only audition provides specific directional information needed to identify the location of a specific activity. Audition also allows hearing species to “see through visual obstacles” when some activity takes place within the FOV, but is invisible due to other objects obscuring its visibility or due to the opaqueness of the environment (e.g., nighttime, smoke, smog, mist, and fog). Auditory information can assist in guiding the more precise vision system toward the objects of interest and extends vision beyond the immediate foreground.

As all other senses, audition has its own strengths but also its own limits. The role of human factor engineers and interface designers is to maximize the use of all the senses in the multiple operational environments. The real challenge is to enhance some sensory capabilities without reducing or making other the capabilities useless such that Warfighter safety or overall effectiveness is compromised.

The operational acoustic environment of the Warfighter is likely to be extremely varied. At one extreme, the surrounding environmental noise may be so intense as to preclude normal voice communication and reception of acoustic or audio signals. Such noise can be a serious source of stress as well as interference to both direct and radio communication. At the other extreme is the need for surreptitious activity in a quiet environment where any audible sound generated by the Warfighter or his equipment is to be avoided for security reasons. Either of these ambient conditions can restrict the utility of auditory communication and audio HMDs if they are not properly designed. Yet, in both situations audition may be the only remaining critical capability of the Warfighter. Many common military scenarios are very dependable on availability of auditory information, e.g., a Warfighter in hiding in direct visual proximity of the enemy forces, a combat engineer in a protective suit in chemically challenging environment disassembling an explosive trap, a squad entering a building, a fixed-wing aviator communicating during a 5G turn. Interviews with many Warfighters repeatedly confirmed that in limited visibility environment, they do not rely on the remaining visual information but react to what they hear. Unlike visual information, auditory information comes from all directions and over and "through" visual obstacles. Sound is often the first contact the Warfighter has with the enemy (Monroe, 2004).

As discussed in Chapter 5, *Audio Helmet-Mounted Displays*, the main audition-related challenge of Warfighters and underlying HMD-system developers and tacticians is to achieve an optimal balance between the Warfighter's auditory awareness of the environment and uninterrupted, secure, low-level, and clear audio interconnectivity and to balance this against sufficient amount of hearing protection needed to for the ensure the Warfighter's current and long-term hearing capability. Fortunately, there are some capabilities and technologies that can be combined to meet this challenge. As one example, bone conduction two-way interface permits relatively clear communication even in high level of surrounding noise. Bone conduction interfaces can be very inconspicuous, allowing them to be hidden from visual observation. They also are very sensitive to vocal tract changes and can transmit very low level acoustic signals as voiceless whisper, teeth clicking (e.g., Morse code), or other low level coded vocal signals. Recent research reports revealed that even changes in neural activity resulting in muscle changes during voiceless speech articulation can be used as audio input signal (Simonite, 2007; 2008).

Auditory awareness of the environment and hearing protection during sudden burst of noise, such as own and enemy fire power, explosions, sudden impact sounds, etc., can be optimized by the use of adaptive nonlinear hearing protection, as described in Chapter 5, *Audio Helmet-Mounted Displays*. These devices even may incorporate amplification of environmental sounds when needed for sound detection. However, fixed and permanent amplification of acoustic environment – either directional or not – has to be highly discouraged due to its detrimental effects of auditory distance estimation, general spatial orientation, and loss of hearing sensitivity to the sounds coming from the rear.

As technology and warfare becomes more sophisticated, Warfighter sensory and cognitive workload is steadily increasing, making proper utilization and specialization of sensory inputs critical. As for audition, its main function is to allow the Warfighter to understand the dynamically changing environment and to localize quickly – and this means detect early and localize relatively precisely – all the activities in the surrounding space, regardless of their direction. No other sense can substitute hearing in this capacity. Auditory localization precision degraded to 20 to 30° or an even larger angle can still be sufficient for navigation through a safe environment when the time of arrival is not a factor, but it could lead to increased casualties and a substantial drop in mission effectiveness and is especially detrimental to the operational conditions of the dismounted Warfighters. Sound signature recognition and identification are also important, but they can be supported by other senses, and thus being still important they are less critical in audio HMD design.

Meeting the challenges of the diverse and even sometimes contradictory requirements of an effective audio HMD demands good detectability and localizability of environmental events and should be regarded as the highest priority for audio HMD design. This is followed by effective speech communication and protection against hazardous high level noise. Providing these capabilities requires a variety of acoustic and non-acoustic means to be considered and implemented whenever possible. A concise list of the main of these means includes:

- Exposure of both pinnae to environmental sounds or the use of a truly acoustically transparent headgear covering the ears (although recognized as difficult to achieve)
- An acoustically-optimized shape of the headgear (e.g., shell, headband) to minimize dispersion and shadowing of natural sounds
- Level-dependent in-the ear hearing protection to be used when hearing protection is needed
- No fixed and permanent environmental sound amplification; such devices, if incorporated, should be only turned on occasionally and should provide adjustable directivity enhancement
- Inconspicuous, always-on, audio communication system based on bone conduction or whisper-quality audio interface with an easy to access step-wise volume control
- Secure, fixed, but low pressure contact between the audio transducers and the user's head
- Speech-optimized audio transducers assessed for speech intelligibility using a variety of talkers and speech modes (see Chapter 11, *Auditory Perception and Cognitive Performance*)
- Provisions to use biological and chemical protection gear without detrimental effects on audio communication and noise protection
- Optional inconspicuous always-on environmental microphone to send continuous audio stream to commander or other receiving authority without Warfighter's intervention
- Optional tactile-based sniper detection and master warning interface (see Chapter 18, *Exploring the Tactile Modality for HMDs*)

Biodynamics Guidelines and Recommendations

The primary role of the Warfighter's helmet always has been to provide protection. This role has not changed and instead has been expanded with the introduction of HMDs to where the helmet is expected to serve as a mounting platform for the display without compromising the helmet's primary protective capability. These increasing demands means we must consider impact attenuation, head-supported weight, CM offset, frangibility, fit and comfort, retention and stability and their effects on head and neck biodynamics. Since these have not been addressed elsewhere in this volume, we will explore it in more detail in this section.

The human head weighs approximately 9 to 10 pounds (mass of 4 to 4.5 kg) and sits atop the spinal column. The occipital condyles at the base of the skull mate to the superior articular facets of C1, the first cervical vertebra (Perry and Buhrman, 1997; Melzer, 2006).[33] These two small, oblong mating surfaces on either side of the spinal column are the pivot points for the head. Their approximate location in the X-Z plane may be found by palpating the mastoid process (the pointed, bony structure behind the base of the ear). The CM of the head is located at or about the tragion notch, the small cartilaginous flap in front of the ear. Because this is *up* and *forward* of the head/vertebra pivot point, there is a tendency for the head to tip downwards, were it not for the strong counter force exerted by the neck extensor muscles – hence when individuals fall asleep, they "nod off."

While the mass of a HMD system – and from this point forward, we are talking about the protective helmet and any head or helmet-worn protective gear, plus the display – is distributed over the surface of the wearer's head, a specific location can be defined where the HMD mass can be assumed to be concentrated, which we refer to as the HMD's CM and is expressed relative to a pre-defined coordinate system. For the U.S. Army, CM locations are defined with respect to the human head anatomical coordinate system (Figure 17-10) (Deavers and McEntire, 1993; Rash et al., 1996).

---

[33] There are seven cervical vertebrae. These are designated starting with C1 (the Atlas), C2 (the Axis), through C7, the bottom of which mates to the top of T1, the first thoracic vertebra.

The U.S. Navy and Air Force define CM locations relative to structural and anatomical reference points on the head of a crash test manikin (Albery and Kaleps, 1997; Thornton and Zaborowski, 1992). Adding mass to the head in the form of an HMD or NVGs can move the CM (now the helmet assembly + head) away from the ideal location. High vibration or buffeting, or dynamic events like ejection, parachute opening or crash can result in high accelerations of very short durations which will exacerbate the effect of this extra weight and displaced CM.

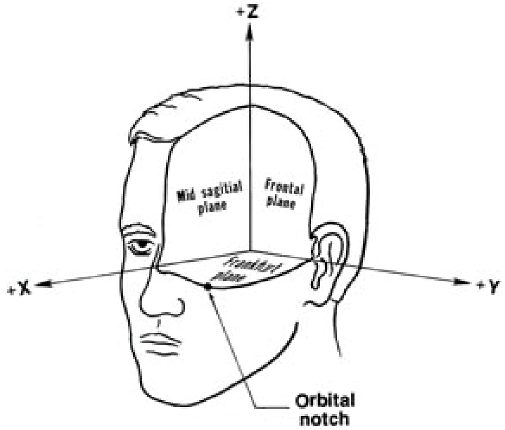


Figure 17-10 The anatomical coordinate system from which the head-supported weight/mass and center-of-mass (CM) requirements are calculated.

Key considerations of head-supported weight and CM must be consistent with levels of human tolerance so that during the course of a mission the wearer is not required to endure mission-compromising levels of head and neck fatigue or to require time afterwards to recover from induced neck pain. Effects can range from fatigue and neck strain to serious or mortal injury (Guill and Herd, 1989), as well as possible long-term cervical and spinal degradation.[34]

An important difference between the aviator and the dismounted Warfighter is the measure of "success" after a dynamic event. A fixed-wing ejection or rotary-wing crash can be considered successful if the pilot can "walk away" even though he may not be considered combat effective immediately. On the other hand, the dismounted Warfighter who parachutes to the ground and is not immediately combat effective cannot be considered to have jumped successfully. This means we must consider different margins of safety between injury and non-injury depending on the Warfighter's task.

Another concern is what effects specific operational factors (e.g., the Warfighter's jump-induced stress or the pilot's cumulative vibration or the pulling of Gs) in combination with the added stress of wearing an HMD may have on long-term degradation of the spine, e.g., causing chronic inflammation and pain. There has been investigation into the long-term effects of extended G-loading on spinal degradation for fixed-wing pilots (Burton,

[34] The human spine was not designed to support upright, bipedal posture. Rather, it was intended as a suspension bridge between the front and rear legs of quadrupeds. It has been shown there is an age-related loss in water content in the spinal disk, which nominally provides intervertebral cushioning. Cadaveric investigations have confirmed that degeneration starts in the late 20's or early 30's, and by age 40, almost all disks will have some form of degeneration - mostly asymptomatic - typically in C4/C5 and C6/C7, although primarily C5/C6 (Burton, 1999a, Burton, 1999b, Burton, 1999c). This degeneration is exacerbated under acceleration, shock and added g-forces (Harms-Ringdahl et al., 1999). Under constant loading, there appears to be long-term mechanical creep of the intra-spinal disk material resulting in reduced intervertebral spacing. Frequent insult can result in osteophytic growth, which further restricts mobility, resulting in an often painful constriction of the nerve root exiting the neuroforamen. This has been found in fighter pilots, rugby players, gymnasts, wrestlers and with automotive injuries (Burton, 1999b). Interestingly, studies of individuals who carry upwards of 200 lb on their heads many times per day show little degeneration, most likely because they maintain an upright, neutral position during the process, minimizing insults to the facet joints and intervertebral disks.

1999a; Burton, 1999b; Hämäläinen and Kuronen, 1999; Harms-Ringdahl, Linder et al., 1999) and some for rotary-wing pilots (Butler and Alem, 1997), though nothing has been done to specifically address the long-term effects of repeated parachute opening shock, parachute landing fall or of long-term helmet wear on dismounted Warfighters. However, given the evidence of long-term effects on pilots and other non-aviation activities, it is difficult *not* to speculate a causal relationship between a Warfighter's activities and some level of long-term degeneration.

Because the environments are different for rotary-wing (long duration missions, high vibration levels), fixed-wing (shorter duration missions, but higher G-loading) and ground combat, fatigue-minimizing design measures, though similar, differ in their specific values. Historically, head-supported weight and CM requirements have been nonexistent or vague. These requirements often were written loosely and based on existing designs. Language in helmet development specifications often resembled statements as "…the helmet CM must be located as close to the head CM as possible," "…lighter and CM no worse than current helmet systems," " provide ease of head movement," and "…(have) reduced bulkiness." These requirements provided little detailed guidance to the design teams and could not be quantitatively evaluated.

During the 1990s, the U.S. military attempted to better define head-supported weight and CM requirements for head-borne devices. In 1991, the U.S. Air Force published interim head-supported weight and CM criteria for its fixed-wing helmets (the "Knox Box" and the "Tolland Box" see Perry and Buhrman, 1997; Knox et al., 1991; MacMillan, Brown and Wiley, 1995; Settecerri, McKenzie, Privitzer and Beecher, 1986); these criteria were developed to keep neck compression loads at an acceptable level during ejection.

In 1998, the USAARL published a set of two curves that defined limits of acceptable longitudinal and vertical CM location as a function of head-supported weight, commonly referred to as the "USAARL curves" (see Figure 17-11, Ashrafiuon, Alem and McEntire, 1997; Barazanji and Alem, 2000; Harding, et al., 1998; McEntire, 1998c; McEntire and Shanahan, 1998). These were developed to provide HMD designers with guidance that would minimize performance degradation during typical helicopter flight scenarios (Alem and Meyer, 1995; Butler, 1992), as well as minimize the risk of acute neck injury during severe, but survivable, helicopter mishaps (McEntire and Shanahan, 1998). Since their publication, the USAARL curves have become the de facto standard for the design of rotary-wing aviation helmet-HMD systems.

Studies conducted by all three U.S. military services since the USAARL curves were published have investigated the effects of head-supported weight and CM location on the risk of neck injury during dynamic events (Bass et al., 2006; Brolin et al., 2008; Doczy, Mosher and Burhman, 2004; Halldin et al., 2005; Merkle, Kleinberger and Uy, 2005). Additionally, several of these studies have shown crash severity and head-supported weight to have the greatest influence on the risk of neck injury (Brolin et al., 2008; Halldin et al., 2005; Paskoff, 2004).

The USAARL and the U.S. Air Force Research Laboratory (AFRL) also have conducted studies investigating the effects of variables such as head-supported weight, CM position, and gender on wearer fatigue and performance (Barazanji and Alem, 2000; Eveland et al., 2008; Eveland and Goodyear, 2001; Fraser et al., 2006). Barazanji and Alem (2000) determined that the biomechanical response of female subjects wearing varying head-supported weight while subjected to simulated helicopter environments was similar to that of males (Butler, 1992) and recommended there should be no gender-specific head-supported weight and CM criteria. Fraser, Alem and Chancey (2006) found that increased head-supported weight and anterior CM position had significant adverse effects on Warfighter performance in visual tracking tasks. Conversely, AFRL research has shown that head-supported weight and CM location did not have a significant effect on performance in tracking tasks during exposure to sustained acceleration (e.g., as experienced during air combat maneuvering) (Eveland and Goodyear, 2001).

Biomechanics research at the U.S. Air Force's Wright-Patterson Laboratories has established head-supported mass and CM boundaries for fixed-wing HMDs, essentially a refinement of the "Knox Box."[35] Using the occipital

---

[35] Plotting the CM spatial limits on an X-Z plane yields a rectangle commonly known as the *Knox Box* (Knox et al., 1991).

condyles (the pivot of the head about the C1 cervical vertebra – located approximately between the two points of the mastoid process just behind the ears) as the origin point, these are:

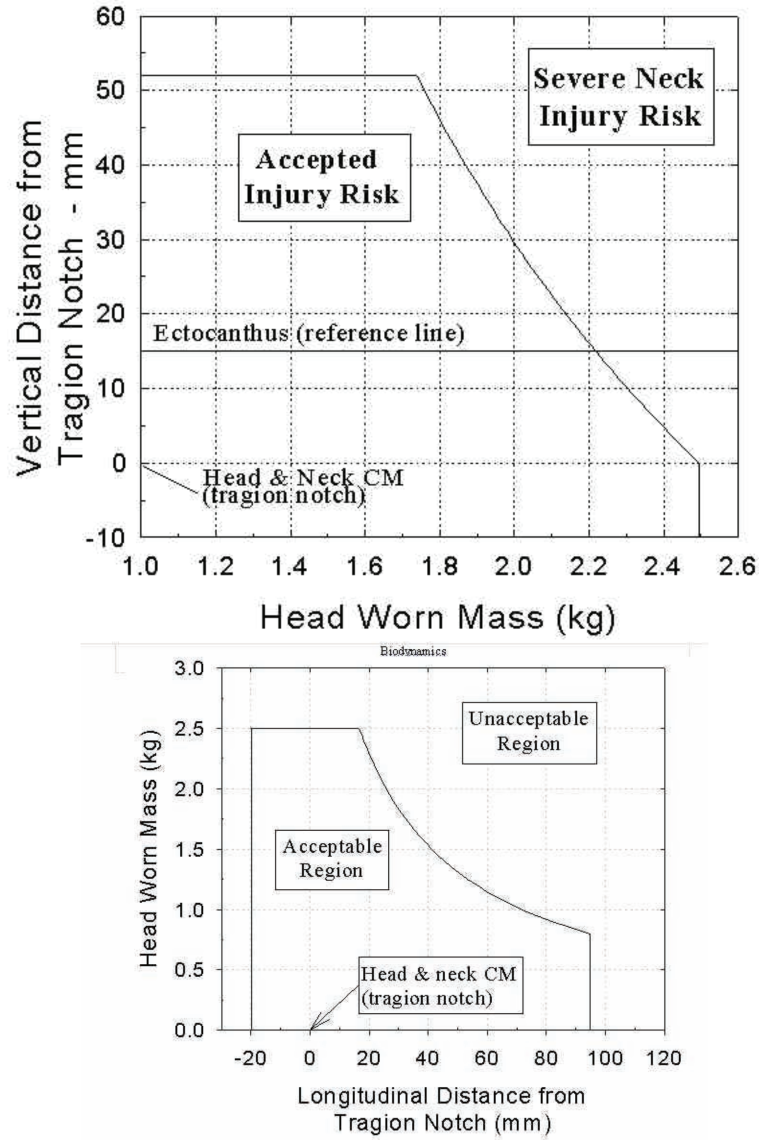


Figure 17-11. Vertical CM location as a function of head-supported weight/mass (top); allowable head-supported weight/mass as a function of longitudinal CM location (bottom).

- Maximum head-supported weight of 5 lb (2.5 kg) (including MBU-20/P oxygen mask and 3 inches [7.6 cm] of hose, helmet, visor and HMD components).
- Vertical CM limits (z-direction): Between +0.5 inches (1.3 cm) and +1.5 inches (3.8 cm) above the occipital condyles.
- *Threshold* CM horizontal limits (x-direction): between +0.5 inches (1.3 cm) forward and -0.8 inches (2 cm) aft of the occipital condyles.

- *Objective* CM horizontal limits (x-direction): between +0.2 inches (0.5 cm) forward and -0.8 inches (2 cm) aft of the occipital condyles.
- Lateral CM limits (y-direction): ± 0.15 inches (± 0.4 cm).

For the purposes of repeatable measurements and proper fitting, the test uses the size large Adam manikin head. This eliminates the impact of variations in human anthropometry (Buhrman, 2009).

Recommendations for dismounted Warfighters are unfortunately less well-defined. Researchers have examined injuries during dismounted Warfighter parachute operations (Bar-Dayan, Bar-Dayan and Shemer, 1998; Craig and Lee, 2000; Craig and Morgan, 1997; Ekeland, 1997; Lowdon and Wetherill, 1989). In most cases, it was concluded that most parachute jump injuries occurred on landing. The most compelling research to date on the effects of head-supported weight and CM on neck strain for dismounted Warfighter parachute jumps was performed at the USAARL (McEntire, Alem and Brozoski, 2004; McEntire and Alem, 2002; McEntire, Brozoski and Alem, 2003) in response to the additional head-supported weight required for the US Army's Land Warrior HMD program. They concluded that "inertial loads created during parachute opening shock with existing helmets do not frequently exceed human tolerance."

McEntire, Brozoski and Alem (2003) further compared their results to the Federal Motor Vehicle Safety Standards (FMVSS) values for neck injury and found that peak force and moment values were well below the accepted limits. The authors further compared their results with newly-proposed neck injury curves for flexion and extension ($+M_y$ and $-M_y$, forward and rearward bending around the ear-to-ear, or Y axis, respectively) using the Abbreviated Injury Scale (AIS)[36] severity scale of "3" (serious). They found a probability of injury to be less than 0.1%. This seems to agree with the findings of Craig and Morgan (1997), which found a 0.15% rate for back and neck injuries.

McEntire, Brozoski and Alem (2003) make it clear that although cadaveric and instrumented manikin results are useful, the data should be viewed carefully because of the lack of voluntary muscle control exerted during the events. Since the subjects cannot respond, there is also the lack of data regarding the "Ouch" factor, or the resulting instantaneous, post-event or chronic pain (McEntire, 2005, private communication). To address the remaining knowledge gaps, the USAARL is embarking on a three-year research effort into the development of models of cervical spine degeneration. One objective of this research program is to develop head-supported weight and CM location guidelines for mounted and dismounted Warfighters.

Frangibility

Frangibility refers to the ability of an HMD component to break free from the overall helmet-HMD system during a dynamic event. The purpose is to "shed" mass from the HMD system, thereby reducing the risk of neck injury during a dynamic event such as a helicopter mishap or ejection. Frangibility often is desired and even required when the total head-supported weight and CM creates the potential for unacceptable risk of neck injury.

A classic example of frangibility dates to the U.S. Army's early version of NVGs (AN/PNS-5). The AN/PVS-5 NVG was attached to the Soldier's Protective Helmet-4 (SPH-4) aviator helmet with "hook and pile" fasteners and elastic tubing. This did not allow the NVGs to easily or consistently detach during a crash. During ANVIS development, the attachment mechanism was re-designed with a spring-loaded "ball and socket" engagement, allowing the NVG to separate from the mount when exposed to a 10G to 15G load. The IHADSS helmet-display unit (HDU), mounted on the right lower edge of the helmet shell, is also designed to detach from the helmet under crash loadings. Shannon and Mason (1997) concluded a 10-year retrospective database study to determine the

[36] The *Abbreviated Injury Scale* (AIS) is an anatomical scoring system first introduced in 1969 to provide a reasonably accurate ranking of the severity of injury. Injuries are ranked on a scale of 1 to 6, with 1 being minor, 5 severe and 6 an unsurvivable injury.

injury rates of U.S. Army aviators involved in accidents and the relationship to wearing NVGs. Crewmembers wearing the non-frangible AN/PVS-5 NVG were shown to have 162% greater likelihood than non-NVG users to experience head or neck injury. Conversely, crewmembers wearing the frangible ANVIS had only a slightly higher, but *non-significant*, risk of head and neck injury as compared to non-NVG users. This reduced-injury probability was attributed to the frangibility of the ANVIS.

Current U.S. Army aviation frangibility (breakaway) design requirements state that when subjected to an acceleration of 9G or less in any vector within the limits described in Figure 17-12, the designed frangible components will *not* separate. However, separation *must* occur for acceleration of 15G or greater; and during breakaway, the frangible components should not come in contact with the wearer's forehead, eye sockets, or facial regions (Rash et al., 1996).

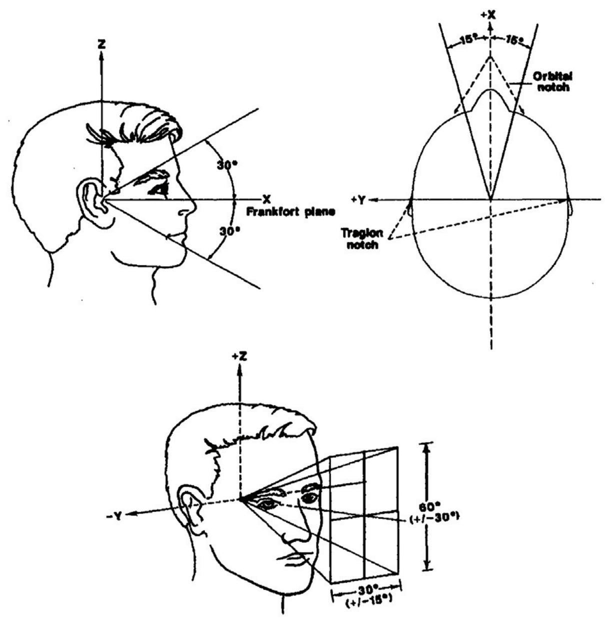


Figure 17-12. Vector limits for HMD breakaway force. If a 9G force occurs anywhere within the shown limits, the components will break away.

Frangibility also may be desirable for HMDs used in ground combat operations. Warfighters exiting military vehicles, moving through vegetation, or performing operations in urban environments are at risk of inadvertently snagging HMD cables (McEntire, 1998b). If the cable snags, tension in the HMD cable could induce excessive neck loading. A cable safety disconnect would reduce the risk of neck injury resulting from these mishaps.

Impact attenuation

Head impact injury is the leading cause of permanent disability and fatality in Army aviation rotary-wing mishaps (Shanahan and Shanahan, 1989; Shannon, Albano and Licina, 1996) and has been a major concern for aviators (Paschal et al., 1990; Trumble, McEntire and Crowley, 2005). This requirement is met with an outer shell and sufficient distance (volume) between the shell and the skull filled with energy-absorbing material (such as expanded polystyrene foam, see Brozoski and McEntire, 2003). The protective shell also resists penetration from sharp or jagged impact surfaces and distributes the load over a greater contact area.

Human tolerance to blunt head impacts is an area of ongoing research. Over the past 40 years, the USAARL has analyzed crash-damaged helmets and has recommended blunt impact performance standards for U.S. Army aviation helmets (McEntire, 1998a; Reading et al., 1984; Slobodnik, 1980). For the current generation of U.S.

Army aviation helmets,[37] the USAARL has recommended test head form accelerations thresholds of 150G to 175G, depending on the impact location. The USAARL-recommended value for the headband region (175G) is based on the concussion threshold to linear accelerations, not on skull fracture, fatality, or rotational acceleration thresholds. The recommended value for the earcup and crown regions (150G) is based on the risk of basilar skull fracture concomitant with impacts to those areas and the high frequency of occurrence in Army helicopter crashes (Shanahan, 1983). Impact attenuation requirements for military aircrew helmets continue to become more stringent.[38] Table 17-8 compares the impact and penetration resistance requirements for the HGU-56/P (rotary-wing) and the HGU-55/P (fixed-wing) helmets.

Table 17-8.
Differences for impact and penetration resistance values for the HGU-55/P (for fixed-wing applications) and the HGU-56/P (for rotary-wing applications).

Helmet	Impact Resistance	Penetration Resistance
HGU-56/P	Crown - <150g @ 4.8 m/sec Headband - <175 g @ 6.0 m/sec Earcups - <150g @ 6 m/sec	5 kg impactor, dropped from 1.52 m Max tear length <5 cm
HGU-55/P	<150g for less than 6 ms <200g for less than 3 ms <400g maximum	<0.25 inch penetration with 16 oz weight dropped from 10 feet

Fit, comfort and stability

It is difficult to put a precise metric on the fit or comfort of an HMD, though it is always immediately evident to the wearer. Even if the HMD image quality is excellent, the user will reject it if it doesn't fit well. Fitting and sizing is especially critical in the case of a HMD where in addition to being comfortable, it must provide a *precision* fit for the display to remain stable relative to the pilot's eye(s). Important issues for achieving a good fit with an HMD include:

- The user must be able to adjust the display to see the imagery.
- The HMD and helmet must be comfortable for a long duration of wear (4 to 6 hours) without causing "hot spots" and resist heat buildup.
- The HMD and helmet must not slip with sweating or under G-loading, vibration, or buffeting.
- The HMD and helmet must be retained during crash or ejection (except where breakaway capability is required).
- The weight of the head-borne equipment must be minimized.
- The mass-moment-of-inertia must be minimized.
- The mass of the head-borne components should be distributed to keep the CM (center-of-gravity) close to that of the head alone.

With the emphasis now being placed on blunt impact protection for ground Warfighters, combat helmet fitting systems are required to also provide blunt impact protection (Department of the Army, 2007). This is a paradigm

[37] The current standard US Army aviator's helmet is the HGU-56/P Aircrew Integrated Helmet System (AIHS), which is worn by all helicopter pilots with the exception (as of this writing) of the AH-64 Apache pilots.
[38] The Modular Aircrew Common Helmet (MACH) is a tri-service program intended to produce one common fixed- and rotary-wing helmet for the Army, Navy, and Air Force, to reduce the number of helmet configurations in the Department of Defense (DoD) inventory, reduce the logistical footprint, provide an effective platform for helmet mounted devices and increase aircrew safety.

shift from the aviation environment where the fitting system is not designed to contribute to the impact energy attenuation capabilities of the helmet. In addition to comfort, stability, and impact attenuation, additional attributes of fitting ease, sanitation, adjustability, durability, and maintainability should be considered when selecting a fitting system.

Another parameter is the anthropometric range[39] for the subject population and the number of helmet sizes needed. Fewer helmet sizes suggest the fitting system accommodate a greater anthropometric range. If designing a helmet system with a restricted exit pupil location, numerous helmet sizes may be required with a minimal thickness fitting system. One of the most common mistakes made by designers is to assume a correlation between various anthropometric measurements, because almost all sizing data are *univariate* – that is, they are completely uncorrelated with other data. For example, a person who has a 95th percentile head circumference will not necessarily have a 95th percentile interpupillary distance (Whitestone and Robinette, 1997). This was shown in a bivariate study that attempted to correlate head length and head breadth for male and female aviators, showing a large spread of data (Barnaba, 1997). Table 17-9 presents Gordon's et al. (1989) univariate (uncorrelated) anthropometric data for key head features for the range of sizes of the 5th percentile female up to the 95th percentile male.

Table 17-9.
The univariate (uncorrelated) anthropometric data for key head features.
Range of sizes for the 5th percentile female up to the 95th percentile male.
(Gordon et al., 1989) (expressed in cm)

Critical head dimensions (cm)	5% female	95% female	5% male	95% male
Interpupillary distance (IPD)	5.66	6.85	5.88	7.10
Head length	17.63	19.75	18.53	20.85
Head width	13.66	15.25	14.31	16.08
Head circumference	52.25	57.05	54.27	59.35
Head height (ectocanthus to top of head)	10.21	12.09	10.89	12.77

Note: The *head length* and *head height* measurements are head-orientation dependent.

Numerous methods of achieving a custom helmet fit have been devised. These include foam pads, sling suspension systems, and mesh-and-drawstring systems like that used with the AH-64 Apache helicopter specific IHADSS (Rash, 2001), as well as fitting systems that line the entire interior contour of the helmet such as the Thermoplastic Liner™ and the ZetaII™ (Figure 17-13). While differing in concept, these fitting systems each act as a physical interface between the wearer's head and the interior contour of the helmet's energy-absorbing liner (Rash, 2001).

Emerging HMD systems such as the TopOwl® and the Advanced Distributed Aperture System (ADAS™) simplify the fitting systems by creating custom energy-absorbing liners (EALs) for individual wearers, using data from three-dimensional laser scanners to create a model of the wearer's head. The data are used by a computer-controlled milling machine to carve the EAL from a block of expanded polystyrene foam (Brozoski and McEntire, 2003). The result is an EAL with a customized interior contour that matches contour of the intended wearer's head. Helmet manufacturers must still determine the minimum allowable EAL thickness needed to meet the impact attenuation requirements of the helmet. Knowing this, helmet designers work from the inside out to determine the number and size of helmet shells needed to accommodate the anthropometric range of the user population.

---

[39] *Anthropometry* – "the measure of Man" – is the compilation of data that define such things as the range of height for males and females, the size of our heads and how far apart our eyes are. Used judiciously, these data can help the HMD designer achieve a proper fit, though an over-reliance can be equally problematical.

Figure 17-13. Thermoplastic Liner™ (left) and ZetaII™ (right) helmet fitting systems.

Helmet retention

Helmet retention refers to the ability of the helmet to stay in place on the wearer’s head during dynamic events such as helicopter mishaps, ground vehicle accidents, or high speed ejections. This is critical because the helmet cannot perform its protective function if it has departed the wearer’s head or it has rotated to a position that leaves the skull open to direct impact.

The helmet retention system and, to a lesser extent, the fitting system are responsible for keeping the helmet in place during dynamic events. The *fitting system* provides frictional resistance to helmet motion relative to the head. The *retention system* performs the primary role of keeping the helmet in place by “anchoring” the helmet to prominent anatomical regions on the head. Reading et al. (1984) showed that the helmet retention system failure was a significant factor in mishaps where helmet losses occurred. Typically, modern retention systems consist of an integrated nape strap and a chinstrap. The nape strap runs behind the head just under the occipital region. The chinstrap runs under the chin, being careful to avoid the areas around and about the trachea. Properly designed, a retention system will prevent excessive forward or rearward rotation of the helmet when the head is exposed to dynamic accelerations (Hines et al., 1990).

Helmet retention requirements have typically consisted of a chinstrap load test. In these tests, the helmets were fixed in place while the chinstrap was loaded to a predetermined level. These tests checked the structural integrity of the retention system stitching and fastening systems, as failures of these components were identified as causes of retention system failure and helmet loss (Reading et al., 1984; Vyrnwy-Jones, Lanoue and Pritts, 1988). While chinstrap strength and elongation play a part in helmet retention, these quasi-static tests do not replicate the inertial loading experienced by the helmet-HMD system during aviation or ground vehicle mishaps. For this reason, dynamic retention tests like those developed by the USAARL (Brozoski and Licina, 2006) also are being incorporated into modern helmet performance specifications (Department of Defense, 2007).

Biodynamic and protection recommendations

- Because inertial loading has been shown to play a significant role in the risk of acute neck injury, head-supported weight and CM for the helmet and HMD combination must be tightly controlled relative to established standards for the Warfighter’s specific environment.

- Where appropriate, inertial neck loads resulting from adverse head-supported weight can be reduced through the use of frangible devices. Properly designed, these have been shown to reduce the risk of neck injury.
- Blunt impact protection should be a primary consideration from the outset of any HMD design. If an HMD is to be designed for use with an existing helmet, modifications that will degrade the impact attenuation properties of the helmet – especially those that reduce the stopping distance between the outer shell and the skull – must be avoided.
- The range of head sizes for the targeted user population and the role the fitting system will play in the HMD system must be well defined. The expected anthropometric range of wearers and minimum thickness of the fitting system will dictate the number of helmet sizes needed to fit the expected users. An over-emphasis on anthropometric data and an under-emphasis on fitting have resulted in extra helmet sizes and user rejection for comfort (Whitestone and Robinette, 1997).
- Retention system materials must be of sufficient strength to withstand the expected inertial loads without failure. The angular displacement of the helmet-HMD system relative to the head, resulting from the dynamic pulse, should not result in any optical systems, supporting surfaces or frames contacting the face, eyes, or forehead regions.

Perceptual and Cognitive Recommendations

HMD design is multidisciplinary involving both technology and the human perceptual system as discussed at length throughout this volume (see also Chapter 19, *The Potential of an Interactive HMD* for more information on cognitive processes and HMD interaction). Technology will change over time – more pixels in the displays, faster processors, lighter weight materials, smaller packaging and lower power consumption – but the human user will not. Rather, research will continue to uncover more about how humans interact with technology – a relatively new field called neuroergonomics (Parasuraman, 2003) – and provide us with better insight into how to design the human interface to the technology.

The HMD offers an opportunity to provide information to the Warfighter that uniquely replicates the way humans explore the environment by moving their head and eyes. This allows the Warfighter to remove the restrictions of the limited visual field of the cockpit, ground vehicle or hand-held display while enabling the ability to create situation awareness (SA) through the repeated information gathering, updating and prediction cycles necessary to accomplish his mission. Since the HMD must not compromise safety and it shares the valuable space on the head with a protective helmet, display components must earn their way onto the head by contributing to SA without incurring size, weight, power and cognitive overload penalties.

Too often HMDs provide only a re-mapping of head-down display information, placing the burden on the user to quickly process the raw *metaknowledge* with a minimum expenditure of cognitive resources. This is not always the case. The HMD can enable the user’s *filter* (by directing his attention to key events) and *fuel* (energizing his perceptual and cognitive resources) aspects of attention by the use of a head orientation tracker to quickly and accurately register the line-of-sight. If Earth-referenced (or conformal) or vehicle-referenced imagery is displayed on a see-through HMD, time-critical data can be cognitively obtained without placing additional workload on the user. One example is the Advanced Non-Distributed Flight Reference symbology (Jenkins, 2003; Jenkins, Turling & Brown, 2003; Jenkins, Sheesley and Bivetto, 2004, and see Figure 19-3) that quickly and intuitively informs the pilot of flight attitude and orientation. This provides the ability to maintain SA cycle without having to move the head back to the narrow FOV of the HUD or without the workload penalty from with switching attention (Spence and Driver, 1997). There also has been considerable research on the efficacy of using cross-modal cueing such as auditory as an alert for an impending time-critical visual event. It has also been found that 3-D audio cueing can enhance SA by superimposing geospatial directionality on cues or communications (Bolia, 2004).

User Adjustment Recommendations

One last but very important topic is that of the most direct interface the user has with the HMD, i.e., the controls that provide the user the ability to make adjustments to the display's characteristics. Despite the trend and the various arguments for automatic or self-adapting circuits and systems, the unique environments and situations encountered by the Warfighter, coupled with the potentially severe outcomes, argue for providing the user with the capability to make control inputs for the purpose of optimizing HMD information. Until advances in a number of scientific fields allow what would currently be considered as "futuristic" user-directed inactive control over HMD functions, (see Chapter 19, *The Potential of an Interactive HMD*), such adjustments most likely will be accomplished by hands-on controls, although voice-activation methods are an alternate approach (Baron and Green, 2006; Cohen and Oviatt, 1995; Kamm, 1995).

On HMD devices, both monocular and binocular, there should be mechanical, electronic, and/or optical adjustment mechanisms available for the user to optimize the attributes of the imagery and selection of displayed information. The mechanical adjustments are used primarily to align the optical axes and exit pupils of the device to the entrance pupils and primary lines of sight of the user, if required by the inherent design. The electronic adjustments may include display brightness, contrast, electronic focus, sizing, sensor sensitivity characteristics (gain and off-set for thermal sensors), etc. The optical adjustments may include the focus adjustments for the eyepieces and sensor objective lens, and magnification selection for targeting and pilotage sensors.

Recommendation Summary

Throughout this chapter, we have presented the various options HMD designs emphasizing the need to understand the user, their required tasks and environment – a human-centered design focus. In so doing, we identify the key requirements that are absolutely necessary – a process known as sub-optimization:

- *Ocularity* – If the Warfighter needs to only briefly view imagery such as maps or text, then a simple monocular HMD is sufficient (e.g., SO35A for Land Warrior). It will also be the lightest, least complicated and the least expensive. For longer term viewing, a monocular HMD may be acceptable (i.e., IHADSS or JHMCS), but a binocular design is best (i.e., HIDSS or JSF RSV), especially for fully immersive simulation and training applications (e.g., SR100A). This latter configuration provides best viewing comfort and improved detection, though it may be heavier, more complex and more expensive.
- *Field-of-view* – A large FOV (i.e., >60°) provides a sense of "being in" an image, key for immersive training and simulation applications. While also desired for flight applications, this must be balanced with biodynamics considerations that limit the horizontal FOV to approximately 40° (e.g., IHADSS and ANVIS NVGs). This must also be weighed against resolution requirements because the larger the FOV, the lower the resolution. Partial binocular overlap and optical tiling can enlarge the FOV without compromising resolution, but these do have their limitations.
- *Resolution* – While the limit of human visual resolution is 1 minute of arc, few HMDs can provide this primarily due to limitations in sensor and image source technologies. For an aircraft system (e.g., AH-64 Apache), it is not sufficient to specify just the HMD image source pixel count. Rather, we must compute the contributions from all subsystems from sensor to eye.
- *Pupil-forming versus non-pupil-forming optical design* – A non-pupil-forming design is best for a simple HMD viewing application. It will also be the most compact, the lightest, the least expensive and the imagery is viewable outside the "viewing eye box." The pupil-forming design is heavier and more expensive because of the additional lenses, though the longer path length provides design freedom to package the HMD around the head or protective helmet, moving the CM towards a more compatible location.

- *Exit pupil and eye relief* – These metrics represent the ability to comfortably view imagery, especially in an operational environment. The larger the exit pupil, the more tolerant the HMD will be to movement on the head during use. With a pupil-forming design, the image is not viewable outside the design exit pupil, so a minimum of 12 mm to 15 mm should be required, though the (operationally successful) IHADSS HMD has only a 10-mm exit pupil. With a non-pupil-forming design, the image is viewable even when outside the eye box, so the exit pupil may be smaller. The off-axis exit pupil (for both design forms) has a disproportionately large impact on overall optic size and weight, so allowing it to vignette (up to 50%) by truncating the diameter of the lens will reduce weight. A longer eye relief improves viewing comfort and allows users to wear prescription eyewear with the HMD, however, since most combiners are tilted, the classically defined eye relief (along the optical axis) is not always an accurate measure. Rather, we should consider the ECD, the distance from the eye to the closest point of the tilted combiner, which should be a minimum of 25 mm.
- *Optical distortion* – Residual optical aberrations can adversely affect image quality and these can be addressed through thorough evaluation of image quality during the design process. Residual distortion is not as easily managed as it may require additional lenses for correction. In this case, it is possible to pre-distort the image prior to the image source.
- *See-through versus non-see-through considerations* – If the HMD is intended for fixed- or rotary-wing applications, a see-through HMD is needed. This lets us superimpose symbology or imagery (aircraft- versus earth- versus screen-referenced) on a see-though combiner. Doing so unlocks the pilot from the limited and fixed-forward FOV of the HUD or cockpit displays, providing imagery wherever he is looking. Pilots would like to have a combiner with as much see-through as possible to let them see farther, though this has implications for the image source luminance. For a Warfighter viewing text or map information, a non-see-through design may be preferable because: 1) this type of imagery could be confusing against a normal background, 2) the non-see-through design has better image source to eye transmission and will therefore require less power and 3) a non-see-through design will be more covert at night.
- *Luminance and contrast* – In order for the Warfighter to see the HMD imagery, we must know the range of ambient light levels in which he will prosecute his mission. For a pilot to view HMD imagery against a high ambient background, we must determine the transmission values of the canopy, any visor and the combiner to arrive at a value for the image source luminance based upon a minimum contrast ratio requirement. For the dismounted Warfighter wearing a non-see-through display, a value of 35 fL to 50 fL should suffice for daytime viewing with the ability to reduce the luminance to 0.1 fL at night.
- *Considerations for helmet-mounted sensors* – Adding sensors to a Warfighter's helmet (e.g., the AN/PVS-14 or AN/PSQ-20) augments their vision under low light conditions, but when configured in line with their eyes may have negative head-supported weight and CM/CG effects. Mounting the sensors at a more favorable location on the top or sides of the helmet may present adverse perceptual artifacts such as offset hand-eye coordination (monocular) or hyperstereopsis (binocular). Limited data indicate that a perceptual re-calibration is possible, though with unknown residual aftereffects. Research is continuing in this area.
- *Acoustics/Auditory* – The helmet/head-gear component of the HMD system should optimally allow for exposure of both pinnae to environmental sounds or for the use of an acoustically transparent headgear covering the ears. This headgear should have an acoustically-optimized shape to minimize dispersion and shadowing of natural sounds. Level-dependent, in-the ear hearing protection is recommended when hearing protection is required. An always-on audio communication system based on bone conduction or whisper-quality audio interface with an easy to access step-wise volume control should be employed. Designs should include provisions to use biological and chemical protection gear without introducing detrimental effects on audio communication and noise protection.

- *Biodynamics* – The primary purpose of the Warfighter's helmet is protection and only secondly as a mounting platform for the HMD components, which – because of the additional head-supported weight and CM – can contribute to increased fatigue and injury potential. Strict guidelines for head-supported weight and CM have been established which will minimize this likelihood for pilots. For the dismounted Warfighter, these guidelines have not been as firmly established. Though most head and neck injuries occur upon parachute landing fall, strict limits are still being determined. In all cases, the implications of long-term wear of a helmet with HMD components have not been established, though research is also continuing in this area.
- *Perceptual/Cognitive* – All HMD components should earn their way onto the head because they reduce Warfighter workload and enable him to accomplish his mission. Information should not simply be a re-mapping of available cockpit information, but should be cognitively pre-digested to ease the transfer of information while not overloading working memory.
- *User adjustment* – The selection and implementation of available adjustments must allow for individual differences while carefully avoiding complexity and minimizing the potential for misadjustment.

References

Albery, C.B., and Kaleps, I. (1997). A procedure to measure the mass properties of helmet systems. *Proceedings of the design and Integration of Helmet Systems Symposium*. Framingham, MA. 2-5 December.

Alem, N.M., and Meyer, M.D. (1995). Effects of head supported devices on pilot performance during simulated helicopter rides. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 95-37.

Arbak, C. (1989). Utility evaluation of a helmet-mounted display and sight. *Proceedings of the SPIE, Helmet-Mounted Displays*, 1116, 138-141.

Ashrafiuon, H., Alem, N.M., and McEntire, B.J. (1997). Effects of weight and center of gravity location of head-supported devices on neck loading. *Aviation, Space and Environmental Medicine*, 68, 914.

Azuma, R. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355-385

Bahill, A.T., Adler, D., and Stark, L. (1975). Most naturally occurring human saccades have magnitudes of 15 degrees or less. *Investigative Ophthalmology*, 14, 468-9.

Balram, N., and Olsen, B. (1996). *Proceedings of the Society for Information Displays, SID 96 Digest*, Paper P-10, 429-432

Barazanji, K.W., and Alem, N.M. (2000). Effects of head-supported devices on female aviators during simulated helicopter rides. Part I: Biomechanical response. Fort Rucker, AL: U. S. Army Aeromedical Research Laboratory. USAARL Report No. 2000-16.

Bar-Dayan, Y., Bar-Dayan, Y., and Shemer, J. (1998). Parachuting injuries: a retrospective study of 43,542 military jumps. *Military Medicine*, 163, 1-2.

Barnaba, J.M. (1997). Human factors issues in the development of helmet mounted displays for tactical, fixed-wing aircraft. *Proceedings of the SPIE, Head-Mounted Displays II*, 3058, 2.

Baron, A., and Green, P. (2006). Safety and usability of speech interfaces for in-vehicle tasks while driving: A brief literature review. Ann Arbor, MI: Transportation Research Institute, University of Michigan. Technical Report UMTRI-2006-5.

Bass, C.R., Donnellan, L., Salzar, R., Lucas, S., Folk, B., Davis, M., Rafaels, K., Planchak, C., Meyerhoff, K., Ziemba, A., and Alem, N. (2006). A new neck injury criterion in combined vertical/frontal crashes with head supported mass. *Proceedings of the 2006 International IRCOBI conference on the biomechanics of impact*. Madrid, Spain. 20-22 September.

Belt, R., Kelley, K., and Lewandowski, R. (1998). Evolution of helmet mounted display requirements and Honeywell HMD/HMS systems. *Proceedings of the SPIE, Helmet- and Head-Mounted Displays III*, 3362, 373-384.

Bertelson, P., and de Gelder, B. (2004). The psychology of multimodal perception. In: Spence, C., and Driver, J. (Eds.), *Crossmodal Space and Crossmodal Attention*. Oxford: Oxford University Press.

Birt, J.A., and Task, H.L. (1973). Proceedings of a Symposium on Visually Coupled Systems: Development and Application. Held at Brooks Air Force Base, Texas 8-10 November 1972. Brooks AFB, TX: Aerospace Medical Division.

Boff, K.R., and Lincoln, J.E. (1988). *Engineering Data Compendium, Human Perception and Performance*. Wright-Patterson AFB, OH: Harry G. Armstrong Aerospace Medical Research Laboratory.

Bolia, R.S. (2004). Special issue: Spatial audio displays for military aviation. *International Journal of Aviation Psychology*, 14, 233-238.

Breitmeyer, W.A., and Reetz, F. (1985). Visual and spectroradiometric performance criteria for night vision goggles (NVG) compatible aircraft interior lighting. *AGARD Conference Proceedings, No. 379, Visual Protection and Enhancement*, Athens, Greece, 22-24 April.

Brennesholtz, M.S., and Stupp, E.H. (2008). *Projection Displays.* John West Sussex, UK: Wiley and Sons.

Brolin, K., Hedenstierna, S., Halldin, P., Bass, C., and Alem, N. (2008). The importance of muscle tension on the outcome of impacts with a major vertical component. *International Journal of Crashworthiness*, 13(5), 487-498.

Brozoski, F.T., and Licina, J.R. (2006). Static and dynamic retention assessment of the HGU-56/P aircrew integrated helmet system equipped with quick-release ALPHA and snap fastener retention assemblies. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2006-02.

Brozoski, F.T., and McEntire, B.J. (2003). Evaluation of novel energy absorbing materials. *Proceedings of the RTO Applied Vehicle Technology Panel (AVT) / Human Factors and Medicine Panel (HFM) Joint Specialists' Meeting on Equipment for Personal Protection (AVT-097) / Personal Protection: Bio-Mechanical Issues and Associated Physio-Pathological Risks (HFM-102)*. RTO-MP-AVT-097. Koblenz, Germany. 19-23 May.

Buhrman, J. (2009), Biomechanics Branch, Wright Patterson Air Force Base, OH. Personal Communication.

Burton, R.R. (1999a). Acute neck injuries. *Cervical Spinal Injury from Repeated Exposures to Sustained Acceleration*, RTO Technical Report 4 (RTO-TR-4). Neuilly-Sur-Sein Cedex, France.

Burton, R.R. (1999b). Degenerative diseases of the cervical spine. *Cervical Spinal Injury from Repeated Exposures to Sustained Acceleration*, RTO Technical Report 4 (RTO-TR-4). Neuilly-Sur-Sein Cedex, France.

Burton, R.R. (1999c). Summary and conclusions. *Cervical Spinal Injury from Repeated Exposures to Sustained Acceleration*, RTO Technical Report 4 (RTO-TR-4). Neuilly-Sur-Sein Cedex, France.

Butler, B.P. (1992). Helmeted head and neck dynamics under whole-body vibration. Ph.D. dissertation, Ann Arbor, MI: University of Michigan.

Butler, B.P., and Alem, N.M. (1997). Long-duration exposure criteria for head-supported mass. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 97-34.

CAE Electronics. (1989). Wide field of view, helmet-mounted infinity display system development. Brooks AFB, TX: U.S. Air Force Systems Command. AFHRL-TR-84-27.

Cakmakci, O., and Rolland, J. (2006). Head-worn displays: a review. *Journal of Display Technology*, 2(3), 199-217

Chisum, G.T. (1975). Head-coupled display visual design considerations. *Aviation, Space, and Environmental Medicine*, 11, 1373-1377.

Cohen, P.R., and Oviatt, S.L. (1995). The role of voice input for human-machine communication. *Proceedings of the National Academy of Science*, 92, 9921-9927.

Craig, S.C., and Lee, T. (2000). Attention to detail: injuries at altitude among U.S. Army Military static line parachutists, *Military Medicine*, 165, 268-71.

Craig, S.C., and Morgan, J. (1997). Parachuting injury surveillance, Fort Bragg, North Carolina, May 1993 to December 1994, *Military Medicine,* 162, 162-4.

Curry. D.C., Harrington, L.K., and Hopper, D.G. (2006). Dichoptic image fusion in human visual system, *Proceedings of the SPIE, Helmet- and Head-Mounted Displays XI: Technologies and Applications,* 6224, 1-11.

Deavers, M.B., and McEntire, B.J. (1993). An automated method for determining mass properties. US Army Aeromedical Research Laboratory. Fort Rucker, AL. USAARL Report No. 93-4.

Department of Defense, (2001). Department of Defense Standard for Lighting, Aircraft, Night Vision Imaging System (NVIS) Compatible (MIL-STD-3009).

Department of Defense. (2007). Performance specification: Modular aircrew common helmet (MACH). PRF-TA-MACH-001-07. 19 April.

Department of the Army. (2007). Purchase description: Helmet, advanced combat (ACH). Fort Belvoir, VA. CO/PD-05-04. 30 October.

Doczy, E, Mosher, S., and Burhman, J. (2004). The effects of variable helmet weight and subject bracing on neck loading during frontal – Gx impact. *Proceedings of the 42nd Annual SAFE Association Symposium*. Salt Lake City, UT. 27-28 September.

Dornheim, M. (1995). VTAS sight fielded, shelved in 1970s. *Aviation Week and Space Technology,* 23 October, 51.

Dunlap, D. (2002). Noise reduction for improved communications, *Armed Forces International*. Retrieved 20 March 2009 from: www.armedforces-int.com

Ekeland, A. (1997). Injuries in military parachuting: a prospective study of 4499 jumps. *Injury,* 28, 219-222.

Ellis, S.R. (1995). Origins and Elements of Virtual Environments. In: Barfield, W., and Furness, T.A. (Eds.), *Virtual Environments and Advanced Interface Design.* New York: Oxford University Press.

Eveland, E.S., Esken, B., Shouse, N., Goodyear, C., and Kane, M. (2008). Neck muscle fatigue resulting from prolonged wear of weighted helmets under high G acceleration. Wright Patterson Air Force Base, OH: U.S. Air Force Research Laboratory. Report No. AFRL-RH-WP-TR-2008-0085.

Eveland, E.S., and Goodyear, C. (2001). Neck muscle response to changes in helmet loading under +Gz acceleration – Gender differences. Wright Patterson Air Force Base, OH: U.S. Air Force Research Laboratory. Report No. AFRL-HE-WP-TR-2002-0211.

Farrell, R.J., and Booth, J.M. (1984). *Design handbook for Imagery Interpretation Equipment*. Seattle, WA: Boeing Aerospace Co.

Fischer, R.E. (1997). Fundamentals of HMD Optics. In: Melzer, J.E., and Moffitt, K.W. (Eds.), *Head-Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Flanagan, P., Stuart, G.W., and Gibbs, P. (2007a). Hyperstereopsis in helmet-mounted NVDs: Absolute distance perception. *Proceedings of the SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications,* 6557, 65570J1-7.

Flanagan, P., Stuart, G.W., and Gibbs, P. (2007b). Hyperstereopsis in helmet-mounted NVDs: Time to contact estimation. *Proceedings of the SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications*, 6557, 65570O1-7.

Foote, B. (1998). Design guidelines for advanced air-to-air helmet-mounted display systems. *Proceedings of the SPIE, Helmet- and Head-Mounted Displays III*, 3362, 94-102.

Fraser, S., Alem, N., and Chancey, V.C. (2006). Helicopter flight performance with head-supported mass. *Proceedings of the American Helicopter Society 62nd Annual Forum*. Phoenix, AZ. 9-11 May.

Friberg, N. (1997). Helmet Mounted Sight System ODEN, the simple solution. *Proceedings of the SPIE, Head-Mounted Displays II*, 3058, 191-218.

Glumm, M. M., Branscome, T.A., Patton, D.H., Mullins, L.L., and Burton, P.A. (1999). The effects of an auditory versus a visual presentation of information on soldier performance. *Human Factors and Ergonomics Society Annual Meeting Proceedings, System Developmen*t, 1041-1044.

Gordon, C.C., Churchill, T., Clauser, C.E., Bradtmiller, B., McConville, J.T., Tebbetts, I., and Walker, R.A. (1989). 1988 Anthropometric survey of U.S. Army personnel: Summary statistics interim report. Natick, MA: U.S. Army Natick Research and Development Laboratories. Technical Report TR-89/027.

Guill, F.C., and Herd, G.R. (1989). An evaluation of proposed causal mechanisms for "ejection associated" neck injuries, *Aviation, Space and Environmental Medicine*, A26, July.

Hämäläinen, O., and Kuronen, P. (1999). Cervical injury and degeneration in the Finnish Air Force. *Cervical Spinal Injury from Repeated Exposures to Sustained Acceleration*, RTO Technical Report 4 (RTO-TR-4), Neuilly-Sur-Sein Cedex, France.

Harms-Ringdahl, K., Linder, J., Spångberg, C., Burton, R.R. (1999). Biomechanical considerations in the development of cervical spine pathologies. *Cervical Spinal Injury from Repeated Exposures to Sustained Acceleration, RTO Technical Report 4 (RTO-TR-4)*, Neuilly-Sur-Sein Cedex, France.

Halldin, P., Brolin, K., Hedenstierna, S., Aare, M., and von Holst, H. (2005). Finite element analysis of the effects of head-supported mass on neck responses – Complete Phase Two Report. Stockholm, Sweden: Royal Institute of Technology, School of Technology and Health, Division of Neuronic Engineering. Contractor report to U.S. Army. Contract number N62558-03-C-0013.

Harding, T.H., Beasley, H.H., Martin, J.S., Rash, C.E., McLean, W.E., Mora, J.C., and McEntire, B.J. (1998). Optical and biodynamic evaluation of the helmet integrated display sight system (HIDSS) for the RAH-66 Comanche development and validation program phase. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 98-22.

Hedge, A. (2007). Auditory displays. Retrieved 5 July 2008 from: http://ergo.human.cornell.edu/studentdownloads/DEA325pdfs/idauditory.pdf

Hines, R.H., Palmer, R.W., Haley, J.L., and Hiltz, E.E. (1990). Development of an improved SPH-4 retention assembly. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 90-9.

Hixon, J.G., Jacobs, E.L., Vollmerhausen, R.H. (2004). Target detection cycle criteria when using the targeting task performance metric. *Proceedings of the SPIE, Electro-Optical and Infrared Systems: Technology and Applications*, 5612, 275-283.

Hoppe, M.J., and Melzer, J.E. (1999). Optical tiling for wide-field-of-view head-mounted displays. *Proceedings of the SPIE, Current Developments in Optical Design and Optical Engineering VIII*, 3779, 146.

Hughes, R.L., Chason, L.R, and Schwank, J.C.H. (1973). Psychological considerations in the design of helmet-mounted displays and sights: Overview and annotated bibliography. Wright-Patterson AFB, OH: U.S. Air Force Aerospace Medical Research Lab. Technical Report AMRL-TR-73-16.

Jackson, T.W., and Craig, J.L. (1999). Design, development, fabrication, and safety-of-flight testing of a panoramic night vision goggle. *Proceedings of the SPIE, Head- and Helmet-Mounted Displays IV*, 3689, 98.

Jenkins, J.C. (2003). Development of helmet-mounted display symbology for use as a primary flight reference, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays VIII: Technologies and Applications*, 5079, 333-345.

Jenkins, J.C., Thurling, A.J., and Brown, B.D. (2003). Ownship status helmet-mounted display symbology for off-boresight tactical applications, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays VIII: Technologies and Applications*, 5079, 346-360.

Jenkins, J.C., Sheesley, D.G., and Bivetto, F.C. (2004). Helmet-mounted display symbology for enhanced trend and attitude awareness, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays IX: Technologies and Applications*, 5442, 164-178.

Kalawsky, R.A. (1993). *The Science of Virtual Reality and Virtual Environments.* New York, Addison-Wesley Publishing Company.

Kalich, M.E., Rash, C.E., McLean, W.E., and Ramiccio, J.G. (2007). A limited flight study for investigating hyperstereo vision, *Proceedings of the SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications*, 6557, 65570I1-14.

Kamm, C. (1995). User interfaces for voice applications. Proceedings of the National Academy of Science, 72(22), 10031-10037.

Kasper, E.F., Haworth, L.A., Szoboszlay, Z.P., King, and R.D., Halmos, Z.L. (1997). Effects of in-flight field-of-view restriction on rotorcraft pilot head movement. *Proceedings of the SPIE, Head-Mounted Displays II,* 3058, 34.

Klymenko, V., Verona, R.W., Beasley, H.H., and Martin, J.S. (1994). Convergent and divergent viewing affect luning, visual thresholds, and field-of-view fragmentation in partial binocular overlap helmet-mounted displays. *Proceedings of the SPIE, Helmet- and Head-Mounted Displays and Symbology Design Requirements*, 2218, 2.

Klymenko, V., Harding, T.H., Beasley, H.H., Martin, J.S., and Rash, C.E. (2000). Investigation of helmet-mounted display configuration influences on target acquisition, *Proceedings of the SPIE, Head- and Helmet-Mounted Displays,* 4021, 316.

Knox, F.S., Buhrman, J.R., Perry, C.E, and Kaleps, I. (1991). Interim head/neck criteria. Consultation Report. Wright-Patterson Air Force Base, OH: Armstrong Laboratories, Escape and Impact Protection Branch.

Kooi, F.L. (1993). Binocular configurations of a night-flight head-mounted display. *Displays*, 14(1), 11-20.

Landau, F. (1990). The effect of visual recognition performance on misregistration and overlap for a biocular helmet mounted display. *Proceedings of the SPIE, Helmet-Mounted Displays II,* 1290, 173-184.

Leachtenauer, J.C. (2003). Resolution requirements and the Johnson Criteria revisited. *Proceedings of the SPIE, Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XIV,* 5076, 1-15.

Lloyd, J.M. (1975). *Thermal Imaging Systems*. New York: Plenum Press.

Lowdon, I.M., and Wetherill, M.H. (1989). Parachuting injuries during training descents. *Injury, 20*, 257-8.

Mace, T.K., Van Zyl, P.H., Cross, T. (2001). Integration, development, and qualification of the helmet-mounted sight and display on the Rooivalk Attack Helicopter, *Proceedings of the SPIE, Helmet and Head-Mounted Displays VI,* 4361, 12

MacMillan, R.T., Brown, R.W., Wiley, L.L. (1995). Safety of flight testing for advanced fighter helmets. *Proceedings of the SPIE, Helmet- and Head-Mounted Displays and Symbology Design Requirements II,* 2465, 122-129

Merkle, A.C., Kleinberger, M., and Uy, O.M. (2005). The effects of head-supported mass on the risk of neck injury in Army personnel. *Johns Hopkins University Applied Physics Laboratory Technical Digest*, 26(1), 75-83.

McEntire, B.J. (1998a). U.S. Army aircrew helmets: Head injury mitigation technology (Reprint). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 98-12.

McEntire, B.J. (1998b). Ground troop helmet electronic cable safety design issues (Reprint). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 98-24.

McEntire, B.J. (1998c). Biodynamics. In: Rash, C.E. (Ed.), Helmet-Mounted Displays: *Design Issues for Rotary Wing Aircraft*. Bellingham, WA: SPIE Press.

McEntire, B.J. (2005). U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL. Private communication.

McEntire, B.J., Alem, N., and Brozoski, F.T. (2004). Parachutist neck injury risk associated with head-borne weight. *U.S. Army Medical Department Journal*, April-June, 30-34.

McEntire, B.J., and Alem, N. (2002). Land Warrior helmet: biomechanical issues related to airborne operations. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL TM-02.

McEntire, B.J., and Brozoski, F.T., Alem, N. (2003). Neck injury from helmet mounted devices during parachute opening. Proceedings of NATO RTO Specialist Conference, Koblenz.

McEntire, B.J., and Shanahan, D.F. (1998). Mass requirements for helicopter aircrew helmets (reprint). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 98-14.

McLean, B., and Smith, S. (1987). Developing a wide field of view HMD for simulators. *Proceedings of the SPIE, Display System Optics,* 778, 79-82.

Melzer, J.E. (1998). Overcoming the field of view: resolution invariant in head-mounted displays. *Proceedings of the SPIE, Helmet- and Head-Mounted Displays III,* 3362, 284-293.

Melzer, J.E. (2006). Head-Mounted Displays. In: Spitzer, C.R. (Ed.), *The Avionics Handbook,* 2nd Edition. CRC Press: Boca Raton, FL

Melzer, J.E., and Larkin, E.W. (1987). An integrated approach to helmet display system design, *Proceedings of the SPIE, Proceedings of the SPIE: Display System Optics,* 778, 83-88.

Melzer, J.E., and Moffitt, K. (1989). Partial binocular-overlap in helmet-mounted displays. *Proceedings of the SPIE, Display System Optics II,* 1117, 56-62.

Melzer, J.E., and Moffitt, K. (1991). An ecological approach to partial binocular-overlap. *Proceedings of the SPIE, Large Screen Projection, Avionic, and Helmet-Mounted Displays,* 1456, 124-131.

Melzer, J.E., and Moffitt, K. W. (Eds.) (1997). *Head-Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Melzer, J.E., and Moffitt, K. W. (2007). Sensor-offset HMD perception and performance. *Proceedings of the SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications,* 6557, 65570G.

Merryman, R.F.K. (1994). Vista Sabre II: integration of helmet-mounted tracker/display and high off-boresight missile seeker into F-15 aircraft. *Proceedings of the SPIE, Helmet- and Head-Mounted Displays and Symbology Design Requirements,* 2218, 173.

Moffitt, K.W. (1989). Ocular responses to monocular and binocular helmet-mounted display configurations. *Proceedings of the SPIE, Helmet-Mounted Displays,* 1116, 142.

Moffitt, K.W. (1997). Designing HMDs for viewing comfort. In: Melzer, J.E., and Moffitt, K.W. (Eds.), *Head-Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Moffitt, K., and Melzer, J.E. (1993). Method and apparatus to reduce rivalry in a partial overlap binocular display. U.S. Patent 5,233,458.

Moffitt, K. (2008). *Head-Mounted Display Image Configurations*, Unpublished document, downloaded January 2009 from: http://www.kirkmoffitt.com

Mohananchettiar, A., Cevher, V., CuQlock-Knop, G., Chellappa, R., and Merritt, J. (2007). Hyperstereo algorithms for the perception of terrain drop-offs. *Proceedings of the SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications,* 6557, 65570L-1-10.

Mohler, B.J., Thompson, W.B., Creem-Regehr, S.H., Willemsen, P., Pick, H.L., and Reiser, J.J. (2007). Calibration of locomotion resulting from visual motion in a treadmill-based virtual environment. *ACM Transactions on Applied Perception,* 4(1), 4.

Monroe, J.R. (2004). Sound identification recordings for warfighters (SIR). Charlottesville, VA: U.S. Army National Ground Intelligence Center.

National Research Council. (1995). *Human Factors Tactical Display for the Individual Soldiers. Panel on Human Factors in the Design of Tactical Display Systems for the Individual Soldier, Committee on Human Factors.* Washington, DC: National Academy Press.

National Research Council. (1997). *Tactical Display for Soldiers: Human Factors Considerations. Panel on Human Factors in the Design of Tactical Display Systems for the Individual Soldier.* Washington, DC: National Academy Press.

Newman, R.L., and Greeley, K.W. (1997). Helmet-Mounted Display Design Guide. San Marcos, TX: Crew Systems Consultants.

Oldsmobile Club of America. (2006). Oldsmobiles pace the Race. Retrieved 4 July 2008 from: http://www.oldsclub.org/History/PaceCars/IndyOldsPaceCars.htm#1988%20Oldsmobile%20Cutlass%20Supreme

Olson, W.P., and Balram, N. (1996). Lower bound for LCD image quality, *Proceedings of the SPIE, Liquid Crystal Materials, Devices and Applications IV,* 2651, 108-116.

Pankratov, M.M. (1995). New surgical three-dimensional visualization system, *Proceedings of the SPIE, Lasers in Surgery: Advanced Characterization, Therapeutics, and Systems V,* 2395, 143.

Parasuraman, R. (2003). Neuroergonomics: Research and practice. *Theoretical Issues in Ergonomics Science*, 4, 5-20.

Paschal, C.R., Palmer, R.W., Shanahan, D.F., and Haley, J.L. (1990). Evaluation of the head injury hazard during military parachuting. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 90-6.

Paskoff, G. (2004). Cervical injury risk resulting from rotary wing impact: Assessment of injury based upon aviator size, helmet mass properties, and impact severity. Patuxent River, MD: Naval Air Warfare Center. Report No. NAWCADPAX/TR-2004/86.

Patterson, R., Winterbottom, M.D., and Pierce, B.J. (2006). Perceptual issues in the use of head-mounted visual displays, *Human Factors,* 48 (3), 555-573.

Perry, C.E., and Buhrman, J.R. (1997). Biomechanics in HMDs. In: Melzer, J.E., and Moffitt, K.W. (Eds.), *Head-Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Prinzel III, L.P., and Risser, M. (2004). Head-up displays and attention capture. National Aeronautics and Space Administration. NASA/TM-2004-213000.

Priot, A.E., Hourlier, S., Giraudet, G., Leger, A., and Roumes, C. (2006). Hyperstereopsis in night vision devices: basic mechanisms and impact for training requirements, *Proceedings of the SPIE, Helmet- and Head-Mounted Displays XI: Technologies and Applications,* 6224, 62240N1-11.

Proctor, P. (1999). Helmet displays boost safety and lethality. *Aviation Week and Space Technology, 150*(5), 81-83.

Rash, C.E. (Ed.). (2001). *Helmet-Mounted Displays: Design Issues for Rotary Wing Aircraft*. Bellingham, WA: SPIE Press.

Rash, C.E., Kalich, M.E., van de Pol, C., and Reynolds, B.S. (2002). The issue of visual correction compatibility with helmet-mounted displays, Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2003-04.

Rash, C.E., Mozo, B.T., McLean, W.E., McEntire, J.L., and Licina, J.R. (1996). RAH-66 Comanche health hazard and performance issues for the Helmet Integrated Display and Sighting System. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 97-1.

Rash, C.E., and Verona, R.W. (1992). The human factor considerations of image intensification and thermal imaging systems. In: Karim, M.A. (Ed.). *Electro-Optical Displays.* New York: Marcel Dekker, Inc.

Reading, T.E., Haley, J.L., Jr., Sippo, A.C., Licina, J.R., and Schopper, A.W. (1984). SPH-4 U.S. Army flight helmet performance, 1972-1983. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 85-l.

Rogers, P.J., and Freeman, M.H. (1992). Biocular display optics. In: Karim, M.A. (Ed.), *Electro-Optical Displays.* New York, Marcel Dekker, Inc.

Shanahan, D.F. (1983). Basilar skull fracture in U.S. Army aircraft accidents. *Aviation, Space, and Environmental Medicine,* 54(7), 628-631.

Shanahan, D.F., and Shanahan, M.O. (1989). Injury in U.S. Army helicopter crashes fiscal years October 1979-September 1985. *Journal of trauma,* 29(4), 415-423.

Shannon, S.G., Albano, J.P., and Licina, J.R. (1996). Head injury risk in U.S. Army rotary-wing mishaps: Changes since 1980. *Proceedings of the AGARD Specialist's Meeting on Impact Head Injury*, 7-9 November, Mescalero, New Mexico.

Shannon, S.G., and Mason, K.T. (1997). U.S. army aviation life support equipment retrieval program: head and neck injury among night vision goggle users in rotary-wing mishaps. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 98-02.

Slobodnik, B.A. (1980). SPH-4 helmet damage and head injury correlation. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 80-7.

Schor, C., Landsman, L. and Ericson, P. (1987). Ocular dominance and the interocular suppression of blur in monovision. *American Journal of Optometry and Physiological Optics,* 64, 723-30.

Self, H.C. (1986). Critical tolerances for alignment and image differences for binocular helmet-mounted displays, Technical Report AAMRL-TR-86-019, Wright-Patterson AFB OH: Armstrong Aerospace Medical Research Laboratory.

Settecerri, J.J., McKenzie, J., Privitzer. E., and Beecher, R.M. (1986). Mass properties and inertial loading effects of head encumbering devices. *Proceedings 24th Annual SAFE Conference*, December.

Simons, R., and Melzer, J.E., (2003). HMD-based training for the U.S. Army's AVCATT-A collective aviation training simulator. *Proceedings of the SPIE, Helmet- and Head-Mounted Displays VIII: Technologies and Applications,* 5079, 1-6.

Simonite, T. (2007). Thinking of words can guide your wheelchair. *NewScientistTech*. Retrieved 18 March 2009 from*:* www.newscientist.com/article/dn12602

Simonite, T. (2008). Nerve-tapping neckband allows ‘telepathic’ chat. *NewScientistTech*. Retrieved 18 March 2009 from: http://www.newscientist.com/article/dn13449

Smith, G., and Atchison, D. A. (1997). *The eye and visual optical instruments.* New York: Cambridge University Press.

Spence, C., and Driver, J. (1997). Cross-modal links in attention between audition, vision and touch: implications for interface design, *International Journal of Cognitive Ergonomics,* 1, 351-373

Sobel, D. (1996). *Longitude: the true story of a lone genius who solved the greatest scientific problem of his time*, New York: Penguin Books.

Steenblik, J.W. (1989). Alaska Airlines’ HGS. *Air Line Pilot,* December, 10-14.

Stuart, G.W., Flanagan, P., and Gibbs, P. (2007). Hypersteropsis in helmet-mounted NVDs: Slope Perception. *Proceedings of the SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications,* 6557, 65570K1-8.

Sutherland, I. (1968). A head-mounted three-dimensional display. *Proceedings of the Fall Joint Computer Conference, American Federation of Information Processing Societies*, 33, 757-764.

Szoboszlay, Z.P., Haworth, L.A., Reynolds, T.L., Lee, A.G., and Halmos, Z.L. (1995). Effect of field-of-view restriction on rotorcraft pilot workload and performance: preliminary results, *Proceedings of the SPIE, Helmet- and Head-Mounted Displays and Symbology Design Requirements II,* 2465, 142.

Task, H.L. (1997). HMD image sources, optics and visual interface. In: Melzer, J.E., and Moffitt, K.W. (Eds.), *Head-Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Thornton, J.M., and Zaborowski, D.J. (1992). Mass properties test procedure for manikin headforms and helmet systems. Washington, D.C.: Naval Air Systems Command. Report No. NAWCADWAR-92092-60.

Trumble, C.C., McEntire, B.J., and Crowley, J.S. (2005). Blunt head injury protection for paratroopers. Part II: Improved system description. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2005-05.

Velger, M. (1998). *Helmet-Mounted Displays and Sights*. Norwood, MA: Artech House, Inc.

Ververs, P.M., and Wickens, C.D. (1998). Head-up displays: effects of clutter, display intensity, and display location on pilot performance. *International Journal of Aviation Psychology*, 8(4), 377-403.

Vollmerhausen, R.H., Jacobs, E.L., and Driggers, R.G. (2003). New metric for predicting target acquisition performance. *Proceedings of the SPIE, Infrared Imaging Systems-Design, Analysis, Modeling, and Testing XIV,* 5076, 28-40.

Waldelöf, C., and Firberg, N. (1996). Integrated helmet system, the Swedish approach. *Proceedings of the SPIE, Head-Mounted Displays*, 2735, 245-253.

Watson, B.A., and Hodges, L.F. (1995). Using texture maps to correct for optical distortion in head-mounted displays. *Proceedings of the Virtual Reality Annual International Symposium (VRAIS ’95*), 172-178.

Weintraub, D.J. (1992). *Human Factors Issues in Head-Up Display Design: The Book of HUD.* Dayton, OH: Dayton University Research Institute.

Wells, M.J., Venturino, M., and Osgood, R.K. (1989). Effect of field of view size on performance at a simple simulated air-to-air mission. *Proceedings of the SPIE, Helmet-Mounted Displays,* 1116, 126.

Whitestone, J.J., and Robinette, K.M. (1997). Fitting to maximize performance of HMD systems. In: Melzer,J.E., and Moffitt, K.W. (Eds.), *Head-Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Wickens, C.D. (1997). Attentional issues in head-up displays. In: Harris, D. (Ed.), *Engineering Psychology and Cognitive Ergonomics: Transportation Systems (Vol. 1)*. Aldershot, UK: Ashgate.

Wickens, C.D., Fadden, S., Merwin, D., and Ververs, P.M. (1998). Cognitive factors in aviation display design. Savoy, IL: Aviation Research Laboratory, Illinois University.

Wood, R.B. (1992). Holographic head-up displays. In: Karim, M.A. (Ed.), *Electro-Optical Displays*. New York, Marcel Dekker, Inc.

Wood, R.B., and Howells, R. (2001) Head up displays. In: Spitzer, C.R. (Ed.), *The Avionics Handbook*. New York: CRC Press.

Vyrnwy-Jones, P., Lanoue, B., and Pritts, D. (1988). SPH-4 U.S .Army flight helmet performance, 1983-1987. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 88-15.

Yeh, M., Wickens, C.D., and Seagull, F.J. (1998). Effects of frame of reference and viewing condition attentional issues with helmet-mounted displays. Aviation Research Lab, Institute of Aviation, University of Illinois at Urbana-Champaign. Technical Report ARL-98-1/ARMY-FED-LAB1.

18 EXPLORING THE TACTILE MODALITY FOR HMDs[1]

Kimberly P. Myles
Mary S. Binseel

The Concept of Tactile Interfaces

Humans commonly are considered to have five separable named senses, called hearing, sight (vision), smell, taste, and touch, providing information about the external world. These senses, their organs, and respective modalities are listed in Table 18-1. The two most dominant senses are sight and hearing. Equipment designers understand human reliance on these two senses and have used them as the basis for most instruments and controls, as would be expected since humans manipulate and receive most feedback from their external environment based on what they see and hear. However, humans have a limited capacity to receive, hold in working memory, and cognitively process information taken from the environment through any particular sensory pathway. For operators of complex or multiple systems processing demands may be so large that the sole reliance on sight and hearing can lead to an overloading of these two sensory channels (Sorkin, 1987; van Veen and van Erp, 2001).

Table 18-1.
Five senses providing information about the surrounding environment (Modified from Silbernagel [1979]).

Sense	Sense Organ	Sensory Modality
Hearing	Ears	Auditory
Sight (Vision)	Eyes	Visual
Smell	Nose	Olfactory
Taste	Tongue	Gustatory
Touch	Skin	Tactile *

* Sense of touch also provides sensations of heat and pain.

When a single sensory channel is overloaded with information, and the user becomes incapable of processing all incoming information, it can result in a rapid increase in errors (Oviatt, 1999) and a decrease in situational awareness and overall user performance. One way to reduce the sensory overload is to deliver part of this information through unused or underutilized sensory modalities. A multimodal system using several sensory modalities to transmit information between the environment and the user will lessen the chance of any one sensory mode becoming overloaded (Oviatt, 1999; Wickens, 1984, 2002). Oviatt (1999) explains that the goal of multimodal systems should be to "integrate complementary modalities to yield a highly synergistic blend in which the strengths of each mode are capitalized upon and used to overcome weaknesses in the other" (p. 74). Thus, tactile displays take advantage of the sense of touch to distribute the cognitive workload among visual, auditory, and tactile sensory channels.

Smell and taste both involve the analysis of chemical molecules. Humans perceive odors via the sense of smell and the flavor of foodstuffs via the sense of taste. Smell, like vision and hearing, is a spatial telereceptor and has already been considered in virtual reality applications (Psotka, Division and Lewis, 1993). The tongue is rich with

---

[1] This chapter is an extended version of the ARL Technical Report: Myles, K. and Binseel, M.S. (2007). *The Tactile Modality: A Review of Tactile Sensitivity and Human Tactile Interfaces;* ARL-TR-4115. Army Research Laboratory: Aberdeen Proving Ground, MD.

receptors and can discriminate five distinct tastes (salt, sweet, sour, bitter, and umami), which guide our food preferences, especially away from potentially harmful substances (Smith and Margolskee, 2006). Taste has been also reported to be successfully used by the blind as a substitute sense for navigation in space. An array of electrodes placed on a tongue has been used to help blind people to navigate or to catch moving object. (Bach-y-Rita et al., 1998). However, both smell and taste pose challenges for use as informational interfaces. Therefore, of smell, taste, and touch; touch – specifically the cutaneous (related to skin) mechanical, aspect of touch called *taction* – is the most conducive to being used as a human system interface.

Touch is the sense by which external objects or forces are perceived through contact with the body (Stedman's Medical Dictionary, 2004). It is a part of the somatic (somatosensory) system that receives external and internal information about the state of the body. Somatic senses are divided into cutaneous senses (touch), proprioception, and visceral senses. Priopriocepion includes the vestibular system (sense of balance) and kinesthetic senses. Kinesthetic senses inform the brain about relative positions of various body parts and their movement. Visceral senses inform the brain about the state of internal organs and overall body condition (e.g., hunger, fatigue, stomach ache).

The main operational advantages of touch over smell and taste sensory channels are a large area of possible stimulation, relatively large dynamic range, and natural directional capabilities. Touch includes sensations of pressure, heat-and-cold, and pain. The sensation of pressure is referred to as taction or tactition (from Latin *tactus;* touched) or as the tactile modality. A hierarchical taxonomy of terms related to touch and taction is shown in Figure 18-1. The solid lines show the "branch" of taxonomy of interest in this chapter; branches along the dotted lines are informational and not exhaustive.

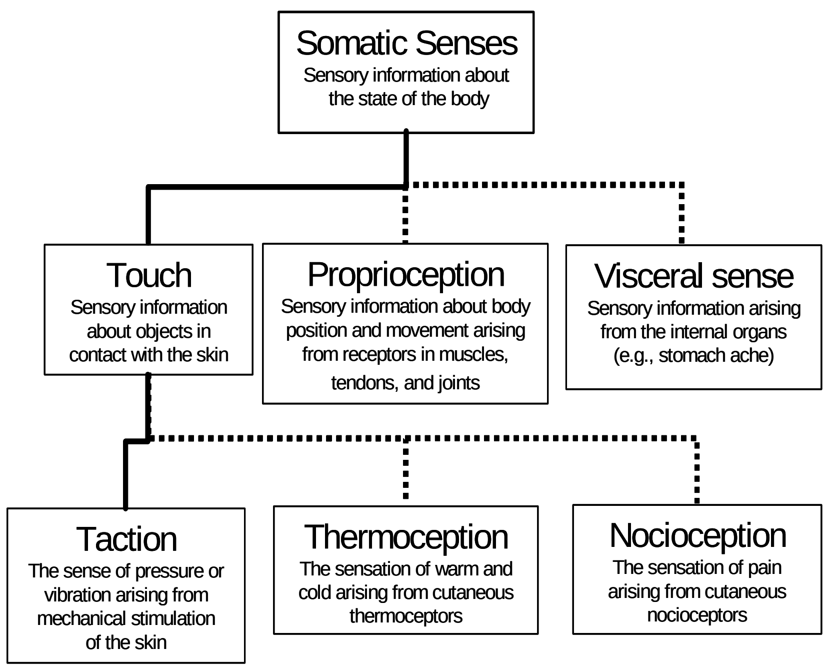


Figure 18-1. Graphical representation of a taxonomy related to touch.

One additional term that is frequently used in relation to touch is *haptics*. Most authors use the words *haptic* and *tactile* as synonymous (Stedman's Medical Dictionary, 2004). However, some authors consider haptic sensation as a combination of tactile and kinesthetic sensations resulting from the opposition to movement of the object touching the skin (Geiser, 1990; Youngblut et al., 1996). Therefore, haptic perception frequently is referred to as the perception of three-dimensional (3-D) objects (size, weight, temperature, texture, etc.) held in or being pushed by the hand. The latter meaning of the term *haptics* is used in this chapter.

As shown in Figure 18-1, the skin responds to several types of stimulation. Receptors in the skin can be stimulated by mechanical, thermal, electrical, and chemical means. However, the latter three types of stimulation are not good candidates for interfaces, since they produce sensations that are hard to precisely localize (e.g., Mauderli et al., 2003). In addition, there are obvious issues with undesirable outcomes of a thermal, electrical, or chemical stimulation, such as pain and dermal damage. Conversely, mechanical receptors have several properties that make them good receptors of communication signals. Mechanical stimulation of the skin may have two forms: *constant stimulation,* resulting in the sensation of skin deflection, and *variable stimulation,* resulting in the sensation of skin vibration (referred sometimes to as "prickling"). The former usually is referred to as tactile or pressure stimulation and the latter as vibrotactile stimulation. An example of a constant tactile signal is a modified computer mouse which raises a piston against the finger when the mouse is positioned over an on-screen button, giving tactile feedback regarding cursor location (Akamastu, MacKenzie and Hasbrouq, 1995).

Vibrotactile stimulation can be used to send various coded signals, since it can vary in frequency, intensity, and temporal pattern. The use of the tactile or vibrotactile interface as an information channel can be beneficial in multimodal systems, especially when the visual and/or auditory channels are heavily loaded (Gemperle, Ota and Siewiorek, 2001; Raj, Kass and Perry, 2000; van Erp, 2001). Vibrotactile displays also have been proven to help fill the communication gap when the visual and/or auditory sensory modalities are weakened (Raj et al., 2000; Schrope, 2001; van Erp and van Veen, 2001). Various types of vibrotactile displays have been used successfully in a number of applications: assistance for the blind, video games, human-machine interfaces, and virtual reality enhancements. An example of a common vibrotactile display is an alert device built into pagers and cellular phones for use when an auditory signal may disturb or alert others. In summary, the tactile modality is beneficial for transmitting environmental information to the user. However, in order to design a useful tactile or vibrotactile interface, the designer needs to understand the limitations of tactile stimulation and the boundaries of useful tactile parameters.

The Physiological Basis of Tactile Stimulation

Skin is a layer of cells that protect the tissue underneath and help to maintain body temperature. Human skin has an area of 1.8 meter^2 (m^2) (19.4 feet^2 [ft^2]), a density of 1250 kilogram/meter^3 (kg/m^3) (78 pounds [lb]/ft^3), and a weight of 5 kg (11 lbs) (Sherrick and Cholewiak, 1986). It is classified as either glabrous (non-hairy) skin, which is found only on the plantar and palmar surfaces, or hairy skin, which is found on the rest of the body. This division is relevant to tactile displays because these skin types differ in sensory receptor systems and tactile sensitivity (Cholewiak and Collins, 1995).

The skin has two primary layers called the epidermis (outer layer) and the dermis (inner layer). In the dermal layer and at the interface of the epidermis and dermis, there are many spatially distributed free nerve endings that collect and disburse information about objects coming in contact with the skin. These free nerve endings are slowly-adapting (SA) receptors that are sensitive to mechanical, thermal, electrical, and chemical energy and convert these to neural signals, producing sensations of heat or pain (Patestas and Gartner, 2006). In addition, groupings of fast acting hair follicle receptors surround skin hair follicles and respond to skin displacement near the base of skin hair when the hair is touched (Cholewiak and Collins, 1991; Sherrick and Cholewiak, 1986).

In addition to free nerve endings and hair follicle receptors, there are four specialized types of mechanoreceptors in the skin that respond to pressure and vibration. These four types of mechanoreceptors are Pacinian corpuscles, Merkel cells (or Merkel disks), Ruffini endings (or Ruffini corpuscles), and Meissner corpuscles. Each of these cell types has a specific sensory nerve channel associated with it called P (Pacinian channel), NPI (Non-Pacinian channel I), NPII (Non-Pacinian channel II), and NPIII (Non-Pacinian channel III), respectively (Bolanowski et al., 1988; Cholewiak and Collins, 1991; Klatzky and Lederman, 2002). The main difference between the channels is their frequency response. A summary of the basic properties of the four mechanoreceptors is shown in Table 18-2.

A cross-section of the skin showing the types and locations of the various mechanoreceptors is shown in Figure 18-2.

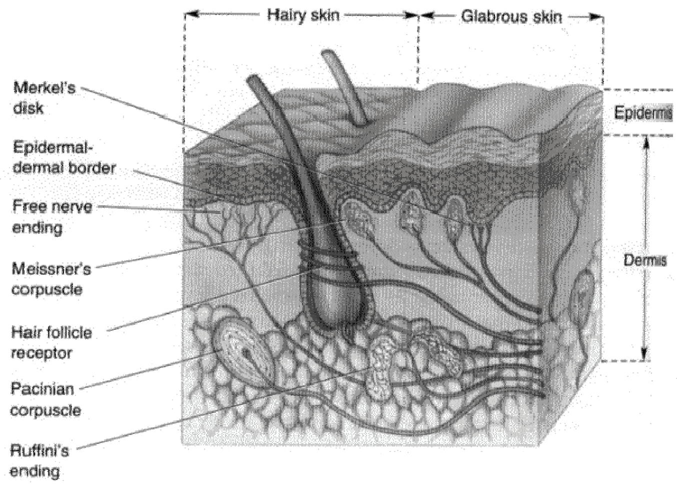


Figure 18-2. The somatosensory receptors of the skin (Kohler, 2001). Somatosensory senses (http://www.humboldt.edu/~jgk5/ cutaneous_senses.htm).

Each mechanoreceptor type (and associated neural channel) has a specific role in the perception of vibration which extends from almost 0 Hz to greater than 500 Hz (Bolanowski et al., 1988; Cholewiak, Collins and Brill, 2001; Gemperle et al., 2003). Some of them are SA cells (pressure detectors), which respond while the stimulus is present with no decrease in firing rate, whereas others are the rapidly-adapting (RA) (change detectors), which respond with bursts of firing in response to a change in stimulation. RA change detectors discharge when the sensory cell is compressed and again when the cell is restored to its resting state. SA sustained pressure detectors discharge when the cell is compressed and continue to discharge until the stimulus cease to act (Patestas and Gartner, 2006).

Pacinian corpuscles are the largest and the most sensitive receptors in the skin (Bear, Connors and Paradiso, 2006; Fig. 12.4, p. 391). They have an oval shape (up to 1 millimeter [mm] x 4 mm [0.04 inch [in] x 0.16 in]) and are located at a moderate depth in the skin (about 2 to 3 mm [0.08 x 0.12 in]) (Sherrick and Cholewiak, 1986). The sensitivity function of Pacinian corpuscles has a “U” shape with maximum sensitivity occurring in the 250 to 300 Hz range (Bolanowski et al., 1988; Lamore and Keemink, 1988; Verrillo, 1962, 1966). The Pacinian corpuscles are rapidly adapting and sensations for these receptors are described as deep and diffuse (Sherrick, Cholewiak and Collins, 1990). Pacinian corpuscles are located deeply in the dermis and have relatively large receptive fields, that is, regions over which skin stimulation excites a primary afferent fiber, of about 100 mm^2 (0.16 in^2) (Youngblut et al., 1996).

Merkel cells are slowly adapting cells, which are sensitive to constant pressure. The cells have a cross-section of 10 to 15 microns (μm), are located in the upper layers of the dermis, and have a small receptive field of about 10 mm^2 (0.016 in^2) (Youngblut et al., 1996). Meissner corpuscles are located just below the epidermis and are sensitive to low frequency vibration below 300 Hz (Sherrick, Cholewiak and Collins, 1990) but are most sensitive to frequencies of stimulation in the 20 to 50 Hz range. The Meissner corpuscles are rapidly adapting receptors having a small receptive field averaging 10 mm^2 (0.016 in^2) (Youngblut et al., 1996). Sensations for these receptors are felt as a gentle skin flutter, sometimes called the “flutter sense” (Sherrick, Cholewiak and Collins, 1990). The size of a Meissner corpuscle is between 30 to 140 μm in length and 20 to 60 μm in width (Guinard et

Table 18-2.
Characteristics of the four types of mechanoreceptors in the human skin.
(Bear et al., 2006; Bolanowski et al., 1988; van Erp and van den Dobbelsteen, 1998)

Location	**Type of Mechanoreceptor**	
	Rapidly-Adapting (RA) Receptors	Slowly-Adapting (SA) Receptors
Superficial skin (small receptive field, e.g., 1 to 3 mm [0.04 to 0.12 in])	Meissner corpuscle (RAI NPI) Operate in the 5-300 Hz range. U shaped sensitivity curve; most sensitive 20-50 Hertz (Hz). Sensitivity independent of temperature. No temporal summation. Cause “flutter” (prickling) sensation.	Merkel cell (SAI, NPIII) Operate in the 0.4-100 Hz range. Relatively flat sensitivity curve. Sensitivity slightly dependent on temperature. Sensitive to tactile form and roughness (texture).
Deeper tissue (large receptive field, e.g., 6-10 mm [0.24 to 0.4 in])	Pacinian corpuscle (RAII, P) Operate in the 10-1000 Hz range. U shaped sensitivity curve; most sensitive 250-300 Hz. Sensitivity very dependent on temperature. Spatial and temporal summation. Non-localized vibration sensation.	Ruffini ending (SAII, NPII) Operate in the 15-400 Hz range; at a much lower sensitivity. Relatively flat sensitivity curve. Sensitivity very dependent on temperature. Sustained pressure receptors. Skin stretch receptors.

al., 2000). They are particularly numerous on extremities but sparse on the skin of the back, and their number decreases with age.

Ruffini endings are slowly adapting receptors responding to constant pressure and very slow vibration. They are also thermoreceptors and are sensitive to directional skin stretch. Ruffini endings have relatively large receptive fields averaging about 60 mm^2 (0.09 in^2) (Youngblut et al., 1996).

As the four mechanoreceptors overlap in their absolute sensitivities and receptive fields, a complex or variable vibratory stimulation will seldom activate one receptor because the energy applied to the skin will move throughout nearby skin tissues (Sherrick and Cholewiak, 1986; van Erp and van den Dobbelsteen, 1998). When constant pressure is applied to the skin, the smallest absolute threshold of sensation is about 0.03 erg and the minimum noticeable difference in stimulus intensity is about 3% (Eysenck, Arnold and Meili, 1972).

Neural stimuli caused by tactile stimulation travel through ascending neural pathways of the dorsal root ganglion, medulla oblongata, and medial lemniscus; and enter the cerebral cortex at the ventral posterior nucleus of thalamus (Bear, Connors and Paradiso, 2006). The primary somatosensory areas of the brain are located in the parietal lobe (Kohler, 2001). The ascending pathways of the neural responses to touch are shown in Figure 18-3.

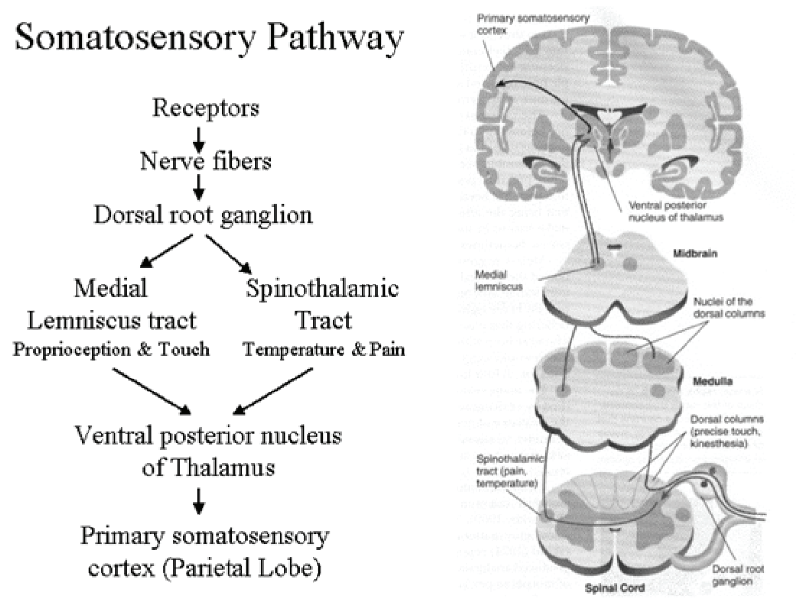


Figure 18-3. Basic ascending pathways of the touch neural impulses (Kohler, 2001) Somatosensory senses (http://www. humboldt.edu/ ~jgk5/cutaneous_senses.htm).

Elements of Vibrotactile Perception

Head-mounted tactile displays, especially vibrotactile displays, are relatively new and sparsely applied; therefore, much of the discussion of vibrotactile perception and interface design included in this chapter is based on knowledge gained from the tactile interfaces designed for other areas of the body. This knowledge nonetheless gives insight regarding vibrotactile sensitivities, spatial resolution, and other parameters of touch to be considered in designing helmet-mounted and head-mounted systems.

Weber's (1834/1978) and Weinstein's (1968) early research on tactile perception provides the basis for what is currently known about relative tactile sensitivity for various body sites. Basic parameters that must be considered in designing head-mounted vibrotactile systems include frequency- and location-dependent pressure sensitivities (vibrotactile thresholds), stimulation localization accuracy, and spatial resolution. Additionally, temporal resolution is of interest to designers so that signals will not be presented so closely in time as to be indistinguishable.

Sensitivity

Similar to the relationship found for the visual and auditory modalities, the threshold of vibrotactile sensation is inversely proportional to the amount of energy applied to the skin (Verrillo, 1966). However, skin sensitivity and mechanical impedance vary in different areas of the body due to differences in skin "thickness, vascularity, density, electrical conductivity, and more derived properties, such as moduli of shear and elasticity" (Sherrick and Cholewiak, 1986; Weber, 1834/1978). Skin vibrations are detected best on hairy, bony skin but are not detected as well on soft, fleshy areas of the body (Gemperle et al., 2003). This means that the head and scalp are parts of the body that are relatively sensitive to vibrotactile stimulation. In addition, skin sensitivity decreases as you move from distal to proximal portions of extremities (Sherrick, Cholewiak and Collins, 1990; Van Erp and van den Dobbelsteen, 1998; Wilska, 1954).

Weinstein (1968) investigated whether tactile sensitivity differed for gender and for the left and right sides of the body (for various locations on the body). He found that women were more sensitive than men to skin stimulation and that skin sensitivity was generally the same for both the left and right sides of the body. For specific body location, the forehead (face), trunk, and fingers were most sensitive and the lower extremities least sensitive to mechanical stimulation (Figures 18-4 and 18-5).

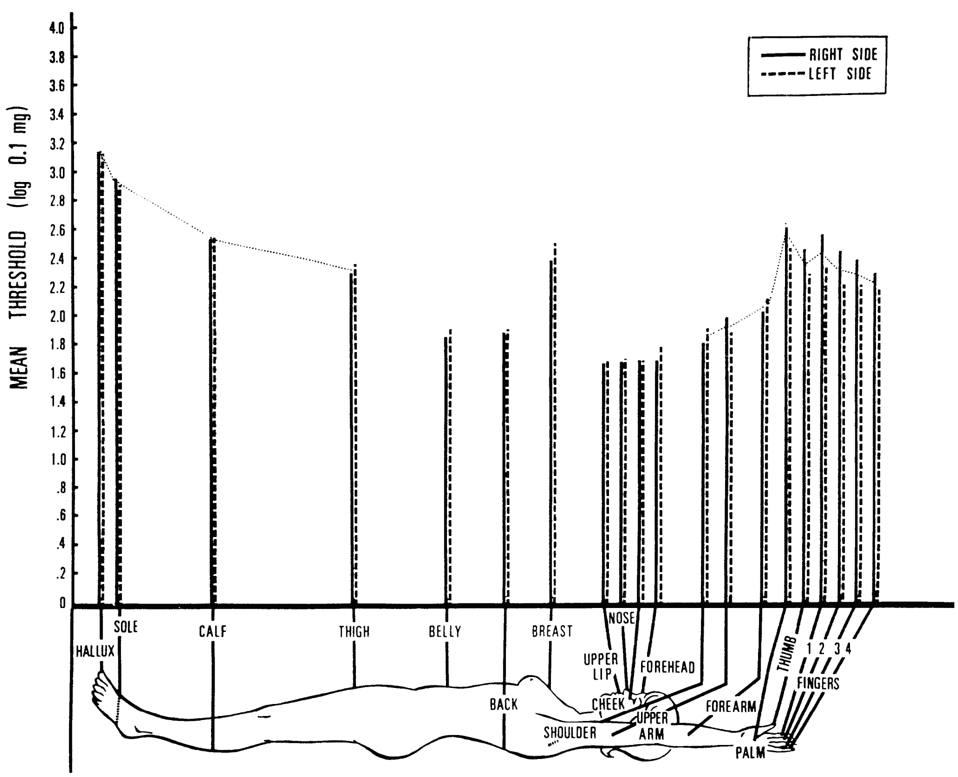


Figure 18-4. Pressure sensitivity thresholds for females for different areas of the body (Weinstein, 1968).

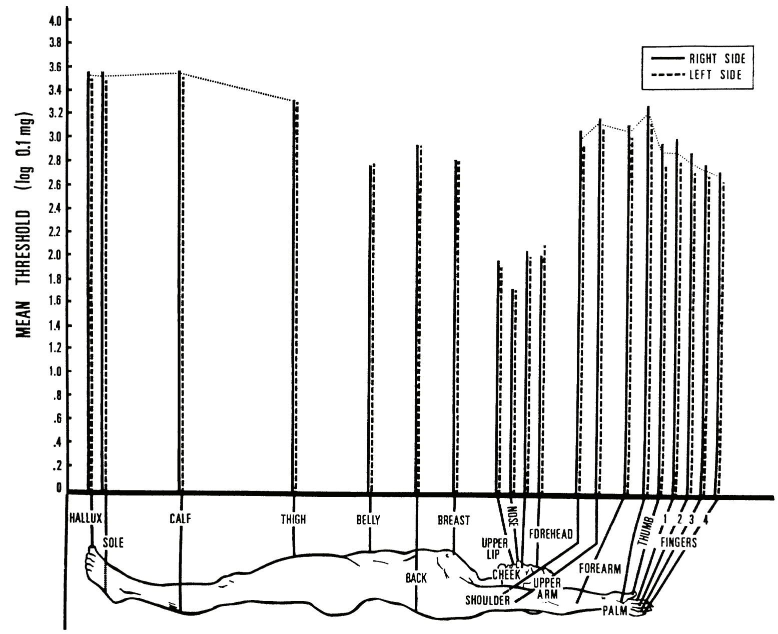


Figure 18-5. Pressure sensitivity thresholds for males for different areas of the body (Weinstein, 1968).

In an attempt to describe vibration sensitivity associated with different regions of the body, Wilska (1954) used a vibrator driven by a sinusoidal alternating current and placed it against the skin of various body regions. He found the hands and soles of the foot to be most sensitive and the gluteus region to be the least sensitive. Body sites including the head, throat, and abdomen were moderately sensitive in comparison to these endpoints. It appears to be no coincidence that most of the body sites involved in tactile parameter estimation in the literature are also those areas of the body that are most sensitive to pressure and stimulus discrimination [finger, Cholewiak and Collins, (1995); Cholewiak and Collins (1997); Goble, Collins and Cholewiak (1996); Horner, 1992; Lamore and Keemink, 1988; Rabinowitz et al., (1987); hand, Bolanowski et al., 1988; Cholewiak and Collins, (1995); Verrillo, 1962; arm, Cholewiak and Collins, 2003; Lamore and Keemink, 1988; Verrillo, 1966]. Some of the lesser-sensitive regions also investigated are the thigh (Cholewiak and Collins, 1995) and torso (Cholewiak, Brill and Schwab, 2004; Cholewiak, Collins and Brill, 2001).

Laidlaw and Hamilton (1937) also explored vibration thresholds for different regions of the body. They found significant variability in threshold measurements across participants within certain regions with specifically higher thresholds among the elderly and obese. These results are in agreement with others who also found an age-related increase in thresholds (Goble, Collins and Cholewiak, 1996; Stuart et al., 2003). For the older group of participants, Stuart et al. (2003) found an increase in threshold for the forearm, shoulder and cheek when compared to younger participants. However, thresholds for the finger were the same for both groups. This finding is not surprising since both Weber (1834/1978) and Weinstein (1968) found this area to be most sensitive to pressure and stimulus discrimination reflecting a high receptor density and making it more resistant to loss of sensitivity with age (Stuart et al., 2003).

Spatial resolution

Two-point discrimination is a measure that represents how far apart two pressure points must be before they are perceived as two distinct stimulation points on the skin (Gemperle et al., 2003). Information about spatial resolution is important for determining the minimum distance between two adjacent points of stimulation. If two tactors (mechanical transducers providing pulse, continuous, or vibrotactile stimuli) are placed too close together and each tactor delivers a unique signal in the scheme of some complex, tactile pattern, the observer will not differentiate between the signals and will miss the underlying message generated with the use of the two signals.

Weber (1834/1978) studied two-point discrimination thresholds for various areas of the body. Using a metal compass (dividers), he touched various areas of the skin with the two points some distance apart and recorded judgments of the distance between the two points. From his findings, Weber (1834/1978) put forth five general propositions, of which the first two stated: (1) various parts of the body are not equally sensitive to the spatial separation of two simultaneous points of contact; and (2) if two objects touch us simultaneously, we perceive their spatial separation more distinctly if they are oriented along the transverse rather than the longitudinal axis of the body. In order of decreasing sensitivity for two-point discrimination, the tongue was found to be most sensitive, followed by the lips, fingers/palm, toes, and forehead. More recently, Gemperle et al. (2003) found that two-point discrimination acuity is 39 mm (1.5 in) for the back, less than 1 mm (0.04 in) for the fingers, 15 mm (0.6 in) for the forehead, 35 mm (1.4 in) for the forearm and 45 mm (1.8 in) for the calf. These observations agree with an earlier report by Weinstein (1968) who found the fingers, forehead, and feet to be most sensitive for two-point discrimination (Figures 18-6 and 18-7).

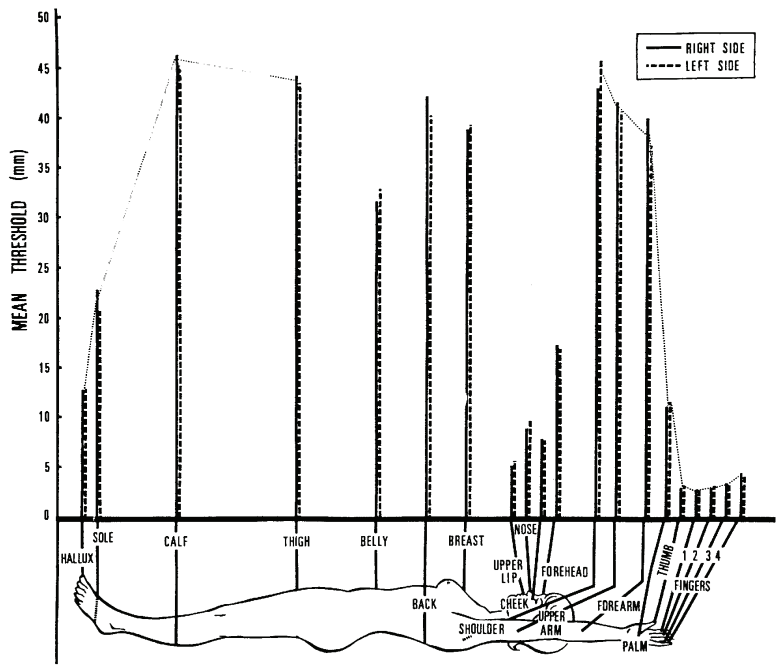


Figure 18-6. Two-point discrimination thresholds for females for different areas of the body (Weinstein, 1968).

Localization accuracy

Localization is defined in this chapter as the ability to identify where on the skin stimulation has occurred. Localization accuracy is typically measured by presenting a stimulus through one tactor among many present on the body site (such as the abdomen) and asking the experimental participant which tactor had been excited. Cholewiak, Brill and Schwab (2004) investigated the vibrotactile localization accuracy for the abdomen using 12, 8, and 6 equidistant tactors, 72 mm (2.8 in), 107 mm (4.2 in), and 140 mm (5.5 in) apart, respectively, arranged circumferentially around the body at abdominal height. They observed that localization accuracy increased as the number of possible locations decreased and reported 74, 92, and 97% identification accuracy for 12, 8, and 6 tactors, respectively. They also found that when the tactor placement included the navel and spine, localization was better than when the tactors were oriented with the navel and spine centered in a gap between tactors. This reflects other studies which indicate that the ability to localize improves when the stimuli are at or near body anchor points, such as joints (Cholewiak and Collins 2003, Weber 1826/1978). Cholewiak and Collins (2003) found that sites on the forearm near the elbow were better localized than those sites farther from the elbow. When increasing tactor spacing from 25 to 50 mm (0.98 to 1.97 in), localization accuracy for the forearm also increased. Hawes and Kumagai (2005) found that soldiers were able to localize an eight-tactor array around the head, with a mean distance of 7.1 centimeter [cm] (2.8 in) (center-to-center) between the tactors. Weinstein (1968) found that the forehead, the fingers, and hallux (big toe) were most sensitive for point localization (Figures 18-8 and 18-9).

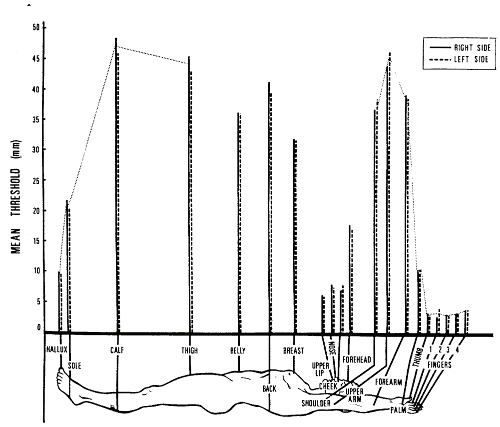


Figure 18-7. Two-point discrimination thresholds for males for different areas of the body (Weinstein, 1968).

Tactile Displays: Tactors

A tactile display consists of one or more tactors and a mounting structure or harness that positions the tactors on the appropriate part of the body. Important decisions that need to be made in designing tactile helmet-mounted display (HMD) systems or other types of tactile displays are the selection of type of tactile transducer and the number of transducers to be used in the design. Decisions regarding the type of tactor involve its size, weight, and power handling capabilities. Especially important are the geometrical properties of the element touching the skin. This element is commonly referred to as a contactor. It has been shown that contactor area and the diameter of the contactor can significantly affect tactile detection thresholds on the skin (Verrillo, 1962). To maintain a proper coupling between the contactor and the skin (especially bony parts of the body such as the skull) it is also important to provide sufficient static force pressing the contactor against the reception area.

In the case of multi-channel tactile interfaces presenting messages in a form of coded patterns, a larger number of tactors on the skin can increase the accuracy of the transmitted information (FreePatentsOnline.com, 2006). However, the designer must determine the maximum number of tactors beyond which no further increase in perceptual ability can be measured (FreePatentsOnline.com, 2006) as the placement of too many tactors can significantly distort the ability to discriminate between signals. In the case of single-channel tactile interfaces either a single tactor or multiple tactors can be used. Single tactor devices can be a simple on-off buzzer or a more complex device that provides a number of signals that carry various types of information. The latter device usually requires large, high powered tactors that are heavy and may become hot during some types of operations. Therefore, they may be replaced by a number of parallel transducers working simultaneously. These multiple tactor single-channel devices are usually used to provide directional information to the user. For such devices, several tactors are placed around the body and used sequentially to indicate specific direction. There are four types of electromechanical transducers that are currently used as tactors: moving coils (magnetoelectric, dynamic) transducers, DC motors with an eccentric weight (e.g., cell phone technology), piezoelectric transducers, and electro-pneumatic transducers (van Erp, 2002).

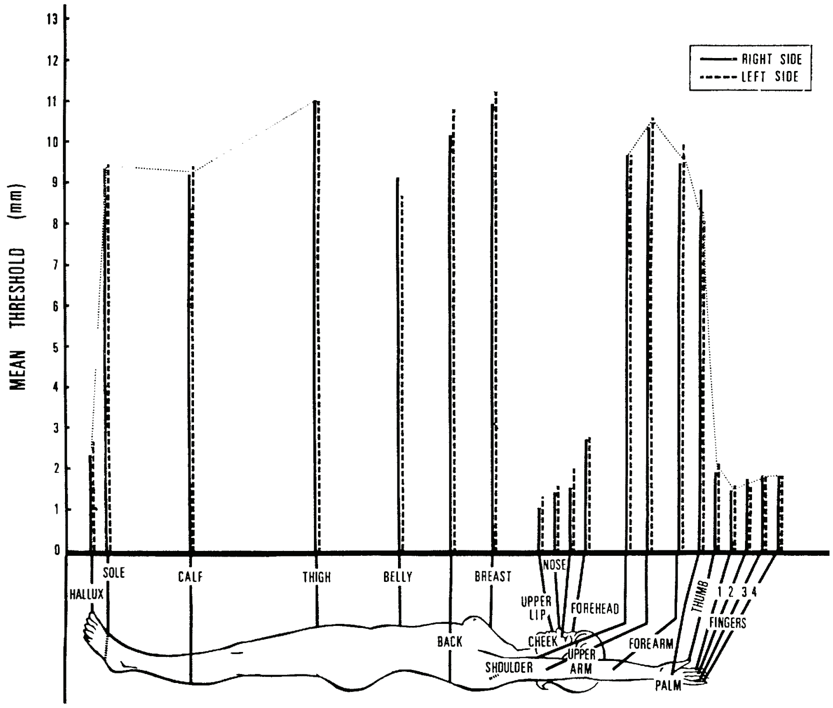


Figure 18-8 Point localization thresholds for females (Weinstein, 1968).

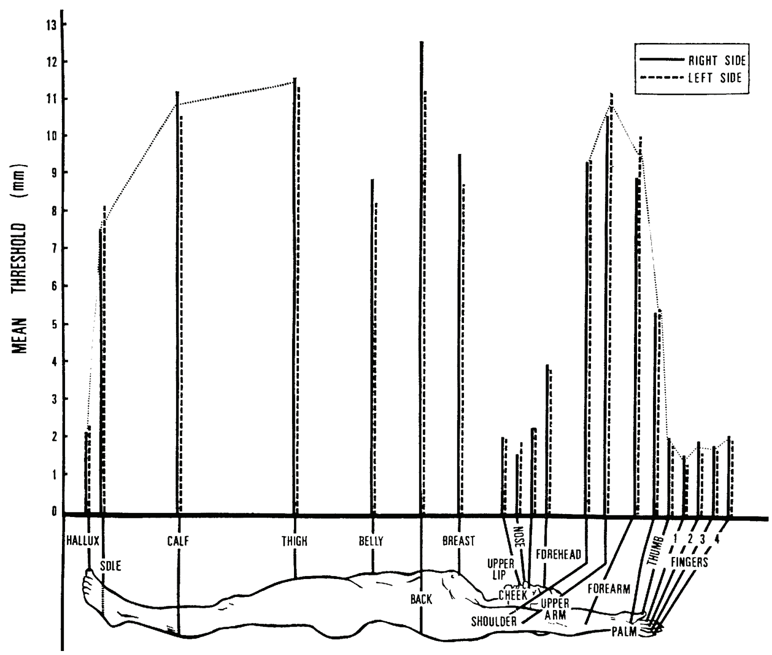


Figure 18-9 Point localization thresholds for males (Weinstein, 1968).

Moving coil transducers

A moving coil transducer is an electromechanic transducer with a stationary magnet and a moving wire coil passing an alternate current. Such transducers are also known as magnetoelectric or dynamic transducers. (Tran, Amrein and Letowski, 2009). An example of a moving coil tactor is the C-2 tactor designed by Engineering Acoustics, Inc. (EAI) (www.eaiinfo.com) and shown in Figure 18-10. The tactor has a moving piston-like element with an attached electric coil. Current passing through the coil creates the contactor movement with displacement proportional to the intensity of the electrical current passing through the coil (Engineering Acoustics, Inc. [EAI], 2006; van Erp, 2002). The contactor displaces the skin with movement similar to a constant pricking or tapping of the skin while additional housing of the contactor shields the surrounding skin from vibration. The housing is used to keep the tactile signal as localized as possible and prevent the signal from radiating to unintended skin surfaces.

Figure 18-10. The C-2 Tactor (EAI) (http://www.eaiinfo.com/EAI2004/Tactor%20Products.htm), 2006.

Direct current (DC) motors

Direct current (DC) motor technology is used in tactors built in cell phones and pagers to alert the user via vibration of an incoming call, text message, calendar alert, etc. The motor produces vibration by rotating an off-center (i.e., eccentric) mass. The rotating mass “creates a centrifugal force that is transmitted through the entire motor as a vibration” (Gemperle, Ota and Siewiorek, 2001). An increase in DC voltage applied to the motor produces an increase in vibration intensity (Cohen et al., 2005; Gemperle et al., 2001). There are two commonly used DC motors: the cylindrical motor and the disk-shaped pancake motor (Figure 18-11). Both motors use an off-center mass to produce vibration. The pancake motor rotates the mass in a plane parallel to the mounting surface and the cylindrical motor rotates the mass in a plane normal to the mounting surface (Piateski and Jones, 2005). The cylindrical motor was found to provide better tactile pattern recognition and to be more reliable than the pancake motor (Bloomfield and Badler, 2006; Piateski and Jones, 2005). However, there are several limitations to using both types of the DC motor to drive the tactors. One limitation is that the off-center mass on the motors is easily impeded by minimal resistance or pressure from fingers or fabric (Gemperle et al., 2001). A second limitation is the tendency for the vibration frequency to be affected by extraneous factors such as how the tactor is mounted and posture changes (Cohen et al., 2005).

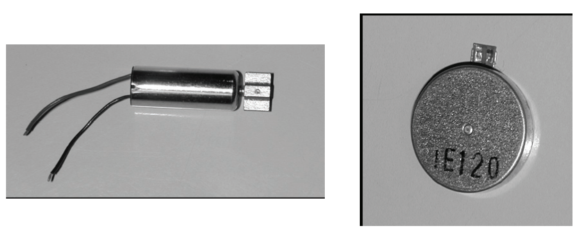


Figure 18-11 The cylindrical (left) and pancake (right) motors often used in cell phones and pagers (http://www.cis.upenn.edu/ ~aaronb/docs/tactorsuit.pdf).

Piezoelectric transducers

Piezoelectric transducers, also called piezoelectric benders, are electromechanical transducers which convert electrical power into mechanical vibration by bending a piezoelectric bimorph (Andersen, 2002; Chilibon et al., 2005). An example of a piezoelectric transducer is shown in Figure 18-12.

The piezoelectric effect is a property of certain crystals to produce static electricity in response to mechanical force (stress) (Andersen, 2002; Phillips, 2000). The effect is reversible. Typical piezoelectric transducers operate when two rectangular piezoelectric plates (and a metallic electrode fixed between them) are glued together and applied pressure (i.e., rubbing) drives one plate to expand while the other plate contracts, forcing the transducer to bend (i.e., deformation), thus creating an out-of-plane motion and vibrations in the range of tens of micrometers (Chilibon et al., 2005).

Figure 18-12. Piezoelectric bender actuators (http://www.physikinstrumente.com/en/products/prdetail.php?&sortnr= 103000).

The piezoelectric benders used for vibrotactile applications are usually not made with crystals but with more effective ceramic materials such as lead zirconium titanate (PZT). The configuration of a piezoelectric transducer can vary depending on the application (GlobalSpec.com, 2006), so these transducers can be designed to fit the system requirements. Other advantages to using piezoelectric bender transducers include: “the ability to generate electrical signals from mechanical and acoustic sources of low impedance”; and “the ability to develop relatively large motions and low forces with small electrical excitation” (Chilibon et al., 2005). Limitations to using piezoelectric bender transducers include the brittleness of the ceramic, which make it prone to breakage over time (Andersen, 2002; Niezrecki et al., 2001), and the unwanted change in displacement over time which may hinder the accuracy of the transducer (Andersen, 2002).

Electro-pneumatic transducers

Electro-pneumatic transducers use air pressure to generate a vibrating sensation on the skin (Figure 18-13). They use devices such as small air jets or air bladders to convert air pressure changes to vibration. When the air jets or bladders are activated by a pneumatic pump the resulting sensation on the skin is perceived as a touch (Enriquez et al., 2001; FreePatentsOnline.com, 2006). Contrary to DC motors, pneumatic jets and bladders do not shake the entire transducer but produce stimulation only within a specific area under the transducer (Enriquez et al., 2001). Drawbacks of electro-pneumatic transducers include possible air leaks in the equipment and the limited range of frequencies available for use (Enriquez et al., 2001). Also, the mechanical aspects of the system (pumps, valves, etc.), mean “all pneumatic tactile systems have an inherently slow response time, which limits the operating bandwidth of these devices, and hence the types of signals that can be sent to the user” (FreePatentsOnline.com, 2006).

Information about selected tactors commercially available at the time of publication is included in Table 18-3. Recently Mortimer, Zets and Cholewiak (2007) described a new class of tactors, which performance is fairly independent of the static pressure acting on the skin.

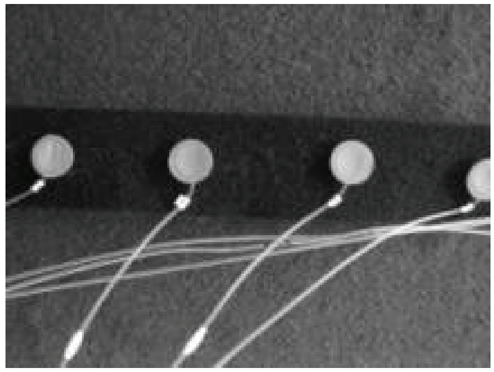


Figure 18-13. Pneumatic tactors (http://www.tactileresearch.org/rcholewi/TRLTactorArrays.html).

General recommendations for using tactors as communication devices were published by van Erp (2002). The author focused on single tactor displays such as those used in mobile phones and computer mice and multiple-element tactor displays worn on the body or used as finger displays. Examples of the later systems are tactile Braille finger displays. van Erp (2002) based his recommendations on broad neurophysiologic and psychophysical data available for human tactile perception. Presented recommendations were divided into stimulus detection guidelines and tactile information coding guidelines. For optimal stimulus detection van Erp recommended tactor placement on glabrous as opposed to hairy skin, 200 to 250 Hz frequency range, and long stimulus duration for frequencies above 60 Hz. For tactile information coding van Erp recommended up to four levels of intensity coding, up to nine levels of frequency, and at least 4 cm (1.6 in) separation between multiple tactors. The difference between the frequencies of the tone signals should exceed 20% and the minimal duration of the signals and the pauses between the signals should exceed 10 milliseconds (ms). However, it has to be stressed that both sets of the above recommendations apply only to body-worn and hand-held tactile displays and are not appropriate for tactile HMD systems where both tactile perception and auditory perception via bone conduction need to be considered together. Tactile signals used in HMDs are discussed in the final part of the chapter.

Tactile Interfaces: Applications

A sensory channel can be used to substitute, reinforce, or add other sensory channels in providing information to the user. Examples include a scanner and pad which convert written material to Braille (substitution); a threat indicator that shows a target on a visual display and also alerts the user via an auditory display (reinforcement); or an auditory alarm for an equipment malfunction (additional channel).

The tactile modality can be used as an operational interface in all these ways. It can be used as an additional, independent input modality to convey information to the user or as a redundant modality to increase information salience of the visual and auditory modalities (Sherrick and Cholewiak, 1986; Sorkin, 1987). For visual- and hearing-impaired users the tactile modality can be used as a substitute channel and can become either the primary or a supplementary channel for the receipt of information. Outside of the visually impaired population, the military has been one of the leading pioneers in the development and use of tactile systems. In military applications, the vibrotactile channel is being used to deliver threat warnings and as an additional sensory input

Table 18-3.
Parameters of selected commercially available tactors.

Tactor/Transducer	Specifications	Cost	Use
Tactors			
 C-2 Tactor www.eaiinfo.com/EAI2004/Tactor%20Products.htm	• Diameter = 30.5 mm (1.2 in) • Height = 7.9 mm (0.31 in) • Weight = 17 grams (0.6 ounces [oz]) • Stimulus Amplitude = > 0.64 mm (> 0.025 in) peak at 230 Hz with 0.25 A RMS drive • Impedance = approx. 7 ohms	On request	• Mount directly on the skin, in a chair, or in clothing • Supplement audio or visual input • Place tactors individually, sequentially, or in groups • Have been used in military, medical, and commercial applications
 VPM2 Vibrating Disk Motor www.solarbotics.com/products/index.php?search_id=1056	• Diameter = 12.7 mm (0.5 in) • Height = 3.4 mm (0.134 in) • Standard Voltage: DC 3.0V • Operating Voltage Gange: DC 2.5 to 3.5V • Power Supply, Voltage Source: DC Power Supply or Battery 3.0V	$3.95	• Mount on skin • Mount in clothing • Place tactors individually, sequentially, or in groups

Table 18-3 (continued).
Parameters of selected commercially available tactors.

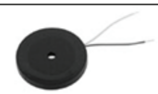
Tactor/Transducer	Specifications	Cost	Use
Tactors			
 Encased Piezo Element www.allelectronics.com/cgi-bin/item/PE-54/search/0.89"#34;_DIA._X_0.13"#34;_ENCASED_PIEZO_ELEMENT_.html	• Diameter = 22.6 mm (0.89 in) • Height = 3.3 mm (0.13 in)	$0.75	• Mount on skin • Mount in clothing • Place tactors individually, sequentially, or in groups
 Coin/Pancake Vibration Motors www.vibratormotor.com/guide-table.htm	• Diameter = 10 mm (0.4 in) • Length = 3 mm (0.12 in) • Frequency Range: 10 to 55 Hz • Operating Voltage Range: 2.5 to 4V DC • Operating Voltage: 3V DC	On request	• Mount directly on the skin, in a chair, or in clothing • Place tactors individually, sequentially, or in groups

Table 18-3 (continued).
Parameters of selected commercially available tactors.

Tactor/Transducer	Specifications	Cost	Use
Vibrotactile Transducers			
 Silver TST229 www.clarksynthesis.com/home-products.php	• Diameter = 20.3 cm (8 in) • Height = 5.7 cm (2.25 in) • Weight = 0.57 kg (20 oz) • Tactile Freq. Response = 15 Hz to 800 Hz • Maximum Power = 200W • Tactile Force Peak = 216 lb-ft • Transduction Force = 1.6 lb-ft/watt • Impedance = 4 ohms	$300	• Home theater • Commercial theater • Virtual reality • Gaming • Amusement parks • Hearing impaired • Studio monitoring • Stage monitoring • Simulators
 Gold TST329 www.clarksynthesis.com/home-products.php	Diameter = 20.3 cm (8 in) Height = 5.7 cm (2.25 in) Weight = 0.57 kg (20 oz) Tactile Freq. Response = 10 Hz to 800 Hz Maximum Power = 200W Tactile Force Peak = 378 lb-ft Transduction Force = 2.8 lb-ft/watt Impedance = 4 ohms	$500	• Home theater • Commercial theater • Virtual reality • Gaming • Amusement parks • Hearing impaired • Studio monitoring • Stage monitoring • Simulators

Table 18-3 (continued).
Parameters of selected commercially available tactors.

Tactor/Transducer	Specifications	Cost	Use
Vibrotactile Transducers			
 Platinum TST429 www.clarksynthesis.com/home-products.php	Diameter = 20.3 cm (8 in) Height = 5.7 cm (2.25 in) Weight = 0.57 kg (20 oz) Tactile Freq. Response = 5 Hz to 800 Hz Tactile Force Peak = 932 lb-ft Transduction Force = 6.9 lb-ft/watt Impedance = 4 ohms	$700	• Home theater • Commercial theater • Virtual reality • Gaming • Amusement parks • Hearing impaired • Studio monitoring • Stage monitoring • Simulators
 VX-GH72 www.vidsonix.com/vx3/vx_inwall.html	• Diameter = 7.6 cm (3 in) • Height = 2.54 cm (1 in) • Weight = 0.5 kg (1.2 lbs) • Freq. Response = 20 Hz to 20,000 Hz • Maximum Power = 50W • Impedance = 8 ohms	$69.95	• Tactile therapy • Home theater • Museum exhibits • Spas • Bathtubs • Pools ** Device can be placed in water

Table 18-3 (continued).
Parameters of selected commercially available tactors.

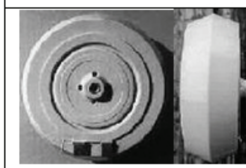
Tactor/Transducer	Specifications	Cost	Use
 Rolen Star Audio Transducer www.invisiblestereo.com	• Diameter = 10.2 cm (4 in) • Height = 4.4 cm (1.75 in) • Weight = 1 kg (2.2 lbs) • Freq. Response = 20 Hz to 20 kHz • Maximum Power = 100W • Impedance = 8 ohms	On request	Home Theater - Mount on chairs, recliners, love seats, small couches
 Aura Interactor Vest www.combatsim.com/htm/mar99/aura2.htm www.avsim.com/pages/0604/aura/aura_interactor.htm migman.com/hw/sound/AuraInteractor/AuraInteractor.htm	• Diameter = 7.6 cm (3 in) • Height = 1.7 cm 0.6875 in. • Maximum Power = 25W • Impedance = 3 - 4 ohms	$150	• Video game market Aura no longer offers this vest but it can be purchased from other sources over the internet.

Table 18-3 (continued).
Parameters of selected commercially available tactors.

Tactor/Transducer	Specifications	Cost	Use
 Crowson Tech TES 100 www.crowsontech.com/go/crowsontech/3342/en-US/DesktopDefault.aspx	Size = 14.5 cm x 12.2 cm x 2.8 cm (5.7 in x 4.8 in x 1.1 in) Weight = 1.6 kg (3.5 lbs) Freq. Response = 1 Hz to 500 Hz Minimum Power = 50W Maximum Power = 500W Impedance = 6 ohms	$349	Home Theater - Mount on chairs, recliners, love seats, small couches
 Aura Pro Bass Shaker www.partsexpress.com/pe/showdetl.cfm?&Partnumber=299-028	Size = 15.7 cm dia x 6.4 cm H (6.2 in dia X 2.5 in H) Weight = 1.7 kg (3.75 lbs) Freq. Response = 20 to 80 Hz Power = 50W RMS/100W max. Force Peak = 30 lbs. per ft. Impedance = 4 ohms	$38.88	- Enable movement from music, movies, video games - Video game chairs

facilitating human orientation, navigation, and communication capabilities (Castle and Dobbins, 2006; Chaisson, McGrath and Rupert, 2002). Some of the specific applications are discussed below.

Spatial orientation

Several tactile vests and belts have been developed in various countries to enhance spatial orientation under adverse operational conditions. Examples of such systems include TNO Tactile Torso Display (van Erp et al., 2003), Carnegie Mellon University (CMU) Wearable Tactile Display (Gemperle at al., 2001), and MIT Wireless Tactile Control Unit (Jones, Nakamura and Lockyer, 2004). The most widely known tactile vest is the Tactile Situation Awareness System (TSAS). The TSAS was developed at the U.S. Naval Aerospace Medical Research Laboratory (NAMRL) to minimize the occurrence of spatial disorientation in rotary-wing pilots, thereby reducing aircraft mishaps (Griffin, Pera, Cabrera and Moore, 2001; McGrath et al., 2004; Nordwall, 2000). The TSAS also helped to ease the visual overload naturally placed on pilots from the visual instruments in the aircraft. The TSAS is a vest filled with 32 tactors, worn on the torso of the pilot and assists the pilot in determining the aircraft's orientation with respect to the ground (Ryan, 2000). The location of a vibration on the torso directly relates to out-of-envelope excursions in aircraft attitude where corrective action is required (Schrope, 2001). For example, a vibration signal applied to the front of the torso indicates a correction is needed to raise the nose of the aircraft (Schrope, 2001). The system has been shown to increase pilot performance over a visual cockpit indicator alone. One pilot even wore the vest while blindfolded with no significant degradation in flight performance (Ryan, 2000). The TSAS confirms the efficient use of tactile systems for the orientation domain.

The success of the TSAS motivated its developers to expand the TSAS tactile vest concept to other applications. The U.S. Navy SEALs have shown interest in the system for use underwater for swimmer navigation and reduction of spatial disorientation, especially at night (Castle and Dobbins, 2006). The system has also been implemented as the Tactor Locator System (TLS) to reduce spatial disorientation for astronauts in the International Space Station (Rochlis and Newman, 2000) and as a ground navigation aid in the Tactile Situation Awareness System for Special Forces (TSAS-SF) (Chiasson, McGrath and Rupert, 2002).

Navigation aids

The TSAS has been applied in air (parachutist), land (dismounted), as well as underwater (diver) navigation (McTrusty and Walters, 1997; Chiasson, McGrath and Rupert, 2002). The results indicated that overall, using tactile feedback for navigation was feasible and beneficial. As an indication of the visual load during land navigation using the TSAS-SF, participants were asked to locate objects placed along the navigation path. Participants located about 80% more objects navigating with the TSAS-SF versus using a Global Positioning System (GPS) with a visual display. A similar system is a tactile belt for the torso, designed for infantry Soldiers to aid in navigation on the battlefield (Elliott et al., 2006; Krausman and White, 2006; Redden et al., 2006).

Researchers from the US Army Natick Soldier Research, Development, and Engineering Center (NSRDEC) and the US Army Research Institute of Environmental Medicine (Mahoney et al., 2007) have examined the effects of movement and physical exertion on vigilance in navigation tasks. Investigators used the tactile modality as a secondary communication source to the visual and auditory modes of communication. Results showed that while traversing a course with obstacles, participants covered less distance when responding to tactile signals than auditory signals.

Command and control

In the two applications discussed above, the tactile inputs convey direct physical information. Tactile displays can also be used to impart more abstract information such as in command-and-control applications. Merlo et al. (2006) reported a study in which four signals ("halt," "rally," "move out," and "nuclear, biological, or chemical warning") were presented three ways: hand signals initiated from the front of the group; hand signals initiated from the rear; and transmitted via a vibrotactile display. The vibrotactile display used was an eight-tactor torso belt worn just above the navel called Tactile Communication System (TACTICS) (Brill et al., 2006). During the study, Soldiers conducted individual movement techniques over obstacles, in many different postures, and with and without combat loads. Results showed faster detection of and response to signals when they were transmitted via the vibrotactile display. Soldiers commented that they preferred the vibrotactile display because it allowed them to use their vision to maintain their situational awareness without the need to frequently check their leader for hand signals. Although four signals were used, the amount of abstract information that can be imparted via vibrotactile could be increased by varying factors such as combinations and patterns of activated tactors, tactor locations, and signal frequency and duration.

Task reinforcement

Tactile inputs can be used to enhance performance for many tasks by reinforcement of other modalities. Akamastu, MacKenzie and Hasbrouq (1995) showed the advantage of incorporating tactile feedback when they asked participants to locate a target using a mouse-type device and to move the cursor inside the target. After the initial visual presentation of the target, participants were given auditory feedback, tactile feedback, color (visual) feedback, combined feedback, or no feedback to alert them that the cursor was placed inside the target. The authors found that the time required to correctly position the cursor was lowest for tactile feedback, showing that the addition of tactile feedback yielded a quicker motor response than other feedback systems for the task.

Tactile Head/Helmet-Mounted Displays (HMDs)

The tactile displays discussed thus far were primarily used for torso, arm, or hand applications. However, the hands are often occupied or unsuitable for use with tactile displays. The arms or the torso as locations for tactile displays have their own limitations, such as display size, bulkiness, thermal comfort, and compatibility with equipment, such as body armor, which can degrade the utility of these displays. Most importantly, the mental mapping of tactile signals can be impacted by head orientation when the display is mounted on the torso. Ho and Spence (2007) found that when the head is not aligned with the body, the perception of the location of a tactile signal is negatively affected. Therefore, in situations where the wearer is actively looking around or when the head is not aligned with the body, applications such as navigation or target cuing could suffer. This makes the head a location of choice for tactile displays aiding in navigation or providing directional information about the environment (e.g., sniper detection).

Tactors mounted on the head can be worn on headband, harnesses, such as used on other body locations, or incorporated into many types of headgear, including helmets. This would eliminate many of the potential problems encountered with torso-mounted displays. Therefore, tactile HMD systems should be considered in applications such as navigation and target/threat cueing, where mental mapping of the stimulus to the physical world is important and head-direction errors undesired.

An analysis of Weber's (1834/1878) data related to head sensitivity indicates that (1) various points on the head differ in tactile sensitivity, (2) the crown is less sensitive than the skin near the forehead, temples, and lower part of the back of the head, (3) spatial resolution is less for locations leading downward from the crown than for areas around the crown, and (4) forehead, and temples are best for tactile acuity. Gilliland and Schleger (1994) used various numbers of tactors (n = 5, 6, 8, 10, and 12) placed over the parietal meridian of the head (i.e., from ear to ear) and investigated tactile detection for a stimulus pulsing at a rate of 4 Hz. They reported that optimal tactile detection and localization accuracy occurred with the use of five tactors. As the number of tactors increased, localization accuracy decreased and reaction time increased.

Despite the many advantages of placing tactile displays on the head, examples of tactile HMD systems are still elusive. Borg, Neovius and Kjellander (2001) used three microphones and four transducers mounted in glasses to provide directional information about sound source (talker) location to hearing impaired and deaf-blind people. Mounting direction sensitive tactor arrays on the head allows the user to quickly orient the head toward the incoming sound source. Cassinelli, Reynolds and Ishikawa (2006) report on a pilot study of an "artificially extended skin" concept they call "haptic radar". In this experiment, six tactors were mounted along the rear hemisphere of the subjects' heads at 30° increments. Collocated with the tactors were infrared proximity sensors with a range of about 80 cm. When the proximity sensors detected an object, a vibrotactile signal proportional to the object's distance was applied to the associated tactor (the closer the object, the more intense the vibration). The experimenters then swung a foam ball at the back of the blindfolded subject's head. The subjects were significantly successful in moving in response to the stimulus; however, they were not significantly better at avoiding contact, as compared to their performance with the system off. These early results show promise for the integration of sensors and tactile displays, and the intuitive response to vibrotactile signals felt on the head.

Another tactile HMD system involved a navigation task (Hawes and Kumagai, 2005). The authors compared the utility of three types of vibrotactile displays: an eight-tactor head-mounted display, a four-tactor head-mounted display, and an eight-tactor chest-mounted display. All of the displays were mounted on circumferential bands with the tactors being placed at essentially equal intervals around the band. The results demonstrated better task performance with the eight-tactor head-mounted variant than the other two displays. In comparing the displays, the soldier participants rated the head- and chest-mounted variants similar in many subjective areas such as ease of use.

There are a number of issues which need to be investigated before a robust understanding of appropriate applications for head-mounted tactile displays is developed. For example, Hawes and Kumagai (2005) reported that even though head-mounted tactors produced better performance in a group of soldiers on a navigation task, and the soldier participants rated the head- and chest-mounted systems similar in many subjective areas such as ease of use, the soldiers showed a preference for the chest-mounted system. In the discussion of the results, the authors note (p. 49):

> "The participants found the vibration of the tactors was too strong on the head compared to the chest. Two participants reported getting headaches and the majority of the soldiers felt the system was too distracting when worn on the head. They reported that there currently tends to be too much information and equipment coming in through the head."

However, reported poor satisfaction with tactile HMD systems were most likely related to some suboptimal conditions of the study such as mounting bands that were too tight or presented signals which were too high in intensity. The main factor, which might have had a large contribution to any dissatisfaction, was that the tactile frequency used was 160 Hz. At this frequency bone conduction response is very strong and completely masks the presence of the cutaneous response on the skin. It has to be stressed that for tactile stimuli with frequencies above 60 Hz cutaneous perception through the skin occurs together with auditory perception through bone conduction pathways. This may or may not be a desirable situation. Since bone conduction perception is more effective than cutaneous perception for higher tactile frequencies, it can mask cutaneous response of the skin. For a tactile HMD system to provide tactile information in the auditory range the system must overcome the masking effect of bone conduction transmission, which may lead to prohibitively large and potentially dangerous tactile stimulation of the head.

Current research in tactile HMD systems is geared toward determining the optimum operational parameters for tactor placement and signal intensity and frequency. One of the projects conducted at the U.S. Army Research Laboratory (ARL) is to determine the optimum synergy between tactile and bone conduction signal reception using the same array of transducers. The concept of this system is an auditory-tactile cueing system using a circumferential tactor display which can also be utilized as a bone conduction communications headset. Early results show that for frequencies above 100 Hz, the bone-conducted sound component resulting from use of vibrotactile tactors is too strong to allow the use of frequencies in that regime for tactile purposes. It appears that the optimum tactile frequency range for head-mounted tactile displays is between 20 to 60 Hz and the shape of the tactile stimulus should have slow on and off transients to prevent generation of auditorily perceived clicks (Kalb, Amrein and Myles, 2008).

In conclusion, head mounted tactile displays offer promise in many single and multi-modal configurations for both civilian and military applications. Recent reports by Kalb, Amrein and Myles (2008) and Myles and Kalb (2009) support the use of such displays for sniper detection and tactical signal displays. By using tactile HMD systems, advantages in equipment compatibility, natural directional cueing, increased situational awareness, and integration of communications and informational displays can be achieved. With the recent explosion in research on tactile displays in general and in head-mounted displays in particular, the promise of these displays may soon be realized.

References

Akamastu, M., MacKenzie, I.S., and Hasbrouq, T. (1995). A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device. *Ergonomics.* 38, 816-827.

Andersen, B. (2002). Piezoelectric actuators move precisely and powerfully. Laser Focus World. September. http://lfw.pennnet. com/Articles/Article_Display.cfm?Article_D=154727&pc=gls

Bach-y-Rita P., Kaczmarek K., Tyler, M., and Garcia-Lara, J. (1998). Form perception with 49-point electrotactile stimulus array on the tongue: A technical note. *Journal of Rehabilitation Research and Development.* 45(4), 427-430.

Bear, M.F., Connors, B.W., and Paradiso, M.A. (2006). *Neuroscience: Exploring the Brain* (5th Ed.). Baltimore, MD: Lippincott, Williams, and Wilkins.

Bloomfield, A., and Badler, N.I. (2006). A Low Cost Tactor Suit for Vibrotactile Feedback. http://www.cis.upenn.edu/ ~aaronb/ docs/tactorsuit.pdf

Bolanowski, S.J., Gescheider, G.A., Verrillo, R.T. and Chechosky, C.M. (1988). Four channels mediate the mechanical aspects of touch. *Journal of the Acoustical Society of America,* 84(5), 1680-1694.

Borg, E., Neovius, L, and Kjellander, M. (2001). A three microphone-system for real-time directional analysis; toward a device for environmental monitoring in deaf-blind. *Journal of Rehabilitation Research and development*, 38(2), 265-272.

Brill, J.C., Terrence, P.I., Stafford, S., and Gilson, R.D. (2006). A wireless tactile communication system for conveying U.S. Army arm-hand signals. *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society*. San Francisco: Human Factors and Ergonomic Society.

Cassinelli, A., Reynolds, C., and Ishikawa, M. (2006). Augmenting spatial awareness with Haptic Radar. *10th IEEE International Symposium on Wearable Computers*. Montreux, Switzerland, October 11-14.

Castle, H., and Dobbins, T. (2006). Tactile display technology: A brief overview of its benefits over visual and audio displays. *Ingenia, Technology and Innovation*, p. 31-34. Retrieved on March 3, 2006 from: http://www.raeng.org.uk/news/ publications/ ingenia/issue20/Castle.pdf

Chiasson, J., McGrath, B.J., and Rupert, A.H. (2002). Enhanced situation awareness in sea, air and land environments. Paper presented at the *RTO HFM Symposium on "Spatial Disorientation in Military Vehicles: Causes, Consequences and Cures."* La Coruna, Spain, April 15-17.

Chilibon, I., Dias, C., Marat-Mendes, J., and Inacio, P. (2005). Lead Zirconium Titanate (PZT) and PolyVinyliDene Fluoride (PVDF) bimorph actuators. *Twelfth International Congress on Sound and Vibration*, July 11-14, 2005, Libson, Portugal. http://www.icsv12.ist.utl.pt/papers/paper.php?id=780

Cholewiak, R.W., Brill, J.C., and Schwab, A. (2004). Vibrotactile localization on the abdomen: Effects of place and space. *Perception and Psychophysics*, 66(6), 970-987.

Cholewiak, R.W., and Collins, A.A. (1991). Sensory and Physiological Basis of Touch. In: M.A. Heller and W.R. Schiff (Eds.), *The Psychology of Touch*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cholewiak, R.W., and Collins, A.A. (2003). Vibrotactile localization on the arm: Effects of place, space, and age. *Perception and Psychophysics,* 65(7), 1058-1077.

Cholewiak, R.W., and Collins, A.A. (1997). Individual differences in the vibrotactile perception of a “simple” pattern set. *Perception and Psychophysics,* 59(6), 850-866.

Cholewiak, R.W., and Collins, A.A. (1995). Vibrotactile pattern discrimination and communality at several body sites. *Perception and Psychophysics,* 57(5), 724-737.

Cholewiak, R.W., Collins, A.A., and Brill, J.C. (2001). Spatial factors in vibrotactile pattern perception. *Eurohaptics*, Birmingham, UK, July.

Cohen, J., Niwa, M., Lindeman, R.W., Noma, H., Yanagida, Y., and Hosaka, K. (2005). A closed-loop tactor frequency control system for vibrotactile feedback. *Association for Computing Machinery Special Interest Group Computer-Human Interaction*, April 2-7. Portland, Oregon. http://web.cs.wpi.edu/~gogo/papers/Cohen_ chi2005.pdf

Elliott, L.R., Redden, E., Pettitt, R., Carstens, C., van Erp, J., and Duistermaat, M. (2006). Tactile Guidance for Land Navigation. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory. ARL-TR-3814.

Engineering Acoustics, Inc. (EIA). (2006). Retrieved 10 Jan 2006 from: http://www.eaiinfo.com/EAI2004/Tactor%20Products.htm

Enriquez, M., Afonin, O., Yager, B., and Maclean, K. (2001). A pneumatic tactile alerting system for the driving environment. *Perceptive User Interfaces*, p. 1-7.

Eysenck, H.J., Arnold, W., and Meili, R. (1972). *Encyclopedia of Psychology, Vol. III,* Eysenck, H.J., Arnold, W., and Meili, R. (Eds.). New York: The Seabury Press.

FreePatentsOnline.com (2006). Virtual environment tactile system. United States Patent 5583478. Retrieved 12 March 2009 from: http://www.freepatentsonline.com/5583478.html

Geiser, G. (1990). *Mensch-Maschine Kommunikation*. München: Oldenbourg.

Gemperle, F., Hirsch, T., Goode, A., Pearce, J., Siewiorek, D., and Smailigic, A. (2003). Wearable vibro-tactile display. Carnegie Mellon Wearable Group, Carnegie Mellon University.

Gemperle, F., Ota, N., and Siewiorek, D. (2001). Design of a wearable tactile display. *Proceedings of the 5th IEEE International Symposium on Wearable Computers*, pp. 5-12. Zurich, Switzerland: IEEE.

Gilliland, K. and Schlegel, R.E. (1994). Tactile stimulation of the human head for information display. *Human Factors*, 36(4), 700-717.

GlobalSpec.com. (2006). Piezoelectric Effect, Theory, Design and Usage.

Goble, A.K., Collins, A.A., and Cholewiak, R.W. (1996). Vibrotactile threshold in young and old observers: The effects of spatial summation and the presence of a rigid surround. *Journal of the Acoustical Society of America,* 99(4), 2256-2269.

Griffin, M., Pera, L., Cabrera, L., and Moore, B. (2001). Operational assessment of the tactile situational awareness system; AFSOC Project 99-060. Hurlburt Field, FL: U. S. Air Force.

Guinard, D., Usson, Y., Guillermet, Ch., and Saxod, R. (2000). PS-100 and NF 70-200 double immunolabeling for human digital skin Meissner Corpuscle 3D imaging. *Journal of Histochemistry and Cytochemistry,* 48, 295-302.

Hawes, V.L., and Kumagai, J.K. (2005). Examination of head and chest located tactile information for infantry wayfinding, DRDC Toronto CR-2005-042. Toronto, Ontario, Canada: Department of National Defence, Defence Research and Development Canada-Toronto.

Ho, S., and Spence, C. (2007). Head orientation biases tactile localization. *Brain Research*, 1144, 136-141.

Horner, D.T. (1992). The effects of complexity on the perception of vibrotactile patterns presented to separate fingers. *Perception and Psychophysics,* 52(2), 201-210.

Jones, L.A., Nakamura, M., and Lockyer, B. (2004). Development of a tactile vest. *Proceedings of the 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 82-89.

Kalb, J.T., Amrein, B.E., and Myles, K. (2008). Instrumentation and Tactor Considerations for a Head-Mounted Tactile Display. ARL-MR-705. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Kohler, J. (2001). Somatosensory senses. Retrieved May 30, 2007 from: http://www.humboldt.edu/~jgk5/cutaneous_senses. htm

Klatzky, T.L., and Lederman, S.J. (2000). Touch. In *Experimental Psychology* (Vol. 4, pp. 147-176), Healy, A. F., and Proctor, R.W. (Eds.). *Handbook of Psychology,* Weiner, I.B. (Editor-in-Chief). New York: Wiley.

Krausman, A., and White, T.L. (2006). Tactile Displays and Detectability of Vibrotactile Patterns as Combat Assault Maneuvers are Being Performed. ARL-TR-3998. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Laidlaw, R.W., and Hamilton, M.A. (1937). Bulletin of the Neurological Institute of New York, 6, 494. Cited in Wilska, A. (1954). On the vibrational sensitivity in different regions of the body surface. *Acta Physiologica Scandinavica*, 31, 285-289.

Lamore, P.J.J., and Keemink, C.J. (1988). Evidence for different types of mechanoreceptors from measurements of the psychophysical threshold for vibrations under different stimulation conditions. *Journal of the Acoustical Society of America*, 83(6), 2339-2351.

Mahoney, C.R., Hirsch, E., Hasselquist, L., Lesher, L.L., and Lieberman, H.R. (2007). Effect of physical exertion on soldier vigilance. *Aviation, Space, and Environmental Medicine*, 78(1), B51-B57.

Mauderli, A.P., Vierck, C.J., Jr., Cannon, R.L., Rodrigues, A., and Shen, C. (2003). Relationships between skin temperature and temporal summation of heat and cold pain. *Journal of Neurophysiology,* 90, 100-109.

McGrath, B.J., Estrada. J., Braithwaite, M.G., Raj, A.K., and Rupert, A.H. (2004). Tactile situational awareness system flight demonstration final report. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2004-10.

McTrusty, T.J., and Walters, C. (1997). Swimmer Inshore Navigation System (SINS) Tactile Situation Awareness System (TSAS) test report. (CSS Test Report 02 August 1997). Coastal Systems Station, Panama City, Florida.

Merlo, J.L., Stafford, S., Gilson, R., and Hancock, P.A. (2006). The effects of physiological stress on tactile communication. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, October 16-20.

Mortimer, B., Zets, G., and Cholewiak, R. (2007). Vibrotactile transduction and transducers. *Journal of the Acoustical Society of America*, 121(5), 2970-2977.

Myles, K., and Kalb, J.T. (2009). Vibrotactile Sensitivity of the Head. ARL-TR-4696. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Niezrecki, C., Brei, D., Balakrishnan, S., and Moskalik, A. (2001). Piezoelectric actuation: State of the art. *The Shock and Vibration Digest*, 33(4), 269-280.

Nordwall, B.D. (2000). Vest enables pilots to feel aircraft attitude. *Aviation Week and Space Technology,* 152(17), 69-70.

Oviatt, S. (1999). Ten myths of multimodal interaction. *Communications of the ACM.* 42(11), 74-81.

Patestas, M., and Gartner, L. (2006). *A Textbook of Neuroanatomy*. Malden (MA): Blackwell Publishing.

Phillips, J.R. (2000). Piezoelectric technology: A Primer. TechOnLine, August 10. http://www.techonline.com/community/ed_resource/ feature_article/8277

Piateski, E., and Jones, L. (2005). Vibrotactile pattern recognition on the arm and torso. *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*.

Psotka, J., Davison, A., and Lewis, S. (1993). Exploring immersion in virtual space. *Virtual Reality Systems*, 1(2), 70-92.

Rabinowitz, W.M., Houtsma, A.J.M., Durlach, N.I. and Delhorne, L.A. (1987). Multidimensional tactile displays: Identification of vibratory intensity, frequency, and contactor area. *Journal of the Acoustical Society of America*, 82(4), 1243-1252.

Raj, A.K., Kass, S.J., and Perry, J.F. (2000). Vibrotactile displays for improving spatial awareness. *Proceedings of the IEA 2000/HFES 2000 Congress*, 181-184.

Redden, E.S., Carstens, C.B., Turner, D.A., and Elliott, L.R. (2006). Localization of Tactile Signals as a Function of Tactor Operating Characteristics. ARL-TR-3971. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Rochlis, J.L., and Newman, D.J. (2000). A tactile display for international space station (ISS) extravehicular activity (EVA). *Aviation, Space, and Environmental Medicine*, 71(6), 571-578.

Ryan, D.M. (2000). Pilots: Close your eyes and fly! *Navy Medicine*. May-June.

Schnupp, J.W.H., Dawe, K.L., and Pollack, G.L. (2005). The detection of multisensory stimuli in an orthogonal sensory space. *Experimental Brain Research*, 162, 181-190.

Schrope, M. (2001). Simply sensational. *New Scientist*, 2 June, 30-33.

Setzepfand, W. (1935). *Zoo Biology*, 96, 236. Cited in Wilska, A. (1954). On the vibrational sensitivity in different regions of the body surface. *Acta Physiologica Scandinavica*, 31, 285-289.

Sherrick, C.E. (1984). Basic and applied research on tactile aids for deaf people: Progress and prospects. *Journal of the Acoustical Society of America*, 75(5), 1325-1342.

Sherrick, C.E. (1985). Touch as a communicative sense: Introduction. *Journal of the Acoustical Society of America*, 77(1), 218-219.

Sherrick, C.E., and Cholewiak, R.W. (1986). Cutaneous Sensitivity. In: Boff, K., Kaufman, L., and Thomas, J.L. (Eds.), *Handbook of Perception and Human Performance*, p. 12-1 to 12-58. New York: Wiley.

Sherrick, C.E., Cholewiak, R.W., and Collins, A.A. (1990). The localization of low- and high-frequency vibrotactile stimuli. *Journal of the Acoustical Society of America*, 88(1), 169-179.

Shimizu, Y., and Wake, T. (1982). Tactile sensitivity to two types of stimulation: Continuous and discrete shifting of a point stimulus. *Perceptual and Motor Skills*, 54, 1111-1118.

Silbernagel, D. (1979). *Taschenatlas der Physiologie*. Berlin, Germany: Thieme.

Smith, D.V., and Margolskee, R. F (2006). Making sense of taste. *Scientific American special edition, Secrets of the senses*. pp. 84-92.

Sorkin, R.D. (1987). Design of auditory and tactile displays. In: Salvendy, G. (Ed.). *Handbook of Human Factors*, p. 549-576. New York: Wiley.

Stedman's Medical Dictionary. (2004). *The American Heritage Stedman's Medical Dictionary* (2nd Ed.). Boston: Houghton Mifflin Company.

Stuart, M., Turman, A.B., Shaw, J., Walsh, N., and Nguyen, V. (2003). Effects of aging on vibration detection thresholds at various body regions. *BMC Geriatrics*, 3(1). http://www.biomedcentral.com/ 1471-2318/3/1

Suzuki, K., and Jansson, H. (2003). An analysis of driver's steering behaviour during auditory or haptic warnings for the designing of lane departure warning system. *Journal of the Society of Automotive Engineering Review*, 24, 65-70.

Tran, P., Amrein, B., and Letowski, T. (2009). Audio HMD Design. In: *Helmet-Mounted Displays: Sensation, Perception and Cognitive Issues*, Rash, C.E., Russo, M., Letowski, T., and Schmeisser, E. (Eds.). Ft. Rucker, AL: U.S. Army Aeromedical Research Laboratory.

van Erp, J.B.F. (2002). Guidelines for the use of vibro-tactile displays in human computer interaction. *Proceedings of Eurohaptics 2002*, pp.18-22. Edinburgh (UK), July 8-10. University of Edinburgh.

van Erp, J.B.F. (2001). Effect of timing parameters on the vibrotactile spatial acuity of the torso. TNO-Report TM-01-A061). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

van Erp, J.B.F., Meppelink, R., and van Veen, H.A.H.C. (2002). A touchy car. TNO-Report TM-02-B001). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

van Erp, J.B.F., and van den Dobbelsteen, J.J. (1998). On the design of tactile displays. TNO-Report TM-98-B012. Soesterberg, The Netherlands: TNO Human Factors Research Institute.

van Erp, J.B.F., and van Veen, H.A.H.C. (2001). Vibro-tactile information presentation in automobiles. *Proceedings of Eurohaptics 2001*, pp.99-104. Birmingham (UK): July 1-4. University of Birmingham.

van Erp, J.B.F., Veltman, J.A., van Veen, H.A.H.C., and Oving, A.B. (2003). Tactile torso display as countermeasure to reduce night vision goggles induced draft. DTIC X5-X5, pp. 1-8. Soesteberg, The Netherlands: TNO Human Factors Research Institute.

van Erp, J.B.F., and Werkhoven, P.J. (1999). Spatial characteristics of vibro-tactile perception on the torso. TNO-Report TM-99-B007. Soesterberg, The Netherlands: TNO Human Factors Research Institute.

van Veen, H.A.H.C., and van Erp, J.B.F. (2001). Tactile information presentation in the cockpit. In: Brewster, S., and Murray-Smith, R. (Eds.), *Haptic Human-Computer Interaction*, *First International Workshop*, pp. 174-181. New York: Springer.

Verrillo, R.T. (1962). Investigation of some parameters of the cutaneous threshold for vibration. *The Journal of the Acoustical Society of America*, 34(11), 1768-1773.

Verrillo, R.T. (1966). Vibrotactile thresholds for hairy skin. *Journal of Experimental Psychology*, 72(1), 47-50.

Weber, E.H. (1978). *The Sense of Touch* (Ross, H.E., and Murray, D.J., Trans.). London: Academic. (Original work published in 1834).

Weinstein, S. (1968). Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In: D.R. Kenshalo (Ed.). *The Skin Senses*. Springfield, IL: Charles C. Thomas.

Wickens, C.D. (1984). Processing resources in attention. In: Parasuraman, R., and Davies, R. (Eds.), *Varieties of Attention*. Hillsdale, NJ: Erlbaum.

Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science*, 3(2), 159-177.

Wilska, A. (1954). On the vibrational sensitivity in different regions of the body surface. *Acta Physiologica Scandinavica*, 31, 285-289.

Wood, D. (1998). Tactile displays: Present and future [Editorial]. *Displays*, 18, 125-128.

Youngblut, C., Johnston, R., Nash, S., Wienclaw, R., and Will, C. (1996). Review of virtual environment interface technology. Report IDA P-3186. Alexandria, VA: Institute for Defense Analysis.

19 THE POTENTIAL OF AN INTERACTIVE HMD

James E. Melzer
Clarence E. Rash

Touted as having wide-spread potential ever since their appearance in the 1960s, helmet-mounted displays (HMDs) can be found in hands-free viewing applications (Melzer, 2006) and in visually coupled systems (Kocian, 1987) for military (Rash, 2001; see also Chapter 1, *The Military Operational Environment* and Chapter 4, *Helmet-Mounted Displays* of this volume), simulation and training (Casey and Melzer, 1991; Melzer and Porter, 2008; Melzer and Simons, 2002), and virtual reality applications (Barfield and Furness, 1995; Kalawsky, 1993). In trying to explain why they have not been more pervasive, Keller and Colucci (1998) identified factors such as cost, lagging technology, and sub-optimal ergonomics. Hopper (2000) suggested that the "visceral dislike" of wearing a monitor on one's head has not yet been countered by an application that sufficiently excites potential users. This is somewhat understandable, because too often HMDs have been developed without a *user-centered design* focus. The result was that some early designs were uncomfortable and caused eye strain (Moffitt, 1997), with a tacit demand that the user had to *adapt to* the technology, essentially becoming a slave to the whims of the hardware designer. This is unfortunate, because fundamentally, the benefit of the HMD lies not in the hardware itself, but in the way it aids users in performing their duties that helps them overlook the added weight, cost and complexity. So while the hardware obviously must meet certain application and user-dependent performance requirements (e.g., field-of-view, luminance, contrast, focus, binocular alignment, fit, weight and balance), to make the technology truly work, we must do more. In this chapter we explore the HMD as part of an interactive system, a role consistent with natural exploratory behavior, as described by the "perceptual loop" in Chapter 2, *The Human-Machine Interface Challenge* (Figure 2-2). We envision the HMD as part of a system that *adapts to* the user – Bonner, Taylor, Fletcher, and Miller, (2000) use the term "Cognitive Cockpit" and Schnell (2008) uses the term "Smart Avionics" – that is, one that enhances situation awareness, encourages or enables correct decision-making and reduces workload.

First, we examine the benefits of the HMD over traditional cockpit displays as enabling the pilot to spend more time looking outside of the cockpit. We then focus on situation awareness (SA), cognitive workload, and the associated information acquisition, model-updating and decision-making loop to examine how overloading the pilot can cause this loop to breakdown. From there, we discuss attention, multiple perceptual and cognitive resources and the implications of cross-modal sensory integration, followed by a discussion of some developments in HMD symbology. Finally, we explore ways in which a feedback loop that includes psychophysiological monitoring (e.g., encephalograms, evoked potentials, and ocular-motor measures) can provide real-time integration into the HMD system, and promises to optimize the human-machine interface enhancing situation awareness without contributing to cognitive overload. Advances such as these will allow HMDs to be taken beyond a hands-free display or a visually coupled system to where it can be considered a *cognitive prosthesis*,[1] assisting pilots in the face of overwhelming workload or physical stress that could compromise their mission or their life (Melzer, 2008).

This chapter is intended to be somewhat speculative, to project applications and enablers of the technology that have yet to be fully realized. While other authors in this volume have dealt with some of the basic perceptual, user interface and hardware-related issues, it is our intention to invoke thought and discussion about the future of

[1] The term *Cognitive Prosthesis* is taken from the brain injury rehabilitation literature. It is a computer-based, assistive, compensatory technology designed for individuals who through either injury or illness have acquired a cognitive deficit, thereby allowing them to participate in and navigate through the everyday world (Cole and Matthews, 1999).

HMDs by framing this chapter within a neuroergonomics[2] context. Thus, a better understanding of the ways humans perceive and react to incoming sensory information will allow designers to "radically rethink the design of human-machine system interfaces to optimize the flow and exchange of data between humans and machines" (Berka et al., 2007). Making HMDs fully interactive in these ways will lead to the emergence of more wide-ranging applications.

Why an HMD?

What makes the HMD better than other cockpit displays such as head-down displays (HDD) or head-up displays (HUD)? For the answer, we need to examine the essence of natural human exploratory behavior. In his classic text, Gibson (1986) describes the human as a *perceptual system*: "… the eye is a part of a dual organ, one of a pair of mobile eyes, and they are set in a head that can turn, attached to a body that can move from place to place." The implication is that the capabilities of this perceptual system are fully exploited only if it is free to explore the environment, a concept consistent with Piaget's (1952) thesis that exploration of the environment is fundamental to cognitive development in infants. Although cockpit displays have advanced from HDDs (requiring the head and eyes to be within the cockpit) to HUDs (allowing the head and eyes to be out of the cockpit, but limited to a single line-of-sight), the information critical to achieving situation awareness is still only available in a small region of the pilot's forward field-of-regard.

If, however, we link an HMD to the aircraft with a head-orientation tracker, it becomes a Visually Coupled System (VCS - see Kocian, 1987) that allows the pilot to take advantage of a fuller array of information by overlaying imagery or symbology that is reactive to head motion and which may be aircraft- or geospatially-referenced.[3] Now the HMD (as part of the VCS) expands the pilot's useful field-of-regard by allowing him/her to turn head *and* eyes to better perceive the environment. This gives the pilot access to information when looking outside the limited field-of-view of the HUD with cues to guide or direct attention to specific objects, landmarks or targets because the pilot's threats are not just in front of the aircraft[4]. A head-tracked HMD also allows the pilot to direct another aircraft or crew member to an object or location, or to bring weapons to bear on a specific off-boresight target simply by looking at it (Arbak, 1989; Merryman, 1994), significantly enhancing the aircraft's effectiveness as a weapons or observation platform. Thus, the HMD aids the pilot by: 1) reducing time spent with head down in the cockpit, 2) reducing perceptual switching time from cockpit to outside world (i.e., attention, vergence and focus), 3) presenting imagery that can be either earth- or aircraft-referenced, and 4) allowing the pilot to be directed to a target of interest and then to track the target as it moves (Yeh, Wickens and Seagull,

---

[2] "Neuroergonomics focuses on investigation of the neural bases of such perceptual and cognitive functions as seeing, hearing, attending, remembering, deciding and planning in relation to technologies and settings in the real world... Knowledge of how the brain processes visual, auditory and tactile information can provide important guidelines and constraints for theories of information presentation and task design…Neuroergonomics has two goals: 1) to use existing and emerging knowledge of human performance and brain function to design technologies and work environments for safer and more efficient operation; and 2) to advance understanding of brain function in relation to human performance in real-world tasks" (Parasuraman, 2003; 2007). Neuroergonomics requires an understanding of how the brain processes auditory, visual and tactile stimuli as a basis for designing interfaces between humans and technology. It is not intended to be just a laboratory science, but one that should form the basis for interaction with technologies in the real world (Hancock and Szalma, 2003)

[3] Imagery on the HMD can be displayed in three frames of reference: 1) aircraft-referenced (such as the shape of the front of the aircraft), 2) earth-referenced (either real objects such as runways or horizon lines or virtual objects such as safe pathway in the sky, threat/friendly locations engagement areas, waypoints, and adverse weather), and 3) screen-referenced (such as altitude, airspeed, or fuel status) (Yeh, Wickens, and Seagull, 1998; Procter, 1999).

[4] In simulation studies with an HMD, pilots spent 70 to 80% of their time *not* looking along the line of sight of the HUD (Arbak, 1989), which is especially critical during nap-of-the-earth (NOE) flight. Geiselman and Osgood (1994) found that when provided with useful ownship information, test subjects look further off-boresight for longer periods of time.

1998). Rogers, Asbury and Haworth (2001) surveyed a group of AH-64 Apache helicopter pilots to explore areas in which HMDs could enhance their abilities. Their list included: 1) aiding in maintaining situation awareness, 2) allowing for improved target acquisition, 3) aiding in moving through their environment, 4) improving symbology without increasing clutter, and 5) providing additional warning information. The results reinforce the intent of this chapter as these aviators had first-hand experience with the Integrated Helmet Display and Sighting System (IHADSS, Rash, 2001) and it reveals something about the support for HMDs by pilots with first-hand knowledge of their capabilities. In the next sections, we will examine ways to further enable these advantages.

Situation Awareness and Cognitive Workload

Achieving situation awareness (SA) for the pilot is the primary and ultimate goal of the HMD designer. A commonly accepted definition of SA divides it into three levels: "Level 1) *the perception of the elements in the environment within a volume of time and space,* Level 2) *the comprehension of their meaning, and* Level 3) *the projection of their status in the near future"* (Endsley, 1995a) (Figure 19-1). This definition has been applied to tasks as diverse as air traffic control, battlefield management, medical procedures, firefighting, weather forecasting, football and any environment where a timely and global understanding of a dynamic situation is vital (Endsley, 2000; Endsley and Hoffman, 2002; Uhlarik and Comerford, 2002). For the pilot, SA can be thought of as a dynamic interpretation of constantly changing information considering the future state of the aircraft and environment, essentially an understanding of the "*whatness, whereness* and *whenness*" (Helmetag et al., 1999) of the environment through which a pilot must fly and fight. To have full SA, the pilot gathers information (Level 1 SA) and creates a mental model of the current state of the aircraft and surrounding environment (Level 2 SA). The information actually used – sometimes inconsistent and disjointed – may include visual, auditory and/or tactile *meta-knowledge.*[5] With this information, pilots use their training and cognitive processing skills (including short-term, working and long-term memory resources) to convert the *navigational* knowledge – derived from an egocentric point of view, generally acquired by scanning the cockpit instruments and the outside world, listening to the multitude of communication channels and sensing the behavior of the aircraft – into *configurational* knowledge or a "bird's-eye" view of the current situation.[6] But since the environment (and aircraft status) is constantly changing, this mental model is both dynamic and accretionary, requiring the pilot to repeat the cycle of information gathering, information digesting, model building and prediction over and over again for the duration of the flight,[7] while using a minimum of workload[8] or effort. The optimal state is where the pilot has full SA but is only under a moderate or light workload.

But depending on the amount of data presented, the way in which it is presented, the state of the aircraft, and the sum of all other distractions, the process of *cognitively digesting* incoming data to produce and update an accurate SA model taxes the pilot and breaks down when his capacity to process the information *exceeds* his resources. In other words, "In the complex and dynamic aviation environment, information overload, task complexity, and multiple tasks can quickly exceed the aircrew's limited attention capacity. The resulting lack of SA can result in poor decisions, leading to human error" (Endsley, 1995b). SA fails most often when cognitive overload causes the pilot to lose touch with Level 1 SA (i.e., perceiving the environment). A recent assessment of

---

[5] The term "meta-knowledge" is used here to mean *knowledge about knowledge* from sensors and cockpit displays or data that may be one step removed from the actual information itself. The intent is to emphasize the additional cognitive processing needed by the pilot to convert it to useful knowledge.

[6] This mirrors a body of work in cognitive mapping, in which someone exploring a new environment is gradually able to create a schematic map of the area in his/her head after having explored it (egocentrically) on foot (Kuipers, 1978).

[7] Note here the similarities here to the perceptual loop described in Figure 2-2 of Chapter 2, *The Human-Machine Interface Challenge* and to John Boyd's OODA Loop (for Observe, Orient, Decide, Act) for fighter pilots (Boyd, 2007).

[8] Workload is a multidimensional construct (Hancock and Szalma, 2003), sometimes called the "flip side of the same coin" as SA (Endsley, 1993). It is commonly defined as the demand on attentional and cognitive resources required maintaining SA.

U.S. air accidents found that 80% occur at this level of perception, with the worst failures falling into the sub category (37%) of "failure to monitor" (Smith, 2006). This happens when aircrews are distracted because of

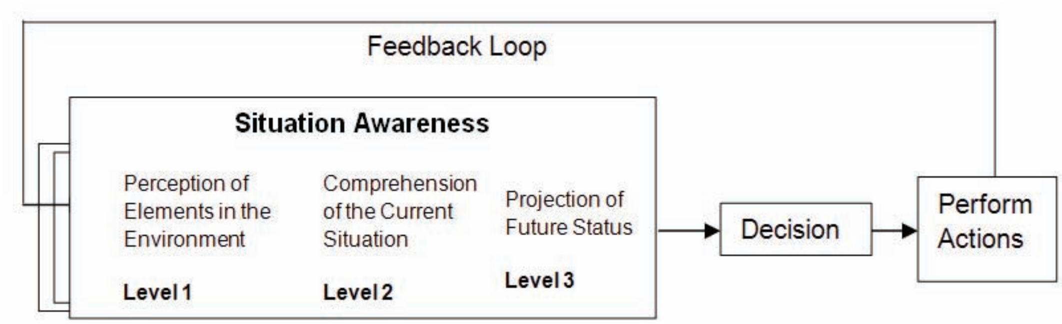


Figure 19-1. Shows a nested Level 1, Level 2 and Level 3 model of Situation Awareness and the continuous feedback loop necessary to maintain SA (after Endsley 1995a).

cognitive overload that they fail to address the real issues at hand. Factors such as divided attention, having too much incoming data, or having to expend too much cognitive effort limits the pilot's ability to monitor the current status and to predict the future state of his aircraft. Thus: "how quickly one converts navigational (egocentric) knowledge to survey ("God's-eye") knowledge and is able to achieve true situational awareness depends partially on *the manner in which the information is presented, the cognitive capabilities of the individual and the amount of cognitive energy the individual is willing to expend in the effort*." (Helmetag et al., 1999, emphasis added). Somehow, we must provide the pilot with information that is easily digested (or perhaps "pre-digested"), to reduce cognitive overload. Endsley and Hoffman (2002) refer to this as the Lewis and Clark Principal: "The human user of the guidance needs to be shown the guidance in a way that is organized in terms of their major goals. Information needed for each particular goal should be shown in a meaningful form, and should allow the human to directly comprehend the major decisions associated with each goal."

Attention, Cognitive Resources and Cross-Modal Integration

The modern pilot is faced with a complex array of tasks (i.e., aviate, navigate, communicate, and systems management – see Wickens, 2007) that by nature require multiple cognitive and perceptual resources, multiple attentional resources and multiple auditory and physical responses. The problem is how to direct the pilot's attention, enable the perception or acquisition of critical information (Level 1 SA), encourage the synthesis of the information (Level 2 SA) and provide a mechanism for the pilot to take action based upon the prediction of future state (Level 3 SA), within the pressures of flying, and the added limitation that the pilot's visual and auditory channels become progressively saturated. The goal is to find methods of presenting information, or ways to capture and guide attention that will not overwhelm the pilot.

Wickens and McCarley (2008) discuss five discrete types of attention[9] though they also provide a simpler definition which divides attention into just two categories: *filter* and *fuel*. Humans are faced with a constant barrage of stimuli which, if not *filtered* by attentional resources, would rapidly be overwhelming. If, however, when attending to a specific stimulus, available perceptual and cognitive resources can be energized to address the implications of this stimulus. This is the *fuel* aspect of attention. Improperly directed attention can reduce situation awareness (i.e., see Smith, 2006 and the implications for Level 1 SA for "failure to monitor") or increase workload by having the pilot attend to too many stimuli varying widely in priority. This filtering is the function of executive control, that part of the brain which allows attention to be directed to the stimulus of choice. When

[9] These are: focused, selective, switched, divided and sustained (Wickens and McCarley, 2008).

attention (or task or perceptual modality – Koch, 2001; Spence and Driver, 1997) is switched, there is an associated time penalty because of the serial steps involved in doing so (*goal shifting* followed by *rule activation* – Rubinstein, Meyer and Evans, 2001). With multiple attention shifts, we pay a larger penalty, and it raises the possibility that the goals will not be remembered upon returning to the original task. This can be overcome using mental models, frequently observed in the different ways experts and novices perform in high workload situations with high information load. Experts use shortcuts such as prioritization, and "gistification" to achieve Level 1 and Level 2 SA (Endsley, 2000). Experts may also pattern match, then load response scripts to prototypical situations or schema. Doing so may allow the pilot to achieve Level 3 SA without having to overload working memory. But in drawing upon schema, the pilot may be subject to situational biasing, possibly reducing his responsiveness to novel situations or stimuli.[10] Thus the pilot also needs to recognize when the information is in conflict with previously learned models and to modify the response,[11] though there is an associated increase in workload and communications (Marshall, 2007b). Problems arise when the executive control function is continuously over-tasked, and the pilot does not have enough reserve capacity to plan out behaviors required to accomplish the complex task of flying. *Unlike* someone who has suffered brain damage, this is a temporary affliction. However, *like* someone who has suffered brain damage, they lack the resources to make complex decisions associated with the aviate, navigate, communicate and systems management tasks. This is where the cognitive prosthesis approach can help (Cole and Matthews, 1999 and Melzer, 2008), by lightening their cognitive load and properly guiding them through difficult situations.

Models in the literature provide a better understanding of the issues surrounding the ways humans perceive and react to incoming information and how to enable the human-machine interface without causing cognitive overload. In his Multiple Resources Theory (MRT), Wickens (1980; 1984; 2002a) provides a framework for predicting performance effects when the pilot is required to execute multiple simultaneous tasks and distinguishes between and within three stages of cognitive processing. He posits that there will be greater interference (and subsequent increased workload) between two tasks if they share the same pool of resources which draw upon physically separate cortical functions:

- *Input perceptual or sensory modalities* (auditory vs. visual) – It is easier to divide attention between hearing and seeing (i.e., auditory/visual) tasks than between two auditory (auditory/auditory) or two visual (visual/visual) tasks, because the sensory modalities require separate resources (drawing upon the separate auditory and visual sensory cortices).
- *Central processing stages* (perceptual and central processing/cognitive vs. response) – Working memory resources used for perceptual and cognitive activities are the same; and they are separate from those that help in executing responses (drawing upon the right and left hemispheres).
- *Response codes* (spatial versus verbal) – Verbal and spatial processes or codes used in perception, working memory, or responses depend on separate resources, which can account for individuals' ability to simultaneously perform well manually and verbally (because they draw upon the different hand and mouth/respiratory regions of the motor and pre-motor cortex).
- *Channels of visual information* (focal versus ambient) – There are two channels of vision, the focal and the ambient that utilize separate resources (nominally in the central and peripheral areas of our vision, respectively).

[10] It is also important to consider the (possibly undesirable) implicit feedback and filtering loops between each nested element within SA. These may manifest themselves when expectations bias perceptions, with at times disastrous consequences (see previous chapters) because it may cause the pilot to reject valid inputs and "loose touch" with Level 1 SA.

[11] Here the SA loop starts to overlap with sensemaking. While the former is generally associated with fitting of data into an already-established model, sensemaking is the attempt to find understanding of disparate and disjointed information by creating a new model (Weick, Sutcliff & Obstfeld, 2005; Leedom, 2001).

MRT says that resources needed for perception and cognition are the same, both of which involve working memory. Thus the resources required to gather knowledge which forms the basis of Level 1 SA are those same resources needed to create and manipulate the model in working memory, which is key to Level 2 SA, the understanding of the current situation. In addition, any complex mental manipulations of the data needed to arrive at that determination will be the same resources needed to determine the future state of the aircraft.

Wickens (2002a) also separates the visual channels into focal and ambient modes. The *focal* mode generally lies in the central region of vision and is dedicated to answering the “What?” about our environment. It typically requires our attention, is sensitive to light level (and our inherent refractive error) and a full range of spatial frequencies (Leibowitz, Shupert and Post, 1985). Under stress from shifts in attention, the individual may suffer a visual narrowing in the focal visual mode due to shifts in attention, which may also contribute to change blindness (Wickens, 2002b).

The *ambient* mode of vision, on the other hand, addresses the question of “Where?” Though overlapping somewhat with the focal mode, it is generally found in the periphery of our vision, and acts together with our vestibular system to help with spatial orientation. It requires only low spatial frequency information, and is more susceptible temporal frequency such as movement and flicker, though less sensitive to refractive error and ambient light level. The importance for HMDs is that the ambient visual mode is thought to be “pre-attentive” or automated and therefore may require *no cognitive resources at all* (Uhlarik and Comerford, 2002). Thus the ambient mode of vision will likely not suffer from attentional narrowing due to overload and may be an important path to improving SA without increased workload.

Wickens (1980; 1984) states that separating the sensory modalities – auditory versus visual versus tactile – allows attention to be divided. Spence and Driver (1997), however, take issue with Wickens’ interpretation of absolute separation of resources and posit that there are limitations on their independence due to cross-modal linkages between these covert (i.e., internal processing) visual, auditory and tactile attentional resources. For example, if the separate tasks (which use separate resources) place high demand on the individual – as in the case of time-sensitive responses – subjects will tend to serialize their responses rather than operate in a truly parallel manner. The distinction may be a bit more subtle, though, in that Wickens’ resource separation focuses on the perception, cognition and response to *continuous* tasks versus Spence and Driver’s focus on discrete tasks requiring attentional shifts. These latter researchers point out that if an event is expected in one sensory modality, and it occurs in another, there is an attentional penalty due to the modality shift. They demonstrated that if a subject was expecting a cue in an auditory or visual modality, but it occurred as a tactile cue, there was a 16% performance lag. Furthermore, they found that “pre-cueing” in one mode can enhance the attentional resources and perception of an event in another mode. This is especially true for auditory and visual events that occur from the same spatial location, though there is still an attentional advantage even if they don’t. Thus, the three different sensory modalities can act effectively as pre-cueing “notifiers” of an event in another modality in various combinations. The most effective appears to be an auditory cue, especially when used to notify the subject of a time-critical event, provided it is presented within 300 milliseconds (ms) of a visual event (Pouget, Deneve and Duhamel, 2004). Hameed et al., (2007) found that a directional tactile cue improved visual detection rates by 43%. This process which combines visual, auditory and tactile sensory signals relating to the same object in time and space appears to be something humans excel at, taking advantage of multi- and intermodal redundancies.[12] When integrating audio earcon and visual icon cues[13] into a display, it is important that we understand these

[12] The only notable exception is that a visual notifier *does not* effectively cue an auditory event (Spence and Driver, 1997).

[13] *Earcons* are abstract sounds where the meaning must be learned and where the meaning forms a hierarchical structure. The typical example is groups of musical notes to designate types of input errors. *Auditory icons* are natural sounds that have a meaning associated with the object they represent. Throwing a document in the desktop trashcan can be accompanied by a crumpled-paper sound to symbolize deleting a file within the context of the desktop metaphor (Houtsma, 2003). Care must be taken, however, to ensure that the meaning is clear, that the messages are synchronized and that there a valid perceptual *co-occurrence* between them (Bertelson and de Gelder, 2004)

issues of multisensory integration so the pilot can make accurate and meaningful statistical inferences (Pouget, Deneve, Duhamel, 2004) about the intent of the multimodal stimulus.

Research has shown that spatial or three-dimensional (3-D) audio[14] can dramatically improve safety and performance, decreasing workload and improving SA by superimposing geospatial directionality on radio communications and by using the audio cues redundantly with visual cueing to direct the pilot's attention for alerts and warnings (see Bolia, 2004, for an excellent collection of papers on the subject). The benefit is to increase situation awareness and decrease workload by decreasing audio clutter, by providing an intuitive spatial location for warnings and alerts, and by redundantly coding external threats and waypoints as an audio cue to direct visual attention. 3-D audio cueing especially when used with an HMD, reduces search time and improves situation awareness for the user (Bolia, D'Angelo, and McKinley, 1999; Flanagan et al., 1998; Houtsma, 2003; see also Chapter 14 of this volume, *Auditory-Visual Interactions*).

We perceive the direction of sounds ("the eyes follow the ears" – Wenzel, 1992) by processing temporal, intensity, phase and spectral differences between the sounds reaching our left and right ears. These differences result from the interference of the head, pinnae, and torso with a sound wave, a transform called the Head Related Transfer Function, or HRTF.[15] Accuracy is less with auditory tracking than with visual tracking so relying on the former for accurate cueing is not appropriate since this is not how – ecologically speaking – we search and navigate through and within the real world.

Spatial hearing also allows the advantage of discriminating sounds in the presence of noise. Providing a spatial separation between the audio source of interest and interfering noise improves the listener's ability to detect and understand the audio content, much like the so-called *Cocktail Party effect*, where we can listen to different conversations within a crowded room simply by attending to them (Cherry, 1953). Similarly, spatial hearing improves the understanding of speech when there are competing sources such as multiple talkers. Assigning a distinct spatial direction (and location) for each source dramatically improves intelligibility compared to when they originate from the same location. Such an advantage would seem natural in an aviation cockpit; though it appears there is much improvement needed with spatialization protocols.[16]

Examples of HMD Imagery

Information displayed to the pilot must be only that which is essential for the task at hand and must be presented so that interpreting the data does not overload the pilot's already-taxed perceptual and cognitive resources. In this section, we present examples of HMD symbology that have been shown to improve performance, i.e. imagery that: 1) provides cognitively pre-digested information, 2) provides stable frames of reference and 3) stimulates the

---

[14] 3-D audio refers to radio channels, cockpit warnings, threat and target designations that have a spatial direction and range (also discussed elsewhere in this book – see Chapter 5, *Auditory Helmet-Mounted Displays*).

[15] The Head-Related Transfer Function (HRTF) refers to binaural hearing effects resulting from the location of our ears on either side of our head. The HRTF consists of three components: Interaural Time Delay (ITD - sounds reach the closest ear first, followed after a short time delay by the sound reaching the other ear), Interaural Intensity Difference (IID - the closest ear hears the full intensity of the sound, the farthest ear, shadowed by the head, hears a reduced intensity of the sound), and finally, spectral filtering from the pinnae (the outer ear filters certain frequencies depending on their fore/aft or up/down location). Because the HRTF differs from person-to-person, it is difficult to generate a *generic* HRTF that will accurately restore "hear-through" for all users (Chapin et al., 2004) though there are ongoing efforts in this area to overcome these limitations (McIntire et al., 2008).

[16] There is no standard or protocol for assigning radio channels or avionics warnings to either relative or absolute geo-spatial locations. For example, should the wingman or the control tower audio come from the correct geospatial location relative to ownship or from some standardized location? In addition, there is no standardized set of non-speech audio warnings and alerts, such as low oil, threats, low fuel, weapon status, and most aviation helmet systems are do not support 3-D audio because they are monaural.

ambient mode of vision. In all cases, we will assume that the HMD is part of visually coupled system (Kocian, 1987; Rash, 2001) in which a tracker communicates helmet-referenced orientation data to a sensor, a computer or a mission processor.

Early HMDs used a simple reticle similar to the one shown on the left in Figure 19-2. The targeting cross in the center is boresighted to the aircraft's weapons and the small diamonds around the edges indicate a "look-to" direction for the pilot. This simple symbology unlocks the pilot from the forward line-of-sight of the aircraft HUD, giving him the ability to designate targets and to aim and deliver weapons off-boresight, and has been shown to have profound implications as a force multiplier (Arbak, 1989; Merryman, 1994).

Compare this with a more sophisticated symbology set on the right side of Figure 19-2 that could be found on a more recent fixed-wing pilot's HMD. The circle within a box at the end of the "look-to" arrow is the target designator box (or "TD box") which combines the center cross and directionality diamonds of the early version. The later version also provides more flight data such as altitude, airspeed, heading, attitude, and weapon status. Having this information readily available anywhere the pilot is looking frees him from

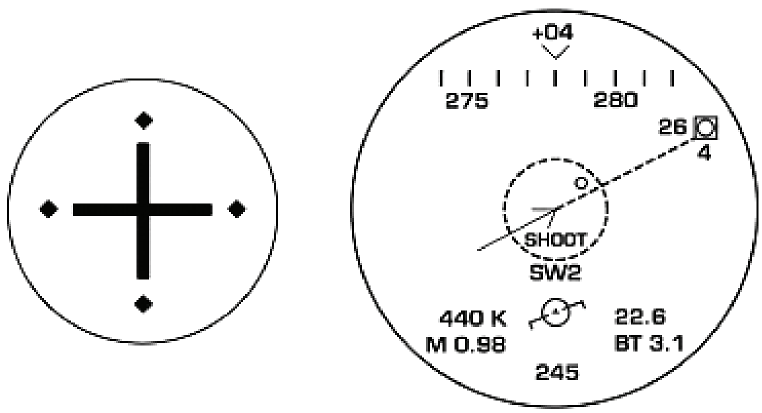


Figure 19-2. Comparison of an early HMD reticle (left) with a more sophisticated symbol set intended for use on fixed-wing fighter HMDs (after Melzer, 2006).

having to look inside the cockpit or forward at the HUD to gather that same Level 1 SA information. But with the exception of the improved targeting reticle, it is only a re-mapping of the information that might normally be found on the HUD and results in a cluttered out-the-window view. With the introduction of HMD-based off-boresight tracking and targeting, the U.S. Air Force has been examining ways to ease the pilot's transition from on-boresight HUD symbology, because pilots complain that there is too much symbology on their HMD. One solution is to simply de-clutter the imagery when the pilot looks off-boresight, with the standard symbology returning when the pilot looks back "on boresight" (Albery, 2007), or to permit the pilot to customize the declutter mode depending on preference and situation.[17] In a series of papers, Jenkins and his colleagues (Jenkins, 2003; Jenkins, Turling and Brown, 2003; Jenkins, Sheesley and Bivetto, 2004 and see also Albery, 2006) evaluated the Advanced Non-Distributed Flight Reference (Advanced NDFR) for displaying ownship status information that is easily read without cluttering up the HMD field-of-view, but which provides sufficient information to allow the pilot to feel confident enough to spend more time off-boresight. The key is an open circle – the arc segment attitude reference (ASAR) originally conceived by Dornier in 1987 – which changes as a function of aircraft attitude as shown in Figure 19-3. At straight and level, the only part showing is the bottom 180°. As the pilot climbs, the circle gradually closes until it becomes a full circle at a 90°-climb. Likewise, as the pilot dives, the circle gradually shrinks until it is only a small segment at a 90°-dive. Jenkins and his colleagues

[17] The only caveat is that critical data must be "re-cluttered" at some point so the pilot does not miss a key piece of information.

improved the original NDFR by adding digital flight path angles, altitude, airspeed and heading to display of rate-of-change data. Simulation studies and flight test results indicate that this was well accepted by pilots and allowed them to spend more time looking off-boresight, out of the aircraft.

In a 1998 study, several alternate HMD imagery concepts were investigated for fixed-wing pilots at the U.S. Naval Weapons Center and Boeing's Phantom Works (Proctor, 1999), including geostationary "X-ray vision" imagery that allowed pilots to see through hills and ridges when flying terrain-masking routes, "message bubbles," virtual sign posts and geospatially-fixed synthetic grids placed over *actual terrain contours*. Message bubbles and other message icons were placed in the display where no other key information was located, freeing the pilot from having to mentally "declutter" the imagery. It allowed the pilots to go quickly from *egocentric* knowledge to *survey* knowledge with a minimum of cognitive processing, and is consistent with our previous contention that pre-digesting the information eases the transition from Level 1 to Level 2 SA.

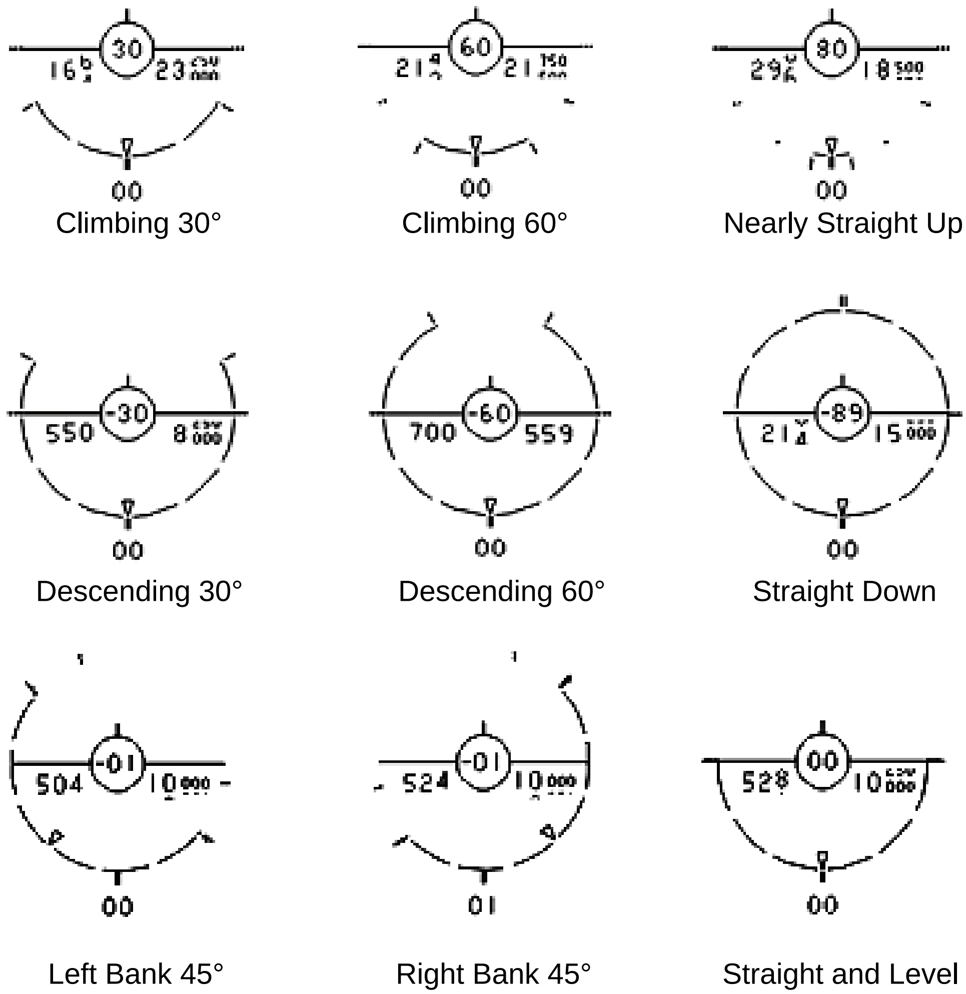


Figure 19-3. Advanced Non-Distributed Flight Reference symbology for fixed-wing HMDs shown in various phases of flight orientation. The number in the central circle indicates is the digital flight path angle. The numbers to the left and right are the airspeed and altitude, respectively. The number at the bottom shows the heading. (Used with permission, U.S. Air Force, 711th HPW/RHCV.)

Many domestic helicopters – with the notable exception of the AH-64 Apache – are equipped for night flight with an HMD in the form of the cathode-ray tube (CRT)-based NVG-HUD mounted on the Aviator's Night Vision Imaging System (ANVIS) goggles. The symbology is not head-tracked and thus neither geo- nor aircraft-stabilized (Yona, Weiser and Hamburger, 2004). Rather, it is generally a re-mapping of the head-down display information that would otherwise be readily accessible to the pilot during daytime flight. As part of the Air Warrior Block 3 program, the U.S. Army specified that the next generation of HMDs provide "*intuitive situational and system awareness displays* that permit pilots to fly the aircraft continuously with heads-up, eyes out regardless of environmental conditions" (U.S. Army, 2003, emphasis added). While helpful, it is generally felt by many pilots that this version of the NVG-HUD does not meet the definition of "intuitive situational awareness displays."

Still and Temme (2001) developed a symbology set called "OZ" to provide a graphical depiction of aircraft position and orientation (Figure 19-4). Their concept uses a star-field metaphor to map the external world into a coordinate system that displays both translations and rotations, shows the aircraft's attitude and location within the external world and takes advantage of the natural human perception of flow fields (Gibson, 1986). OZ enables traditional instrument panel information to be obtained at-a-glance instead of requiring the pilot to sequentially scan and interpret the individual dials and gauges.[18] In a more recent study, Still and Temme (2008) expanded the OZ symbology as an aid to helicopter pilot trainees learning the difficult task of hovering. Their results showed a reduction in training time to reach proficiency because the OZ symbology helped the students learn to interpret the complex motion cues in a helicopter. Though specifically designed for use with a HDD, there is fundamentally no reason why this same symbology could not be used with an HMD.

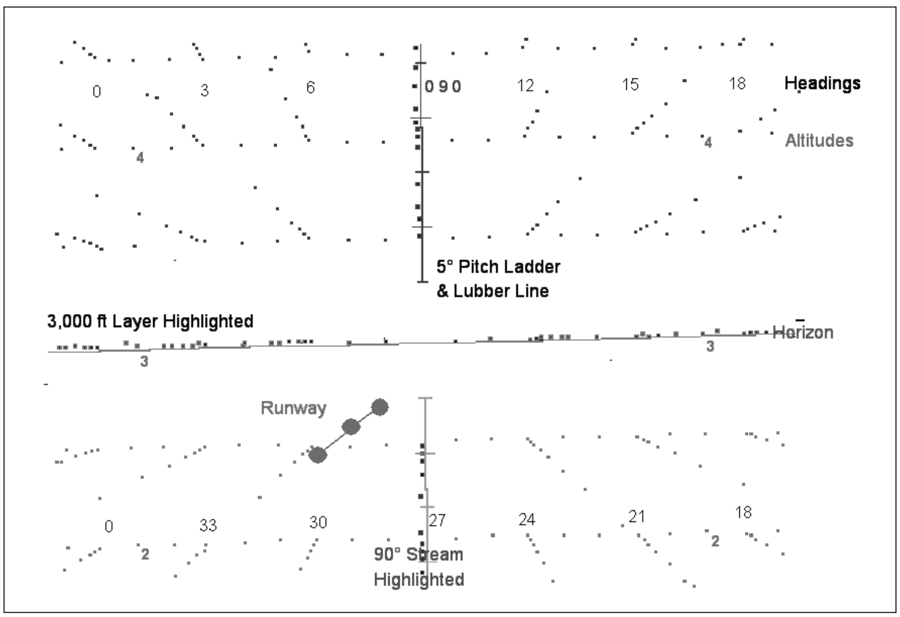


Figure 19-4. OZ symbology set which uses a star field metaphor to show flowfield, elevation and attitude. (Imagery courtesy of Dr. David Still and Dr. Leonard Temme, U.S. Army Aeromedical Research Laboratory, Ft. Rucker, AL, used with permission.)

Rogers and Asbury (2007) created a clock obstacle warning icon as part of their Rotorcraft Obstacle Avoidance Display (ROAD) (Figure 19-5) that could be unobtrusively located on the pilot's HMD to indicate the relative

---

[18] Hansen, Rybacki and Smith (2006) use the term: "synthesize the dials" to describe this part of the process.

location of a possible collision threat. This simple icon was very well received by the test (pilot) subjects who were impressed with how intuitive it was. Note the "splat" marker at the upper right hand side that indicates the direction of a potential collision.

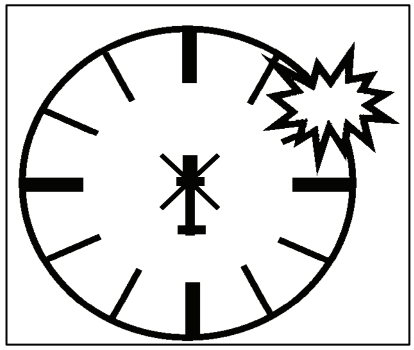


Figure 19-5. A Clock Obstacle Warning used in the Rotorcraft Obstacle Avoidance Display. Note the orientation of the helicopter and the potential collision direction. (Rogers and Asbury, 2007, used with permission).

A special imagery set designed by *Primordial* (Milbert, 2005) for ground soldier applications takes advantage of both conformal symbology and lessons-learned from the video game industry by indicating key points of interest or navigational information and their location relative to the soldier's "forward" position. A small, semi-transparent display window in the lower corner rotates as the soldier turns his head and body, providing a survey map view of the surrounding environment with forward indicated as the "up" direction (Figure 19-6), giving the soldier a better understanding of the surrounding environment.

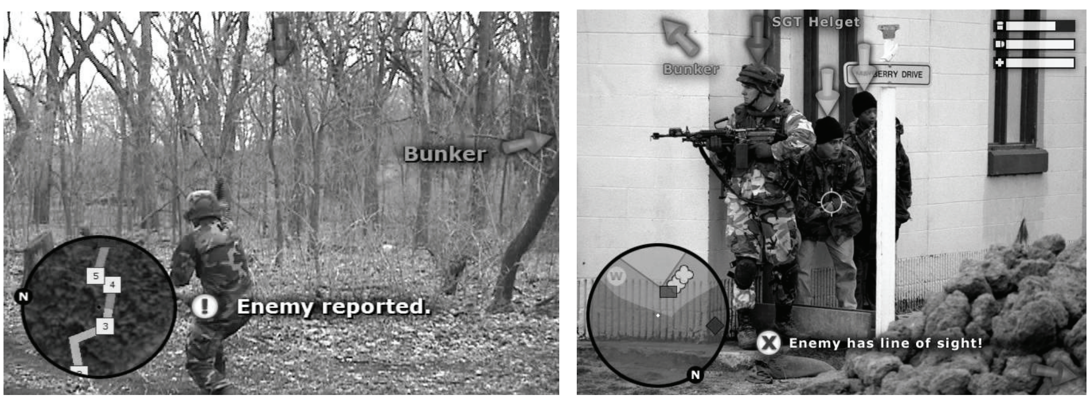


Figure 19-6. Conformal symbology for the ground soldier with a survey map view in the lower left that provides orientation of threats or waypoints in space (from Milbert, 2005, used with permission).

One finding throughout the literature is the benefit of locating conformal imagery or intuitive icons (e.g., virtual sign posts, synthetic grids, threats, safe path in the sky, horizon, ground, other aircraft, or landing field) in a geo-stabilized mode placed where they actually are in space. Wickens (2007) contends that conformal HUD imagery is more readily understood because the earth-referenced information is easily fused by the pilot - simplifying the Level 1 and Level 2 SA steps – because the outside world object moves with the imagery and the pilot intuitively links the two together (Yeh, Wickens, and Seagull, 1998). Doing so intuitively transforms the cockpit-derived *meta-knowledge* to earth-referenced data so the pilot is not required to derive their real location in space. A simulation study by Rogers, Asbury and Haworth (1999) demonstrated the efficacy of this concept by presenting earth-referenced symbology for waypoints and engagement areas (EA) on a head-tracked HMD with dramatic improvements in pilot performance. Using experienced AH-64 Apache aviators, their study demonstrated an impressive 300% improvement (287 feet vs. 878 feet) in waypoint accuracy and 430% improvement (262 feet vs. 1,130 feet) in landing point accuracy, a 12,000% improvement (14 feet vs. 1,666 feet) in engagement area fire sector identification accuracy with a *55 to 69%* reduction in overall workload (when using the waypoint symbols and EA symbols, respectively).

In work intended for general aviation application, Theunissen et al., (2005) created a predictive pathway in the sky to show the pilot the future position of the aircraft, making this part of his Level 3 SA process easier. Rogers, Asbury, and Szoboszlay, (2003) took this a step further and created a Flight Path Marker to overcome some of the problems with previous "pathway in the sky" efforts which only showed a projected tangent to the current (usually) curving flight path, not the actual predicted flight path itself. In experiments with experienced helicopter pilots, Rogers and his colleagues validated their approach with statistically significant improvements in: 1) minimizing the number of ground strikes, 2) mean roll direction changes, and 3) mean overall workload rating, clearly showing how pre-processing the flight path data allows more of the pilot's energies to be spent flying the aircraft than thinking about the future position.

In this same study, Rogers, Asbury, and Szoboszlay, (2003) displayed a set of concentric rings that were always oriented parallel to the horizon, located at a virtual separation of 50 feet in elevation and displayed out to the edges of the 60° HMD field-of-view. The rings provided the pilot a simple method of determining his aircraft altitude and attitude relative to level ground, avoiding the traditional ground versus figure confusion ("Am *I* tilted or is the ground?"). Their findings were significant in terms of: 1) touchdown groundspeed, 2) touchdown pitch error, and 3) overall workload rating. Because the rings were displayed as a wide field-of-view image, it also helped stimulate the ambient visual mode, the peripheral process which does not require conscious attention of the pilot, by pre-processing key pieces of flight imagery such as orientation relative to the horizon. Their results conclusively demonstrated that the pilots maintain their situation awareness with a reduction in workload.

A compounding factor derives from the pilot's seating position in the aircraft that may be a few meters removed from the actual location of a nose-mounted sensor, a situation that is exacerbated in low level flight or when the pilot turns his head 90° to the left or right (Antonio, 2008). One concept investigated on the – since cancelled – RAH-66A Comanche helicopter program was to display a stabilized wireframe outline of the forward aircraft structure. This was felt to be especially beneficial when the pilot was relying on the HMD for all imagery such as flying at night by giving the pilot a sense of orientation relative to the front of the aircraft.

Albery (2007) reported on a multi-sensory cueing system for fixed-wing aircraft called the Spatial Orientation Retention Device (SORD) where the pilot is provided visual, tactile and auditory cues. On- and off-boresight HMD symbology using the Non-Distributed Flight Reference (see also Jenkins, 2003; Jenkins, Turling and Brown, 2003) gives the pilot innovative and intuitive visual references to determine flight attitude using a relatively narrow field-of-view display. Tactile cueing augments the visual cues via torso-mounted tactors so as to convey aircraft attitude.[19] Out of normal attitudes are communicated by localized cueing on the pilot's chest. Further cues are provided with a 3-D audio system which indicates right or left banking. Combined with the Disorientation Analysis and Prediction System and EEG data, the SORD takes advantage of the multiple human sensor modalities to enhance situation awareness for the pilot, while reducing workload. As of this writing, the SORD has been transitioned to a Rotary-Wing Brownout program.

[19] See Albery (2006) and McGrath, et al, (2004) for a description of the Tactical Situation Awareness System (TSAS).

Measuring Workload and Situation Awareness in Real Time

The fast pace of modern aviation requires the pilot to remain engaged in key tasks that contribute to achieving all three levels of SA. Traditional methods of measuring SA and workload such as efficiency ratings, external observations of experts or self-evaluation can be tainted by bias and are certainly not conducted in real-time. Delayed or after-action indication of cognitive overload may be inadequate to capture time-sensitive loss of SA and to act upon it proactively to ensure mission success or to save lives. Researchers have investigated the use of neural and psychophysiological measures such as eye behavior (pupil diameter, blink and gaze), electroencephalography (EEG), heart rate, galvanic skin response (GSR), and functional near infrared imaging (fNIR) to identify cognitive states of workload, task engagement and fatigue (Craven et al., 2006; Schnell, Keller and Macuda, 2007; Wickens and McCarley 2008). Correlating these measures with their respective cognitive states could provide important benefits in training and flight. In the 1990s, the U.S. Air Force attempted to use EEG signals as a means to control complex aviation systems (Tepe-Nasman, Calhoun and McMillan, 1997[20]). As tantalizing as it appeared, it was felt by some to be ambitious for the time. A more direct approach may be to use these complex signals as operator status indicators and as inputs to a closed loop assessment-mitigation process.[21] This could provide an indication of problems such as cognitive overload (or *underload*[22]), fatigue, disorientation or a missed attentional cue and precisely when this occurred. The goal is to ensure that auditory, visual or tactile cueing will grab or channel the pilot's attention so that we avoid "inattention blindness" or the effect of "looked-but-failed-to-see." It may be possible to detect this change blindness using real-time measures of psychophysiological responses, because it is this lack of noticing – or change blindness *blindness* (Yeh, Wickens, and Seagull, 1998) – that is one of the first steps in the breakdown of the SA cycle.

Eye metrics such as pupil size, eye movements and blinks have been used to identify cognitive states such as engagement in problem solving, driving, and alertness/fatigue (Marshall, 2007a; Tsai et al., 2007). Beatty (1982) reviewed the task-evoked pupillary dilation data, finding a strong correlation of workload or cognitive processing load and the increase in pupil diameter that occurs within 100 and 200 ms of the task onset. He showed that the magnitude of pupil dilation is directly correlated with the magnitude of the effort required to address the task with the slope of the diameter increase directly correlated with task difficulty. He also found pupil dilations for near threshold detection of auditory and visual cueing signals as well as peak amplitudes of pupil dilation with memory tasks (increasing up to an asymptote of 7 digits), language related tasks (grammatical reasoning was found to be most difficult), arithmetic reasoning (difficult multiplications were found to be most demanding and resulted in the largest pupil increase) and difficult sensory discrimination tasks.

Marshall (2007a; 2007b) has developed the Index of Cognitive Activity (ICA) to effectively determine levels of cognitive workload from high-frequency increases in pupil dilation. The attractive aspect is its insensitivity to increases in light level that might be found in an aviation environment and would thus make the ICA compatible with an operational HMD. Marshall (2007a) also combined the ICA with other eye metrics such as pupil information, eye movements and blink status to determine cognitive states during problem solving (relaxed versus engaged), driving (focused versus distracted attention) and visual search (alert versus fatigued). She found that combining these measures made for a more robust assessment across individuals in the study rather than relying on any one metric individually.

---

[20] This was, perhaps, a tribute to the 1982 film, *Firefox*, in which the aircraft is controlled by the pilot's EEG-interpreted thoughts.

[21] This is the focus of DARPA's Augmented Cognition (AugCog) program. "The new field of augmented cognition takes psychophysiological measurement to the next level by integrating continuous monitoring into closed-loop systems. By using the operator states as inputs, adaptively automated systems respond to user overload or under load, and react appropriately" (Berka et al., 2007).

[22] Cognitive underload refers to the state where the pilot is not fully engaged in critical tasks, possibly resulting in complacency and a failure to notice important events.

Synchronizing observed behavior with EEG data such as the time-based increase or decrease of the different brain wave rhythms[23] or various ratios of their values have allowed researchers to identify cognitive states including workload, distraction, drowsiness and training levels. Wickens and McCarley (2008) found a correlation of workload with increases in Theta band and decreases in Alpha band. Using a dense EEG sensor array (128 electrodes), Schnell, Keller and Mancuda (2007) found that the ratio of Beta/Alpha is indicative of cognitive workload and that Theta waves measured in the midline correlate with monitoring and memory tasks. Berka and her colleagues (Berka et al., 2004; 2006; 2007) have reported success in assessing cognitive states using a sparser EEG array (three to twelve sensors). Real-time EEG markers have also been found which directly correlate with levels of visual workload and situation awareness (Berka et al, 2007). Still other research has found specific EEG markers of spatial disorientation (Albery, 2007; Viirre, et al., 2006).

The N1 and P3 Event Related Potentials (ERP)[24] have been shown to be associated with the allocation of attentional resources and perceptual-cognitive resources, respectively (Hancock, 2007). Because it is often observed after an “oddball” sensory stimulus, the P3 (resulting from an auditory, tactile or visual stimulus and strongest when the stimulus occurs in an attended sensory modality – Driver and Spence, 1998) is thought to be related to unexpected occurrences and has been used successfully in Rapid Serial Visual Presentation (RSVP) (Gerson, Parra and Sajda, 2006) to triage large imagery mosaics. Further, because of the oddball stimulus correlation, the P3 may be applicable in aviation where the pilot observes something but because he is attending to other duties, may not notice or react to it in time. If the system recognizes the characteristic P3 signal without an accompanying pilot reaction, it may be possible to alert the pilot to the presence of an un-attended object or event that requires attention. Peterson, Allison, and Polich (2006) found that workload-related Alpha signals have an inverse correlation with P3 signal during computer games of various workload levels and they recommend monitoring these various spectral signatures simultaneously to improve accuracy. Trejo, et al. (2006) studied the impact of mental fatigue on EEG rhythms and found an increase in frontal Theta and parietal Alpha power, though their ERP (N1, P2 and P3) data were inconclusive. By monitoring various ERPs, it may be possible to use the information to monitor operator state – the intended goal of AugCog – to determine what cue or event was attended to, or whether it was missed, and when.

Using Real-Time Measures to Improve Training Performance

During a training session, an individual requires mental effort to acquire the skills necessary to complete the task. However, as they go through the three levels of skill development,[25] they require less and less effort to do so until they reach the point of automaticity.[26] Stevens, Galloway and Berka (2006) demonstrated that as trainees acquired expertise, their engagement and workload decreased as noted by their EEG patterns. Berka et al. (2006) noted differences in the Theta band EEG signals between individuals who made correct and incorrect decisions thus providing a potential metric to determine true skill level. Marshall, Pleydell-Pearce and Dickson (2002) found that as individuals gain proficiency in a task, they may change their strategy of where, when, and for how long they gaze at various instruments. By measuring the gaze point during the training sessions, it can be determined when the individual gains insight and understanding of the structure of the task and develops a new strategy which may

---

[23] Brain wave rhythms are divided into: Delta (0.5 to 3 Hertz [Hz]), Theta (4 to 7 Hz), Alpha (8 to 12 Hz), Beta (13 to 30 Hz), and Gamma (greater than 30 Hz), (Scerbo, Freeman, Mikulka, Parasuraman, DiNocero, and Prinzel 2001).

[24] Event Related Potentials (ERP) are non-volitional EEG responses that generate a voltage – either negative (N) or positive (P) occurring within a specific timeframe – after an observed event. The P3 (also called the P300) is a positive voltage that occurs roughly 300 milliseconds after a sensory stimulus and the N1 is a negative voltage that occurs roughly 100 milliseconds after a stimulus.

[25] These are: the *initial learning or cognitive stage* where the trainee assembles new knowledge, the *associative stage* where the trainee begins to automate the learned steps and the *autonomous stage* where the trainee executes the steps with minimal conscious mental effort.

[26] See also “chunking,” a mnemonic device sometimes used to enable the intermediate learning steps.

indicate a change in their level of expertise. For example, if the pilot changes from a general sweep of all cockpit instrumentation, and starts relying more on the *predictive* instruments (Wickens, 2007; Endsley, 2000), it may be a sign that they have reached a new level of expertise in the task. All aspects of training may be affected by the ability to acquire real-time assessments of vigilance, workload, fatigue, engagement, and the ability to assess task proficiency status by observing an increase in SA, a drop in workload or a change in strategy. Rather than relying on outcome-based performance measures, which may inaccurately reflect skill level, it allows the training curriculum to be assessed for statistical timelines and effectiveness. These could be applied during training scenarios to ensure that it is having maximum impact on the trainee without the adverse "cognitive states such as distraction, boredom, confusion and frustration" (Stevens, Galloway and Berka, 2006) by capturing real-time EEG indicators such as engagement (involving information-gathering, visual scanning, and sustained attention) and workload index (which increases with working memory load and with increasing difficulty level of mental arithmetic). It may also be possible to use psychophysiological monitoring on test subjects to evaluate display modalities, symbology, and procedures and be able to capture – in real time – the points during the presentation where workload is high and situation awareness is low.

Adaptive Automation

Automation in advanced technology is occurring, dictated by the continuous movement towards more complex systems. While this has worked well in areas such as the automotive industry with the automatic transmission and anti-lock braking, it has also had negative consequences in situations where the human is excluded from the loop and serves simply as a system monitor. Doing so can have negative consequences because it engenders a time penalty required for the human to notice, understand and react to an important event as well as: 1) loss of vigilance and increased complacency (by placing too much trust in the automation), 2) loss of SA by becoming a passive observer rather than an active participant, and 3) the changed nature of the information or feedback available to the operator (Endsley, 1996). A newer approach is *adaptive automation*, where the level of automation is dynamically initiated and adjusted either by the system or by the operator to optimize engagement or vigilance without producing cognitive overload. Here, the support is enabled when workload is high or when some impairment becomes evident (Hancock, 2007); similarly to the way a pilot would off-load tasks to another crewmember.[27] Traditional automation rigidly changes the role of the user from that of an active participant to that of a passive observer, potentially disengaging them and opening up the possibility that they might miss key events or signals or critical warning signs. Adaptive automation, however, changes the paradigm by enabling assistive automation only when necessary.

While the details of how to enable adaptive automation in the cockpit is beyond the scope of this chapter, it would appear that the HMD can play a key role as part of the system, perhaps acting as the portal through which automation-level-dependent information could *flow to* the pilot (in the form of cognitively pre-digested cues and symbology) and simultaneously, key psychophysiologically-measured operator status data (such as EEG, ERP or eye metrics) could *flow back to* the system (Schnell, 2008). Since the response time between the event and the psychophysiological marker can be on the order of seconds or less, having these real-time indicators could very rapidly invoke the required automation to either immediately reduce pilot workload or take over aspects of the aircraft as necessary. Future research could indicate not only *when* the pilot is overloaded, but *which* of the pilot's resources may be affected, what Scerbo et al. (2001) refer to as Operator Modeling, where an impaired status indicator (from eye metric, EEG or ERP signal) initiates the automated response.

In the studies at Boeing's Phantom Works, information displayed during the simulation would "grey-out" when the pilot subjected himself to a high-g loading in a manner similar to what they would actually experience (Proctor, 1999). In a system equipped with adaptive automation, the aircraft would determine or sense the pilot's

[27] With an accompanying "I've got it" from the automated system.

physiological state as a result of excessive g-loading and simplify or reduce the HMD symbology or, alternatively, take over aircraft control entirely to prevent a catastrophe.

Bonner, Taylor, Fletcher and Miller, (2000) have designed a system called the Cognitive Cockpit intended to adapt to the cognitive state of the pilot by off-loading the more routine flight activities at need. This allows the pilot to focus more energy on the tactical aspects of the situation. The Tasking Interface Monitor ensures that mission goals are maintained and allows the system to assume control of generic tasks that are more rule-based and skill-based.

Albery and colleagues at the Air Force Research Lab have created the Disorientation Analysis and Prediction System (DAPS) as part of the Spatial Orientation Retention Device (SORD) to calculate a "disorientation index" and provide multi-sensor cueing to the pilot that recovery from a non-normal flight attitude may be required should the pilot be disoriented or be unaware of the problem (Albery, 2006, 2007).

Summary

- The HMD provides a unique method of presenting information to the pilot that replicates natural human exploratory behavior, allowing movement of head and eyes outside the limited field-of-regard of typical cockpit displays as the pilot navigates through the environment.
- Situation awareness is the ultimate goal of the display designer. The problem for the pilot is that there is often too much unprocessed data and not enough distilled information to be able to arrive at situation awareness through the information gathering, model making/updating and predicting cycle. Information must be presented in such a way as that it will be easy to understand to make the SA cycle easier and more intuitive, requiring less of the pilot's already-taxed cognitive resources
- HMD symbology should be used to present flight and aircraft status that is not just a re-mapping of the internal cockpit display information but which is cognitively processed so as to provide useful predictive information without cognitive overload and which will allow the pilot to spend more time looking outside the cockpit to reduce the workload associated with the three steps in the situation awareness loop.
- There has been considerable study in the areas of attention, multiple resources and cross-modal integration which can explain how we can sometimes multi-task efficiently, but at some point become cognitively overloaded due to executive control overload. These models can also help identify ways to improve pilot performance using cross-modal cues as notifiers of an event in a complementary sensory modality, such as a 3-D audio cue directing the pilot's attention to a visual event.
- Psychophysiological monitoring (such as eye metrics, respiratory and skin response and EEG or ERP signals) has been shown to accurately measure SA status, fatigue, disorientation, cognitive overload and *underload*, task expertise and correct or incorrect responses in various situations, with the HMD serving as a convenient platform for the sensors. Using these measures as system inputs – the focus of the AugCog program – can provide a real-time understanding of operator status during flight and training.
- Using an operator performance model and real-time psychophysiological measures of the pilot's physical or cognitive state, immediate steps can be taken to allocate or off-load less urgent tasks to the aircraft system or to control the aircraft when the pilot becomes physically or cognitively incapacitated.

From advances in neuroergonomics – the science of understanding the way in which humans perceive information with a look towards improving our interaction with technology in the real world – valuable insights into how the HMD can advance past its current state as an extension of the aircraft display suite can be gained. We can start to improve integration with the aircraft through new developments in symbology, addition of ancillary cueing from tactile or 3-D audio and real-time operator status monitoring where the HMD – now a cognitive prosthesis – provides real-time assistance by closing the loop between the pilot and the aircraft.

References

Albery, W.B. (2006). Development of Multisensory Orientation Technology for Reducing Spatial Disorientation, Final Report, *Report AFRL-HE-WP-TR-2006-0024, Air Force Research Laboratory, Human Effectiveness Directorate, Wright-Patterson Air Force Base, Ohio*

Albery, W.B. (2007). Multisensory cueing for enhancing orientation information during flight. *Aviation, Space and Environmental Medicine*, 78, B186-B190.

Antonio, J.C. (2008). Display content in advanced NVG and HMD systems: a pilot/flight surgeon's concerns, *Proceedings of the SPIE: Head- and Helmet-Mounted Displays XIII: Design and Applications*, 6955, 1-9.

Arbak, C. (1989). Utility Evaluation of a Helmet-Mounted Display and Sight. *Proceedings of the SPIE: Helmet-Mounted Displays*, 1116, 138-141.

Barfield, W., and Furness, T.A. (1995). *Virtual Environments and Advanced Interface Design*. New York: Oxford University Press.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources, *Psychological Bulletin*, 91(2), 276-292.

Berka, C., Levendowski, D.J., Cvetinovic, M.M., Petrovic, M.M., Davis, G., Lumicao, M.N, Zivkovic, V.T., Popovic. M.V., and Olmstead, R. (2004). Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. *International Journal of Human-Computer Interaction*, 17(2), 151-170.

Berka, C., Levendowski, D.J., Davis, G., Whitmoyer, M., Hale, K., and Fuchs, K. (2006). Objective Measures of Situational Awareness Using Neurophysiology Technology. In: Schmorrow, D., Stanney, K., and Reeves, L. (Eds.), *Augmented Cognition: Past, Present and Future.* Arlington VA: Strategic Analysis, Inc.

Berka, C., Levendowski, D., Lumicao, M.N., Yau, A., Davis, G., Zivkovic, V.T., Olmstead, R.E., Tremoulet, P.D., and Craven, P.L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks, *Aviation, Space and Environmental Medicine,* 78(5), B231.

Bertelson, P., and de Gelder, B. (2004). The psychology of multimodal perception. In: Spence, C., and Driver, J. (Eds.), *Crossmodal Space and Crossmodal Attention.* Oxford: Oxford University Press.

Bolia, R.S. (2004). Special issue: Spatial audio displays for military aviation. *International Journal of Aviation Psychology*, 14, 233-238.

Bolia, R.S., D'Angelo, W.R., and McKinley, R.L. (1999). Aurally aided visual search in three-dimensional space. *Human Factors*, 41, 664-669.

Bonner, M.C., Taylor, R.M., Fletcher, K., and Miller, C. (2000). Adaptive automation and decision aiding in the military fast jet domain. In: *CD Proceedings of the Conference on human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium* (154-159). Savannah, GA. Retrieved from: http://www.dtic.mildticasd/sbir/sbir041/srch/af071.pdf

Boyd, J.R. (2007). Patterns of Conflict, edited by Richards, C., and Spinney, C. Retrieved from: http://www.d-n-i.net/dni/john-r-boyd/

Casey, C.J., and Melzer, J.E. (1991). Part-task training with a helmet integrated display simulator system. *Proceedings of the SPIE: Large-Screen Projection, Avionic and Helmet-Mounted Displays*, 1456, 175.

Chapin, W.L., Jost, A.R., Cook, B.A., Cook, B.A., Surucu, F., Foster, S., Bolas, M., McDowall, I., Lorimer, E.R., Zurek, P.M., Desloge. J.G., Beaudoin, R.E., Shinn-Cunningham, B., and Durlach, N. (2003). Concept and technology exploration for transparent hearing, project final report. *Scorpion Audio Team Final Report.* Natick, MA: Natick Soldier Center.

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustic Society of America*, 25, 975-979.

Cole, E., and Matthews, M.K. (1999). Cognitive prosthetics and telerehabilitation: Approaches for the rehabilitation of mild brain injuries. In: Murner, J., and Ettlin, T.M., (Eds.), *HSW-Distorsion and Leichte Traumatische Hirnverletzung Behandlungskonzepte*. Basel, Switzerland.

Craven, P.L., Belov, N. Temoulet, P., Thomas, M., Berka, C., Levendowski, D., and Davis, G. (2006). Cognitive workload gauge development: Comparison of real-time classification methods. *Proceedings of the 2nd International Conference on Augmented Cognition.*

Driver, J., and Spence, C. (1998). Cross-modal links in spatial attention, *Philosophical Transactions of the Royal Society of London B,* 353, 1319-1331.

Endsley, M.R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32-64.

Endsley, M.R. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors*, 37(1), 65-84.

Endsley, M.R. (1996). Automation and Situation Awareness. In: Parasuraman, R., and Mouloua, M. (Eds.), *Automation and Human Performance: Theory and Applications*. Mahwah, New Jersey, Lawrence Erlbaum.

Endsley, M.R. (2000). Theoretical underpinnings of situation awareness: A critical review. In: Endsley, M.R., and Garland, D.J. (Eds.), *Situation Awareness Analysis and Measurement*. Mahway, NJ: Lawrence Erlbaum Associates.

Endsley, M.R., (1993). Situation awareness and workload: Flip sides of the same coin. *Proceedings of the 7th International Symposium on Aviation Psychology.* Columbus, OH: Ohio State University.

Endsley, M.R., and Hoffman, R.R. (2002). The Sacagawea Principal. *Proceedings of the IEEE: Intelligent Systems*, November, 80-85.

Flanagan, P., McAnally, K.I., Martin, R.L., Meehan, J.W., and Oldfield, S.R. (1998). Aurally and visually guided visual search in a virtual environment, *Human Factors*, 40, 461-468.

Geiselman, E.E., and Osgood, R.K. (1994). Utility of off-boresight helmet-mounted symbology during a high angle airborne target acquisition task, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays and Symbology Design Requirements*, 2218, 328-338

Gerson, A.D., Parra, L.C., and Sajda, P. (2006). Cortically-coupled computer vision for rapid image search, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2), 174-719.

Gibson, J.J. (1986). *The Ecological Approach to Visual Perception.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Hameed, S., Jayaraman, S., Ballard, M., and Sarter, N. (2007). Guiding visual attention by exploiting crossmodal spatial links: An application in air traffic control, *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, 220-224. Santa Monica, CA: Human Factors and Ergonomics Society.

Hancock, P.A. (2007). On the process of automation transition in multitask human-machine systems, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 37, 586.

Hancock, P.A., and Szalma, J.L. (2003). The future of neuroergonomics, *Theoretical Issues in Ergonomic Science*, 44, 238-249.

Hansen, A.J., Rybacki, R.M., and Smith W.G. (2006). Synthetic vision in the cockpit: 3-D systems for general aviation. Retrieved on 6 July 2008 from MetaVR, Inc., web site: http://www.metavr.com/technology/papers/syntheticvision.html

Helmetag, A., Halbig, C., Kubbat, W., and Schmidt, R. (1999). Cognitive mapping based on synthetic vision. *Proceedings of the SPIE: Enhanced and Synthetic Vision 1999*, 3691, 48.

Hopper, D.G. (2000). Reality and surreality of 3-D displays: Holodeck and beyond. Invited Keynote Paper in the *Electronic Information Display Conference,* Society for Information Display ExCel, England UK, November, p. 21-23.

Houtsma, A.J.M. (2003). Nonspeech audio in helicopter aviation. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report 2004-03.

Jenkins, J.C. (2003). Development of helmet-mounted display symbology for use as a primary flight reference, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays VIII: Technologies and Applications*, 5079, 333-345.

Jenkins, J.C., Thurling, A.J., and Brown, B.D. (2003). Ownship status helmet-mounted display symbology for off-boresight tactical applications, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays VIII: Technologies and Applications*, 5079, 346-360.

Jenkins, J.C., Sheesley, D.G., and Bivetto, F.C. (2004). Helmet-mounted display symbology for enhanced trend and attitude awareness, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays IX: Technologies and Applications,* 5442, 164-178.

Kalawsky, R.A. (1993). *The Science of Virtual Reality and Virtual Environments*. New York: Addison-Wesley Publishing Company.

Keller, K., and Colucci, D. (1998). Perception in HMDs: What is it in head-mounted displays (HMDs) that really make them all so terrible?, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays III*, 3362, 46-53.

Koch, I. (2001). Automatic and intentional activation of task sets, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27, 1474-1486.

Kocian, D.F. (1987). Design considerations for virtual panoramic display (VPD) helmet systems. *AGARD Conference Proceedings No. 425, The Man-Machine Interface In Tactical Aircraft Design and Combat Automation*, 22-1.

Kuipers, B. (1978). Modeling Spatial Knowledge, *Cognitive Science*, 2, 129-153.

Leedom, D.K. (2001). Sensemaking Symposium: Final Report. Command and Control Research Program of the Assistant Secretary of Defense for Command, Control, Communications and Intelligence, Vienna, Virginia, Retrieved 23-25 October 2001 from: http://www.dodccrp.org/events/2001_sensemaking_symposium/index.htm

Leibowitz, H.W., Shupert, C.L., and Post, R.B. (1985). Emergent techniques for assessment of visual performance. Retrieved 27 July 2008 from: http://www.nap.edu/openbook/POD227/html

Marshall, S.P. (2007a). Identifying cognitive state from eye metrics, *Aviation, Space and Environmental Medicine*, 78(5), B165.

Marshall, S.P. (2007b). Measures of attention and cognitive effort in tactical decision making. In: Cook, M., Noyes, J., and Masakowski, Y. (Eds.), *Decision Making in Complex Environments.* Burlington, VT: Ashgate Publishing Company.

Marshall, S.P., Pleydell-Pearce, C.W., and Dickson, B.T. (2002). Integrating psychophysiological measures of cognitive workload and eye movements to detect strategy shifts. IEEE Computer Society, *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS).*

McGrath, B.J., Estrada, A. Braithwaite, M.G., Raj, A.K., Rupert, A.H. (2004). Tactile Situation Awareness System Flight Demonstration Final Report, Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report 2004-10.

McIntire, J.P., Havig, P.R., Watamaniuk, S.N.J., and Gilkey, R.H. (2008). Aurally aided visual search performance in a dynamic environment, *Proceedings of the SPIE: Head- and Helmet-Mounted Displays XIII, Design and Applications*, 6955, 695501.

Melzer, J.E. (2008). Toward the HMD as a cognitive prosthesis. *Proceedings of the SPIE: Head- and Helmet-Mounted Displays XIII, Design and Applications*, 6955, 69550L-69550L-9.

Melzer, J.E. (2006). Head-Mounted Displays. In: Spitzer, C.R. (Ed.), *The Avionics Handbook*, 2nd Ed. Boca Raton, FL: CRC Press.

Melzer, J.E., and Porter, J. (2008). Helmet-mounted display upgrade for the US Army's AVCATT simulation program, *Proceedings of the SPIE: Head- and Helmet-Mounted Displays XIII, Design and Applications*, 6955, 4-8.

Melzer, J., and Simons, R. (2002). Displaying the Battlespace Environment: Evolution of a Head-Mounted Display. Presented at *The Interservice/Industry Training, Simulation and Education Conference (I/ITSEC), Orlando, FL.*

Merryman, R.F.K. (1994). Vista Sabre II: integration of helmet-mounted tracker/display and high off-boresight missile seeker into F-15 aircraft, *Proceedings of the SPIE: Helmet-and Head-Mounted Displays and Symbology Design Requirements*, 2218, 173.

Moffitt, K.W. (1997). Designing HMDs for viewing comfort. In: J.E. Melzer, K.W. Moffitt (Eds.), *Head Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Milbert, R.L. (2005). Primordial Soldier 2.0 – Routing soldiers around combat obstacles and differentiating friend from foe. *Proceedings of the Second Americas Display Engineering and Applications Conference,* Portland, OR.

Parasuraman, R. (2007). Neuroergonomics of visual cognition: research and applications. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* 1311. Santa Monica, CA: Human Factors and Ergonomics Society.

Parasuraman, R. (2003). Neuroergonomics: Research and practice. *Theoretical Issues in Ergonomics Science*, 4, 5-20.

Peterson, D.A., Allison, B.Z., and Polich, J. (2006). Alpha power complements P300 in cognitive workload assessment, *2006 Neuroscience Meeting, Society for Neuroscience*: Atlanta, GA.

Piaget, J. (1952). *The origins of intelligence in children*. New York: Longman.

Pouget, A., Deneve, S., and Duhamel, J. (2004). A computational neural theory of multisensory spatial representations. In: Spence, C., and Driver, J. (Eds.), *Crossmodal Space and Crossmodal Attention.* Oxford: Oxford University Press.

Proctor, P. (1999). Helmet displays boost safety and lethality, *Aviation Week and Space Technology*, 150(5), 81-83.

Rash, C.E. (2001). Introductory overview. In: Rash, C.E. (Ed.), *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft*. Bellingham, WA: SPIE Press,

Rogers, S. P., Asbury, C. N., and Haworth, L. A. (1999). Evaluation of earth-fixed HMD symbols using the PRISMS helicopter flight simulator, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays IV*, 3689, 54-65.

Rogers, S.P, Asbury, C.N., and Haworth, L.A. (2001). Evaluations of new symbology for wide-field-of-view HMDs, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays VI*, 4361, 213-224.

Rogers, S.P., Asbury, C.N., and Szoboszlay, Z.P. (2003). Enhanced flight symbology for wide field-of-view helmet-mounted displays, *Proceedings of the SPIE: Helmet- and Head-Mounted Displays VIII: Technologies and Applications*, 5079, 321-332.

Rogers, S.P., and Asbury, C.N. (2007). Obstacle display for hover in degraded visual conditions. *Final Report: SBIR Topic No. A04-078, Report: RDECOM TR06-D-54, US Army Research, Development and Engineering Command.* Moffett Field, CA: NASA-Ames Research Center.

Rubinstein, J.S., Meyer, D.E., and Evans, J.E. (2001). Executive control of cognitive processes in task switching, *Journal of Experimental Psychology: Human Perception and Performance,* 27, 763-797.

Scerbo, M.W., Freeman, F.G., Mikulka, P.J., Parasuraman, R., DiNocero, F., and Prinzel, L.J. (2001). The efficacy of psychophysiological measures for implementing adaptive technology. *NASA Center for AeroSpace Information report*. NASA/TP-2001-211018.

Schnell, T. (2008). Operator Performance Laboratory, University of Iowa, internal presentation to Rockwell Collins.

Schnell, T., Keller, M., and Macuda, T. (2007). Application of the Cognitive Avionics Tool Set (CATS) in airborne operator state classification, *Proceedings of the 4th International Conference on Augmented Cognition.*

Stevens, R., Galloway, T., and Berka, C. (2006). Integrating EEG Models of Cognitive Load with Machine Learning Models of Scientific Problem Solving, *Proceedings of the 2nd International Conference on Augmented Cognition*, 55-65.

Smith, D.J. (2006). Situation(al) awareness (SA) in effective command and control. Retrieved 12 July 2008 from: http://www.smithsrisca.demon.co.uk/situational-awareness.html

Spence, C., and Driver, J. (1997). Cross-modal links in attention between audition, vision and touch: implications for interface design, *International Journal of Cognitive Ergonomics*, 1, 351-373

Still, D.L., and Temme, L.A. (2001). Oz: A human-centered computing cockpit display. *Proceedings of the Interservice/Industry Training, Simulation and Education Conference.* Orlando FL, 297-307.

Still, D.L., and Temme, L.A. (2008). Hover training display: Rationale and implementation. *Display Technologies and Applications for Defense, Security, and Avionics II, Proceedings of the SPIE,* 6956, 695607-15.

Tepe-Nasman, V., Calhoun, G.L., and McMillan, G.R. (1997). Brain-actuated control and HMDs. In: Melzer, J.E., and Moffitt, K.W. (Eds.), *Head Mounted Displays: Designing for the User*. New York: McGraw-Hill.

Theunissen, E., Koeners, G.J.M., Roefs, F.D., Rademaker, R.M., Kinkins, R.D., and Etherington, T.J. (2005). Guidance, situation awareness and integrity monitoring with an SVS+EVS. Paper AIAA 2005-6441, *American Institute of Aeronautics and Astronautics Guidance, Navigation and Control Conference and Exhibition,* 15-18 August, San Francisco, CA.

Trejo, L.J., Kochavi, R., Kubitz, K., Montgomery, L.D., Rosipal, R., and Matthews, B. (2006). EEG-based estimation of mental fatigue. Retrieved 25 July 2008 from: http://publications.neurodia.com/Trejo-et-al-EEG-Fatigue-2006-Manuscript.pdf

Tsai, Y., Viirre, E., Strycacz, C., Chase, B., and Jung, T. (2007). Task performance and eye activity: Predicting behavior relating to cognitive workload, *Aviation, Space and Environmental Medicine*, 78(5), B176.

Uhlarik, J., and Comerford, D.A. (2002). A review of situation awareness literature relevant to pilot surveillance functions. *Final Report: DOT/FAA/AM-02/3, Office of Aerospace Medicine.*

United States Army (2003). Operational Requirements Document (ORD) Update for Air Warrior (AW) ACAT III Prepared for Milestone C Decision, Change 2, January 2002.

Weick, K.E., Sutcliff, K.M., Obstfeld, D. (2005). Organizing and the Process of Sensemaking, *Organizational Science*, 16(4), 409-421

Wickens, C.D. (1980). The structure of attentional resources. In: Nickerson, R.S. (Ed.), *Attention and Performance*, 8, 239-257. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Wickens, C.D. (1984). Processing resources in attention. In: Parasuraman, R., and Davies, D.R. (Eds.), *Varieties of Attention*. San Diego, CA: Academic.

Wickens, C.D. (2002a). Multiple resources and performance prediction, *Theoretical Issues in Ergonomics Science*, 3, 159-177.

Wickens, C.D. (2002b). Spatial Awareness Biases. Moffett Field, CA: NASA Ames Research Center. Technical Report AHFD-02-6/NASA-02-4.

Wickens, C.D. (2007). Aviation. In: Durso, F. (Ed.), *Handbook of Applied Cognition: Second Edition*. New York: John Wiley and Sons.

Wickens, C.D., and McCarley, J.S. (2008). *Applied Attention Theory*. Boca Raton, FL: CRC Press, Taylor and Francis Group.

Viirre, E., Wing, S., Huang, R., Strychacz, C., Koo, C., Stripling, R., Cohn, J., Chase, B., and Jung, T. (2006). EEG Markers of Spatial Disorientation. *Proceedings of the 2nd International Conference on Augmented Cognition.*

Wenzel, E.M. (1992). Localization in virtual acoustic displays, *Presence*, 1, 80-107.

Wightman, F., and Kistler, D. (1998). Of Vulcan ears, human ears and earprints, *Nature Neuroscience*, 1, 337-339.

Yeh, M., Wickens, C.D., and Seagull, F.J. (1998). Effects of frame of reference and viewing condition attentional issues with helmet-mounted displays. Aviation Research Lab, Institute of Aviation, University of Illinois at Urbana-Champaign. Technical Report ARL-98-1/ARMY-FED-LAB1.

Yona, Z., Weiser, B., and Hamburger, O. (2004). Day/Night ANVIS/HUD-24 (Day HUD) flight test and pilot evaluations. *Helmet- and Head-Mounted Displays IX: Technologies and Applications, Proceedings of the SPIE*, 5442, 225- 235.

Abbreviations and Acronyms

2AFC Two-Alternative, Forced-Choice
2-D Two-Dimensional
3-D Three-Dimensional
4-D Four-Dimensional
µM Micron
° Degree
λ Wavelength
ABR Auditory Brainstem Response
ACCES Attenuating Custom Communication Earpiece System
ACES Advanced Concept Ejection Seat
ACG Angle Closure Glaucoma
ACGIH American Conference of Governmental Industrial Hygienists
ACH Advanced Combat Helmet
ACT-R Adaptive Control of Though-Rational
ADAS Advanced Distributed Aperture System
ADHD Attention Deficit Hyperactivity Disorder
AFQT Armed Forces Qualification Test
AFP Advanced Flat Panel
AFRL (U.S.) Air Force Research Laboratory
AGC Automatic Gain Control
AGSM Anti-G Straining Maneuver
AHAAH Auditory Hazard Assessment Algorithm for the Human Ear
AHMD Advanced Helmet Mounted Display
AI Articulation Index
AIHS Aircrew Integrated Helmet System
AIS Abbreviated Injury Scale
AKE Autokinetic Effect
AM Active Matrix; Amplitude Modulation
AMEL Active Matrix Electroluminescent
AMLCD Active Matrix Liquid Crystal Display
AMOLED Active Matrix Organic Light-Emitting Diode
AMP Ampere
ANR Active Noise Reduction
ANSI American National Standards Institute
ANVIS Aviator's Night Vision Imaging System
AO Atlanto-Occipital
AOI Area of Interest
APL Aeromedical Policy Letter
AR Augmented Reality
ARL Army Research Laboratory
ARMD Age-Related Macular Degeneration
ARSA Arc-Segmented Attitude Reference
ARU Aircraft Retained Unit
ASHA American Speech-Language-Hearing Association
ASIC Application-Specific Integrated Circuit
ASL Above Sea Level
ASRS Aviation Safety Reporting System
ASSIST Automated System of Self Instruction for Specialized Training
ASTMP Army Science and Technology Master Plan
ATC Air Traffic Control
AVCATT-A Aviation Combined Arms Tactical Trainer – Aviation Reconfigurable Manned Simulator
AVCN Anterior Ventral Cochlear Nucleus
BAT Brightness Acuity Tester
BMLD Binaural Masking Level Difference
BTE Behind-the-Ear
C&HPS Communication and Hearing Protection Systems
CABS Cockpit Air Bag System
CAD Computer-Aided Design
CAE Computer-Aided Engineering; Combat Arms Earplug
CANS Central Auditory Nervous System
CAP Compound Action Potential
CB Chemical-Biological; Critical Band
CBT Core Body Temperature
CCRP Command and Control Research Program
CCTT Close Combat Tactical Trainer
CEP Communications Earplug
CEPS Communication Enhancement and Protection System
CF Characteristic Frequency
CFF Critical Flicker Fusion
CG Center-of-Gravity
CIC Completely In-the-Canal
CIE Commission Internationale de l'Eclairage or International Commission on Illumination
CIPIC Center for Image Processing and Integrated Computing
CL Contact Lens
CM Center-of-Mass, Centimeter, Cochlear Microphonis
CMNR Committee on Military Nutrition Research
CNS Central Nervous System

CO	Carbon Monoxide
COTS	Commercial Off-the-Shelf
CPD	Cycles Per Degree
CPR	Conditioned Position Responding
CPS	Cycles Per Second
CR	Critical Ratio; Contrast Ratio
CRE	Control Reversal Error
CRT	Cathode-Ray-Tube
CSF	Contrast Sensitivity Function
CVC	Combat Vehicle Crewman
CY	Cycle
DA	Department of the Army (U.S.)
DARPA	Defense Advanced Research Projects Agency
DASH	Display and Sight Helmet
DB	Decibel (dB)
DBA	Decibel, A-Weighted
DBP	Decibel, Peak
DC	Direct Current
DCN	Dorsal Cochlear Nucleus
DGPS	Differential Global Positioning Satellite
DHTVS	Drivers Head Tracked Vision System
DL	Difference Limen
DLP	Digital Light Processing
DMD	Digital Mirror Display
DOD	Department of Defense
DRL	Differential Reinforcement of Low Response Rates
DS	Directionally Selective
DSHEA	Dietary Supplement Health and Education Act
DVE	Driver's Vision Enhancer
EAL	Energy Absorbing Liner
EC	Equalization-Cancellation
ECD	Eye Clearance Distance
EDM	Electronic Data Manager
EEG	Electroencephalography; Electroencephalogram
EL	Electroluminescent
EM	Electromagnetic; Energetic Masking
ENVG	Enhanced Night Vision Goggles
EPE	Exit Pupil Expander
EPIC	Executive Process/Interactive Control
ERB	Equivalent Rectangular Bandwidth
ERP	Event-Related Potential
ETI	Endotracheal Intubation
F	Frequency
FA	Factor Analysis
FAA	Federal Aviation Administration
FBCB2	Force XXI Battle Command, Brigade and Below
FCS	Future Combat System
FD	Far-Domain
FDA	Federal Drug Administration
FED	Field Emission Display
FFW	Future Force Warrior
FIR	Finite Impulse Response
FL	Foot-Lambert
FLC	Ferroelectric Liquid Crystal
FLCD	Flexible Liquid Crystal Display
FLIR	Forward-Looking Infrared
FM	Frequency Modulation
FMRI	Functional Magnetic Resonance Imaging
FMVSS	Federal Motor Vehicle Safety Standards
FOLED	Flexible Organic Light Emitting Diode
FOM	Figure of Merit
FOV	Field-of-View
FP	Flat Panel
FPD	Flat Panel Display
FPS	Feet Per Second
FSC	Field-Sequential Color
FSXXI	Flight School XXI
FY	Fiscal Year
G	Force of Gravity; Gram-of-Force
G-LOC	Gravity-Induced Loss of Consciousness
GEN	Generation
GHZ	Gigahertz
GLV	Grating Light Valve
GPS	Global Positioning System
H	Horizontal
HACE	High Altitude Cerebral Edema
HAPE	High Altitude Pulmonary Edema
HATS	Head and Torso Simulator
HAV	Hand-Arm Vibration
HDD	Head-Down Display
HDTV	High Definition Television
HDU	Helmet Display Unit
HEA	Head Equipment Assembly
HF	Human Factors
HFE	Human Factors Engineering
HGU	Head Gear Unit
HIDSS	Helmet Integrated Display Sighting System
HIT	Human Interface Technology Lab
HL	Hearing Level
HMCS	Helmet Mounted Cueing System
HMD	Head-/Helmet-Mounted Display
HMDS	Helmet Mounted Display System
HMI	Human-Machine Interface
HMMWV	High Mobility Multi-Wheeled Vehicles
HMS	Helmet-Mounted Sight
HMW	Head-Supported Weight

HPD	Hearing Protection Device
HRTF	Head-Related Transfer Function
HRU	Helicopter Retained Unit
HSM	Head-Supported Mass
HSS	Helmet Subsystem
HUD	Head-Up Display
HVI	Helmet-Vehicle Interface
HWD	Head-Worn Display
HZ	Hertz
I	Intensity; Stimulus Magnitude
I^2	Image Intensification
I^2CCD	Image-Intensified Charge-Coupled Device
I^2T	Image-Intensified Tubes
IAP	Instrument Approach Procedures
IBU	Inshore Boat Unit
ICA	Index of Cognitive Activity
ICAD	International Community for Auditory Displays
ICC	Interaural Cross Correlation
ICWAA	Integrated Caution, Warning and Advisory Annunciator
ID	Discrimination Index
IDH	Integrated Display Helmet
IEC	International Electrotechnical Commission
IED	Improvised Explosive Device
IERW	Initial Entry Rotary-Wing
IFR	Instrument Flight Rules
IFT	Instrument Flight Trainer
IG	Image Generator
IHADSS	Integrated Helmet and Display Sighting System
IHAS	Integrated Helmet Assembly Subsystem
IHC	Inner Hair Cell
IID	Interaural Intensity Difference
IM	Informational Masking
IMC	Instrument Meteorological Conditions
INVS	Integrated Night Vision System
IPD	Interpupillary Distance; Interaural Phase Difference
IR	Infrared
IRA	Incremental Repeated Acquisition
ISI	Inter-Stimulus Interval
ISO	International Organization for Standardization
ITC	In-the-Ear
ITD	Interaural Time Difference
ITH	Improved Tactical Headset
ITO	Indium-Tin-Oxide
JHMCS	Joint Helmet-Mounted Cueing System
JMENS	Joint Mission Element Needs Statement
JND	Just Noticeable Difference
JSF	Joint Strike Fighter
KEAS	Knots Equivalent Air Speed
KEMAR	Knowles Electronic Manikin for Auditory Research
KHZ	Kilohertz
KG	Kilogram
KM	Kilometer
L	Loudness
LAN	Local Area Network
LASIK	Laser-Assisted in Situ Keratomileusis
LC	Liquid Crystal
LCD	Liquid Crystal Display
LCOS	Liquid Crystal on Silicon
LCS	Liquid Crystal Shutter
L/D	Light/Dark
LDL	Loudness Discomfort Level
LED	Light Emitting Diode
LGN	Lateral Geniculate Nucleus
LHX	Light Helicopter Experimental
LL	Loudness Level
LM	Lumen
LMS	Least Mean Square
LP	Line Pair
LSE	Life Support Equipment
LSOC	Lateral Superior Olivary Complex
LTG	Low Tension Glaucoma
M	Magnocellular
MAA	Minimum Audible Angle
MAAWS	Multi-Role Anti-Armor Anti-Personnel Weapon System
MACH	Modular Aircrew Common Helmet
MAE	Motion Aftereffect
MAF	Minimum Audible Field
MAMA	Minimum Audible Movement Angle
MANTIS	Multispectral Adaptive Networked Tactical Imaging System
MAP	Minimum Audible Pressure
MAR	Minimum Angle of Resolution
MCL	Most Comfortable Loudness
MCP	Micro-Channel Plate
MDS	Multidimensional Scaling
MFD	Multi-Function Display
MG	Milligrams
MIBU	Mobile Inshore Boat Unit
MIHDS	Modular Integrated Helmet Display System
MIL-HDBK	Military Handbook
MIL-SPEC	Military Specification
MIL-STD	Military Standard
MLD	Masking Level Difference
MLS	Maximum-Length Sequence

MM Masking Margin; Millimeter
MONARC Monolithic Afocal Relay Combiner
MOPP Mission Oriented Protective Posture
MOS Military Occupational Specialty; Mean Opinion Score
MOUT Military Operations on Urban Terrain
MPS Meters Per Second
MR Milliradian; Mixed Reality
MRE Mission Rehearsal Exercises
MRI Magnetic Resonance Imaging
MRO Maintenance, Repair and Overhaul
MRT Modified Rhyme Test; Multiple Resource Theory
MS Masking Stimulus
MSEC Millisecond
MSL Mean Sea Level
MSOC Medial Superior Olivary Complex
MSTd Medial Superior Temporal Area Pars Dorsalis
MTF Modulation Transfer Function
MTN Multitalker Noise
MURAL Multilevel Auditory Assessment Language
MVA Multi-Domain Vertical Alignment
MWSS Mounted Warrior Soldier System
NA Numerical Aperture
NASA National Aeronautics and Space Administration
NATO North Atlantic Treaty Organization
NATOPS Naval Air Training and Operating Procedure and Standard
NBC Nuclear-Chemical-Biological
NCW Network Centric Warfare
ND Near-Domain
NDFR Non-Distributed Flight Reference
NIN Nicotine-Induced Nystagmus
NIOSH National Institute for Occupational Safety and Health
NIHL Noise-Induced Hearing Loss
NM Nanometer
NOE Nap-of-the-Earth
NOMI Naval Operational Medicine Institute
NRC National Research Council
NRR Noise Reduction Rating
NSRDEC (U.S. Army) Natick Soldier Research, Development, and Engineering Center
NTSB National Transportation Safety Board
NVD Night Vision Device
NVG Night Vision Goggle
OAE Otoacoustic Emission
OBOGS On-Board Oxygen Generating System
OCB Olivocochlear Bundle
ODALab Optical Diagnostics and Application Laboratory
OEF Operation Enduring Freedom
OFT Operational Flight Trainer
OFW Objective Force Warrior
OHC Outer Hair Cell
OIF Operation Iraqi Freedom
OLED Organic Light-Emitting Diode
OSAP Optimum Sighting Alignment Point
OTC Over-the-Counter
OTW Out-the-Window
P Parvocellular
PA Pascal
PB Phonetically Balanced
PC Personal Computer
PDP Plasma Display Panel
PDU Pilot Display Unit
PFE Pilot Flight Equipment
PHODS Portable Helicopter Oxygen Delivery System
PKI Peacekeeping Institute
PKSOI Peacekeeping and Stability Operations Institute
PM-ACIS Program Manager-Aircrew Integrated Systems
PNVG Panoramic Night Vision Goggles
PNVS Pilot’s Night Vision System
POAG Primary Open Angle Glaucoma
PPE Push-Pull Effect
PRK Photorefractive Keratectomy
PRU Pilot Retained Unit
PSQ Perceived Sound Quality
PST Peristimulus
PT Physical Training
PTA Pure Tone Average
PTS Permanent Threshold Shift
PTSD Posttraumatic Stress Disorder
PTT Push-to-Talk
PVCN Posterior Ventral Cochlear Nucleus
PZT Lead Zirconium Titanate
QDC Quick Disconnect
QVGA Quarter Video Graphics Array
R&D Research and Development
RA Rapidly-Adapting
RCTD Reconfigurable Collective Training Device
RDK Random-Dot Kinematogram
RDS Random Dot Stereogram
RE Real Environment
REAT Real-Ear-Attenuation-at-Threshold
RETFL Reference Equivalent Threshold Force

Level
RETSPL Reference Equivalent Threshold Sound Pressure Level
RF Radio Frequency
RGB Red-Green-Blue
RGP Rigid Gas Permeable
RK Radial Keratotomy
RMS Root-Mean-Square
RPA Rotorcraft Pilot's Associate
RPE Retinal Pigment Epithelium
RPM Revolutions Per Minute
RSD Retinal Scanning Display
RSE Redundant Signals Effect
RSV Rotationally Symmetric Visor
RSVP Rapid Serial Visual Presentation
RT Response Time; Reverberation Time
RV Reality-Virtuality
RWS Remote Weapon System
S&T Science and Technology
SA Situation Awareness; Slowly-Adapting
SAT Speech Awareness Threshold
SBU Special Boat Unit
SCN Suprachiasmatic Nucleus
SD Spatial Disorientation
SDT Speech Detection Threshold
SI International System (of units); Speech Intelligibility
SII Speech Intelligibility Index
SIL Sound Intensity Level
SIR Speech Intelligibility Rating
SL Sensation Level
SME Subject Matter Expert
SNR Signal-To-Noise-Ratio
SOA Stimulus Onset Asynchrony
SOC Superior Olivary Complex
SOG Shade-of-Gray
SONAR Sound Navigation And Ranging
SP Summating Potential
SPE Solar Particle Events
SPH Soldier's Protective Headgear
SPL Sound Pressure Level
SR Speech Recognition
SRT Speech Recognition Threshold
SS Standard Stimulus
STI Speech Transmission Index
STN Super Twisted Nematic
STRICOM Simulation, Training and Instrumentation Command
STS Significant Threshold Shift
SVGA Super Video Graphics Array
SWAT Subjective Workload Assessment Technique
SWIR Short-Wave Infrared
SXGA Super Extended Graphics Array
T-NASA Taxiway-Navigation and Situation Awareness
TACTICS Tactile Communication System
TE Time Error
TFT Thin Film Transistor
TLS Tactor Locator System
TLV Threshold Limit Value
TMJ Temporomandibular Joint
TN Twisted Nematic
TNZ Thermoneutral Zone
TPL Thermal Plastic Liner
TRD Temporal Response Differentiation
TS Test Stimulus
TSAS Tactile Situation Awareness System
TSAS-SF Tactile Situation Awareness System for Special Forces
TTS Temporary Threshold Shift
TUC Time of Useful Consciousness
TV Television
TWS Thermal Weapon Sight
UAS Unmanned Aerial System
UAV Unmanned Aerial Vehicle
UHF Ultra High Frequency
UK United Kingdom
UNC University of North Carolina
USAARL United States Army Aeromedical Research Laboratory
USACHPPM U.S. Army Center for Health Promotion and Preventive Medicine
USAF United States Air Force
USAWC United States Army War College
USB Universal Serial Bus
UXGA Ultra Extended Graphics Array
V Vertical
VA Visual Acuity
VCOP Virtual Cockpit Optimization Program
VCS Visually-Coupled System
VDT Video Display Terminal
VE Virtual Environment; Ventriloquism Effect
VFD Vacuum Fluorescent Display
VFR Visual Flight Rules
VGA Video Graphics Array
VHF Very High Frequency
VISAA Visual Survey of Apache Aviators
VMC Visual Meteorological Conditions
VOR Vestibular-Ocular Reflex
VR Virtual Reality
VRD Virtual Retinal Display
VRDA Virtual Reality Dynamic Anatomy

VSI	Vision Systems International
VTAS	Visual Target Acquisition System
W	Watt
WAN	Wide Area Network
WBGT	Wet Bulb Globe Temperature
WBV	Whole-Body Vibration
WPM	Words Per Minute
XGA	Extended Graphics Array
Z	Vertical Direction (Axis)

Glossary

A

A-weighting: A technique used to obtain a single number representing the sound pressure level of a noise containing a wide range of frequencies in a manner approximating the response of the ear: the human ear does not respond equally to sounds of all frequencies, and is less efficient at low and high frequencies than it is at medium or speech range frequencies. Thus, the low and high frequencies are de-emphasized with the A-weighting.
Aberration: Any variance from a perfect reproduction of an image.
Aberrometer: An instrument designed to measure optical aberrations. Ophthalmic aberrometers were developed in order to measure complex refractive errors that cannot be measured by autorefractors or more traditional clinical methods.
Absolute threshold: The smallest value of a stimulus that results in a sensory reaction.
Acclimatization: The physiological adjustment (adaptation) to new physical and/or environmental conditions.
Accommodation: The autofocus process of the eye that helps maintain a clear retinal image for different viewing distances.
Achromats: A combination of lenses (usually in contact) which reduce chromatic aberration.
Acoustic: Pertaining to sound or to the sense of hearing.
Acoustic display: A display presenting acoustic information.
Acoustic field: A description of the behavior of sound in a specific space; the distribution of acoustic pressure generated by one or more sound sources in the specific open, partially bound, or fully enclosed space. An area in space containing sound waves
Acoustic impedance: The ratio of effective acoustic pressure averaged over a given surface to effective volume velocity of acoustic energy flowing through this surface. The units for impedance are Pa-s/m^3 or dyne-s/cm^5, which are called the acoustic ohm (Ω).
Acoustic manikin: A replica of the human head (or the human head and torso) with microphones placed in the ear canals, at the eardrum position, for making acoustic measurements and sound recordings.
Acoustic nerve: [See **Auditory nerve**]
Acoustic pressure: [See **Sound pressure**]
Acoustic reflex: An action of the middle ear muscles that reduces the sensitivity of the ear for high intensity stimuli.
Acoustic signature: Characteristic sound of a given sound source that permits sound source identification.
Acoustic wave: A mechanic disturbance propagating trough an elastic medium.
Acoustics: The science of production, transmission and reception of sound.
Active matrix electroluminescent (AMEL): A type of electroluminescent display where individual pixels are controlled by a dedicated electronic switch, and which are organized in a matrix form (rows and columns).
Active matrix liquid crystal display (AMLCD): A type of liquid crystal display where each individual pixel is controlled by a dedicated electronic switch, which are organized in a matrix form (rows and columns).
Active matrix OLED (AMOLED): A type of organic light emitting displays where individual pixels are controlled by a dedicated electronic switch, and which are organized in a matrix form (rows and columns).
Active noise reduction (ANR): The process of reducing background noise by electronically inverting its phase by 180 degrees and adding this inverted signal to the original noise.
Action space: The area in which an individual moves and makes decisions (within a 2-meter radius).
Actuator: A devices used or intended to be used for moving or controlling something.
Adaptation: An automatic adjustment of the sensory system in response to a prolonged stimulation. [See **Visual adaptation** and **Auditory adaptation**]

Adapter: [See **Interface**]
Adaptive automation: This is a departure from traditional automation in which the operator is taken out of the loop and is simply an observer. In adaptive automation, operator status is constantly monitored and the system dynamically off-loads or loads tasks to prevent operator overload or underload, respectively.
Addressability: The number of discrete horizontal and vertical pixels or subpixels of a matrix display that may be distinctly driven.
Advanced technology demonstration (ATD): Technology demonstrations tightly focused on specific military concepts and that provide the incorporation of technology that is still at an informal stage into a warfighting system. ATDs are of militarily significant scope and of a size sufficient to establish utility.
Aerial perspective: A depth perception cue of closer objects appearing bright and sharp, while distant objects appear pastel and hazy.
Afferent: Leading inwards, toward the center.
Air-bone gap: A difference between the threshold hearing levels for air conduction and bone conduction.
Air conduction: The process by which sound is conducted to the internal ear through sound waves exciting the air in the ear canal.
Aircraft retained unit (ARU): The frontal portion of the Helmet Integrated Display Sight System (HIDSS), consisting of two image sources, and optical relays attached to a mounting bracket.
Airspeed: The magnitude of the speed at which the aircraft moves relative to the air.
Ambient noise: All-encompassing sound at a given location, usually a composite of sounds from many sources near and far.
Ambient visual mode: Generally located in the peripheral portion of our vision, it interacts with our vestibular system to provide orientation and movement cues and is thought to be pre-attentive, that is not requiring any cognitive resources to process the information.
Amplitude modulation (AM): A systematic variation of the magnitude of one signal (carrier) in proportion to the magnitude changes of another signal (modulating signal).
Amplitude spectrum: [See **Spectrum**]
Angular resolution: [See **Spatial resolution**]
Anhedonia: The inability to gain pleasure from enjoyable experiences.
Anterior chamber: The front chamber of the eye formed by the cornea, iris and front surface of the crystalline lens.
Antihelix: A cartilaginous ridge of the pinna that is medial to and parallel to the helix.
Apex (of the cochlea)**:** The far away end of the spiral of the cochlea where scala tympani and scala vestibuli meet.
Apparent motion: The illusory sense that the objects have moved smoothly from one location to the other created by the rapid alternation of objects presented at different spatial locations.
Apparent size: The visual impression of size.
Aqueous humor: The fluid produced by the ciliary body which fills the anterior chamber of the eye.
Articulation: [See **Speech articulation**]
Articulation index (AI): An objective measure of speech intelligibility based on the average speech level and average noise level in 20 frequency bands over the frequency range from 250 Hertz (Hz) to 7000 Hz.
Artificial ear: [See **Ear simulator**]
Artificial intelligence (AI): The effort to computerize those skills that illustrate human intelligence e.g., understanding visual images, understanding speech and written text, problem solving.
Artificial mouth: [See **Mouth simulator**]
Aspect ratio: The ratio of horizontal dimension (width) to vertical dimension (height).
Astigmatism: One kind of refractive error in which optical power varies systematically over different radial meridians. It can be corrected with spectacles or contact lenses that have a corresponding distribution of refractive powers.

Attention: The application of cognitive or perceptual resources to a task; the concentration of mental effort on sensory or mental events. It is generally considered to be selective; only a subset of the stimuli received by our sensory organs is selected to enter the consciousness.
Audibility threshold: [See **Hearing threshold**]
Audio: Pertaining to an acoustic signal encoded in electrical form and to the means of its transmission.
Audio bandwidth: The range of audio frequencies that an electronic system is able to reproduce within predetermined tolerances.
Audio display: An acoustic display generated by audio signals.
Audio frequency: An acoustic frequency at which a sound is normally audible.
Audio frequency range: Frequency range that extends from the lowest to the highest acoustic frequencies perceived by humans, typically from 20 to 20,000 Hz.
Audio signal: An audible acoustic signal recorded or generated in an electrical form and reproduced by loudspeakers, earphones, or bone vibrators.
Audiogram: A graphic representation of the threshold of hearing of a person compared to the standardized threshold for normal hearing (in dB HL
Audiometric zero level: A value of 0 dB arbitrarily assigned to the reference hearing threshold level permitting expression of hearing loss as a number of dB of hearing level (HL) above the audiometric zero. [See **Reference hearing threshold level**]
Audition: A conscious act of hearing a sound; the ability to hear.
Auditory: Pertaining to sense of hearing or act of audition.
Auditory adaptation: A decrease in auditory sensitivity as a result of prolonged auditory stimulation.
Auditory cortex: The part of the brain's cortex that is responsible for processing auditory signals.
Auditory display: A display presenting information capable of being heard.
Auditory icon: A natural, real world, non-speech sound used as a communication signal that has a meaning associated with the object it represents, e.g., throwing a document into the desktop trashcan can be accompanied by a crumpled-paper sound to symbolize deleting a file within the context of the desktop metaphor.
Auditory image: An overall auditory sensation created by a specific acoustic signal during specific listening conditions; an auditory representation of a specific auditory stimulus.
Auditory nerve: An auditory branch of vestibulocochlear nerve**.**
Auditory pathways: The paths traced by the nerves leading from the organ of Corti in the cochlea to the auditory cortex.
Auditory perception: A mental synthesis of auditory sensations based on prior experience and world knowledge to determine meaning of the stimulation. [See **Perception**]
Auditory scene analysis: The process by which the human auditory system organizes sound into perceptually meaningful elements.
Auditory signal: An acoustic, mechanic, or electric form of a message received or intended to be received by the auditory system.
Auditory situation awareness: The component of situation awareness that is derived from the auditory cues. Auditory cues include information about the presence and location of events within the environment, including azimuth, elevation, and distance. It encompasses information from noises in the ambient environment, weapon noise, vehicle sounds, as well as spoken information through speech communications.
Auditory stimulus: An acoustic, mechanic, or electric stimulus received or intended to be received by the auditory system.
Auditory stream: A sequence of events perceived as coming from the same sound source.
Auditory tube: [See **Eustachian tube**]
Augmented cognition (AugCog): Using psychophysiological operator state measures as inputs to an adaptively automated system.

Augmented reality (AR): A display where computer-generated imagery or symbology is superimposed on the real world.
Aural harmonic: A harmonic of a given stimulus generated in the ear of the listener.
Auralization: Creation of virtual acoustic environments by rendering specific sound events on the impulse response characterizing a real or non-existent space.
Auricle: [See **Pinna**]
Autorefractor: An instrument designed to measure the lower-order refractive errors of the eye including myopia, hyperopia and astigmatism.
Aviator's night vision imaging system (ANVIS): A passive, binocular, third generation I^2 system with improved sensitivity and resolution over the second generation I^2 tubes: ANVIS are used extensively in military aviation.
Azimuth: An angle at which the specific sound source is situated in the horizontal plane in reference to the median plane of the listener. Azimuth is measured in angle degrees.

B

Babble: [See **Multitalker noise**]
Backwards masking: Auditory masking observed when a masking stimulus occurs after the test signal.
Bandwidth: The range of frequencies over which a device or system performs within specified limits.
Bark: A unit of the bark scale extending from 1 to 24 corresponding to a critical band.
Bark scale: A pitch scale created by adding side-by-side 24 non-overlapping critical bands and projecting them along the basilar membrane.
Base (of the cochlea)**:** The first and widest coil of the spiral of the cochlea.
Basilar membrane: A membrane along the spiral of the cochlea that is the base of the organ of Corti.
Battle fatigue: A psychological disorder that develops in some individuals who have had major traumatic experiences (e.g., have been in a serious accident or through a war). The person is typically insensitive at first but later has symptoms including depression, excessive irritability, guilt (for having survived while others died), recurrent nightmares, flashbacks to the traumatic scene, and overreactions to sudden noises.
Battlespace: Refers both to the physical environment in which a confrontation (i.e., warfare) will take place and the forces that will participate in the confrontation. All elements that support the warfighting forces (e.g., logistics, intelligence) are included in this definition. This term is replacing the historical terms *battleground* and *battlefield*.
Beam splitter: An optical device that splits a beam of light in two beams.
Beats: Periodic variations in sound intensity resulting from the superposition of two sinusoidal quantities of different but close frequencies.
Biaural: A condition in which the same acoustic signal is presented to both ears of the listener. [See **Diotic**]
Binaural: Pertaining to, using, or involving the functions of two ears.
Binaural advantage: Improvement in the reception of an auditory signal resulting from the interaction of two ears.
Binaural audio: A method for recreating an original sound field by reproducing a binaural recording of the original sound field through earphones.
Binaural dummy head: A replica of the human head (or the human head and torso) with microphones placed in the ear canals, at the eardrum position, for making acoustic measurements and sound recordings.
Binaural fusion: Sensation of a single sound caused by two different sounds delivered to the left and right ears.
Binaural listening: Listening with two ears.
Binaural masking level difference (BMLD): The difference between the binaural masked thresholds of hearing when the binaural masker is in phase and out-of-phase for the two ears.
Binaural mode: A sound delivery mode in which auditory stimuli are delivered to both ears of the listener.

Binaural summation: An improvement in the threshold of hearing and increased sound loudness due to listening with two ears.
Binaural recording: A method of recording an acoustic filed using a replica of the human head with two microphones in the place of the ears.
Binaural signal: A signal recorded with two microphones located at the ears of the listener or at the ears of binaural dummy head.
Biocular display: A term pertaining to optical devices which provide two visual inputs from a single sensor.
Biodynamic: Referring to characteristics of a system that bare related to forces that act on the human body.
Binocular alignment: The condition by which the optical axes of two independent oculars are parallel.
Binocular display: A term pertaining to optical devices which provide two visual inputs from two sensors which are displaced horizontally in space, making stereopsis possible.
Binocular fusion: The process by which two images, one seen by each one eye, are combined, or fused, into a single percept by the visual system.
Binocular overlap: That portion of an HMD's central display field that is observable by both eyes.
Binocular rivalry: The variation or suppression of a discerned image over time between images produced by two different eyes viewing different images.
Biocular HMD: The HMD configuration where both eyes see the same image source through the respective optic channels.
Biofeedback: A training technique that uses brain actuated control (BAC) based on the concept of recognizing alpha and gamma band EEG patterns that are to be used as a control signal.
Bistable stimulation: A form of multistable perception in which an ambiguous pattern of stimulation leads to two mutually exclusive perceptual interpretations. The observer or listener may alternate between, or be biased towards one of, the two interpretations. The Necker cube is a common example of visual bistable stimulation.
Blackout: A loss of consciousness.
Blade slap: The dominant noise produced by helicopters consists of a broadband spectrum generated by vortex formation and shedding in the flow past the helicopter blade. It is a distinctive, low frequency throbbing sound which increases during certain descent, maneuvering and high-speed cruise operations.
Blink: The rapid closing of the eyelids in response to a threat to the eye (reflex blink) or the slow closing of the eyelids to replenish and smooth the tear layer over the eye's surface (normal blink).
Bloch's law: Within a certain critical duration, all the light received by the retina is summed and processed as if it were a single light. Because of this, within the critical duration, a bright flash delivered within a short time has the same effect as a dimmer light delivered over a longer time, as long as the total quantity of light is the same. The critical duration can vary from 10 to 200 milliseconds depending on viewing conditions.
Bone conduction: The process by which sound is conducted through the cranial bones. This term applies to both the external sound and the talker's own speech transmitted through the bones to the internal ear or to the contact microphone located on the skull of the talker.
Bony labyrinth: A cavity within petrous portion of the temporal bone that houses the inner ear.
Boresight: An optical device with reticle used to align line of sight to the aircraft axis.
Bowman's layer: The second layer of the cornea, just below the epithelium. [See **Cornea**]
Brain: The command and control center of the central nervous system contained within the cranium.
Brain scan: A class of techniques in cognitive neuroscience that measure brain behavior and relate it to cognition.
Brainstem: The part of the central nervous system that connects spinal cord and majority of the cranial nerves to the forebrain and cerebrum; the lowest part of the brain.
Brick-wall filter: An informal term for an idealized electronic filter, which has 100% transmission in the pass band, 0% transmission in the stop band, and an abrupt transition(s) between the two bands.

Brightness (Auditory): A subjective percept that correlates with the amount of high frequency energy in the sound.
Brightness (Visual): A subjective percept that correlates with the luminance or intensity of a light. Along with hue and saturation, brightness is one of the attributes used to describe a particular color.

C

Catadioptric optical design: An optical system which utilizes both reflection and refraction.
Cataract: Any opacity in the crystalline lens of the eye. Smaller opacities or cataracts can cause scatter of incoming light resulting in perceived halos or glare and reduced contrast sensitivity. Denser opacities or cataracts can cause more significant reduction in visual quality and decreases in visual acuity.
Cathode-ray-tube (CRT): A display device that produces images by modulating the intensity of a scanning electron beam striking a phosphor coated surface (the screen).
Cent: A unit of musical pitch equal 1/100 of a semitone. Cents are used to measure extremely small intervals or to compare the sizes of comparable intervals in different tuning systems.
Center-of-mass (CM): That point of a body or system of bodies which moves as though the system's total mass was located at that point.
Central auditory nervous system (CANS): A sound processing part of the central nervous system (CNS); a system of neural fibers and nuclei that connect the ear with the brain.
Central masking: Masking that occurs when a masking stimulus is present in one ear and its masking effect is observed in the other ear.
Central nervous system (CNS): The part of the nervous system consisting of the brain and the spinal cord.
Change blindness: An effect of perception and attention where a person fails to see significant changes between two scenes.
Channel: A route through which signal or data pass or progress.
Channel capacity: The maximum data rate that can be attained over a given channel.
Characteristic impedance: [See **Specific acoustic impedance**]
Checkride: A practical test to measure the skills developed throughout flight training. Pass/fail is based on performance against published test standards.
Chromatic aberration: An optical defect of a lens system that degrades image quality and may cause colored fringes around images. It occurs when more than one wavelength of light is used, as in white light, because the focal power of a lens differs for every wavelength. Although an image may be sharply focused for one wavelength, it will be out of focus for other wavelengths.
Chromatic scale: Ascending or descending sequence of 12 music tones separated by semitones; a music scale that consists of 12 equally spaced logarithmic steps (semitones) in an octave. Chromatic scale corresponds to playing all the white and black keys on a piano.
Chromaticity: A description of the color property of light based on hue and saturation.
Cilia: Plural of *cilium***.** [See **Eyelashes** and **Hair cells**]
Ciliary body: The structure within the eye just posterior to the iris and anterior to the retina that contains the ciliary muscle, serves as the base for the iris, and produces aqueous humor.
Ciliary muscle: The muscle within the ciliary body which controls the accommodation of the crystalline lens. Contraction of the muscle releases tension on the zonules and allows the lens to bulge for near viewing; relaxation of the muscle increases tension on the zonules and pulls the lens flatter for distance viewing.
Circumaural earphone: An earphone that presses against the head with little or no contact with the surface of the pinna; the transducer is loosely coupled to the ear by the relatively large volume of air under the ear cup or earmuff.
Clarity: A sensation of the listener of being able to attend to the details of the auditory stimulus.

Classification: An arrangement according to some systematic division into groups of classes.
Coarticulation: An effect of the sequentially produced speech sounds upon each other.
Cochlea: The snail-shaped tube (in the inner ear coiled around the modiolus) where sound vibrations are converted into nerve impulses by the organ of Corti.
Cochlear nucleus: The most caudal auditory nucleus and termination point for all auditory nerve fibers.
Cocktail party effect: The ability to listen to different conversations within a crowded room simply attending to them. It is this ability that we take advantage of when presenting 3-D audio cues that improves speech intelligibility in the presence of noise and multiple talkers.
Cognition: The processes involved in human thought, perception, and action.
Cognitive neuroscience: The study of human cognition with emphasis on relating cognition to the brain.
Cognitive science: The study of cognition.
Cognitive resources: The capabilities and knowledge of information processing that are used to perform mental tasks. There are limited resources for different cognitive systems.
Cognitive tunneling: Difficulty dividing attention between two superimposed fields of information.
Coincidence detectors: Nerve cells that only respond to concurrent signals from more than one neuron.
Colavita effect: The phenomenon whereby participants presented with auditory, visual, or audiovisual stimuli in a speeded response task sometimes fail to respond to the auditory component of the bimodal targets.
Cold stress: Conditions where an individual is exposed for an extended period to temperatures significantly lower than normal body temperatures; can result in hypothermia, a condition marked by an abnormally low internal body temperature.
Collimation: The bringing of the optical components of a telescope into correct alignment. Collimated light is light whose rays are nearly parallel.
Coma: A higher-order optical aberration that causes an asymmetric image blur, such that a point of light is imaged somewhat like a comet. Within the Zernike system, for classifying ocular aberrations, the coma is divided into two sub aberrations labeled Z(3,-1) and Z(3,1).
Combiner: A beamsplitter that reflects a portion of a beam of light and transmits a portion.
Communicability: [See **Speech communicability**]
Communication: The process of transmitting ideas, thoughts, feelings, and opinions by means of signs, symbols, and signals produced consciously or unconsciously.
Complex tone: A sound consisting of several, usually harmonically related, pure tones.
Compression: In the physics of sound, the segment of the longitudinal wave where pressure is increased, the other segment being rarefaction.
Concha: A bowl-shape depression in the pinna surrounding the entrance to the ear canal.
Conductive hearing loss: A hearing loss caused by the problems in transmitting sound from the outer ear to the inner ear. Conductive hearing loss has a mechanical origin.
Cone of confusion: An imaginary cone-shaped surface radiating outwards from each ear and connecting points from which, a sound source would produce identical interaural difference cues, making the use of such binaural cues useless for sound localization.
Cones: Photoreceptor cells located in the retina, responsible for high-acuity vision and color vision in moderate or bright light; their interaction forms the basis of color vision. Distribution of the cone photoreceptor cells varies across the retina. They are most highly concentrated in the fovea.
Connotation: The indirect, associated, or implied meaning of a word or an expression; a meaning suggested or coded by a word or an expression.
Consonance: A relation between two or more tones that form a chord or interval that sounds pleasant. Generally consonance results from intervals composed of tones with simple frequency ratios such as 1:2 or 2:3.

Contralateral: A term generally used to refer to anatomical structures that are "on the opposite side" of the body as another structure (e.g. the left eye is the contralateral eye with respect to the right eye since it is on the left side of the body/brain).

Contrast: A measure of the luminance difference between two areas. Contrast can be formulated in different ways, e.g., contrast ratio, modulation contrast, etc.

Contrast ratio: A mathematical expression of the luminance ratio for two adjacent areas. As used herein, contrast ratio is defined as higher luminance/lower luminance.

Contrast sensitivity: One measure of visual performance that describes how well an eye can see low contrast patterns. An eye with good vision can detect a low contrast pattern, while an eye with poor vision cannot detect a pattern unless it has high contrast.

Convergence: When the two eyes turn inward often used in order to place two images of an object on corresponding retinal locations.

Core temperature: The internal temperature of the body specifically in the deep structures of the body, in comparison to temperatures of peripheral tissues.

Cornea: The clear dome at the front of the eye. The cornea is the transparent collagen structure which serves as the primary focusing surface for the eye and provides about 65% of the eye's total refractive power. The five layers of the cornea are the epithelium at the front surface, Bowman's membrane, the stroma, Descemet's layer and the endothelium at the back surface.

Coronal plane: [See **Frontal plane**]

Cortex: The outer layer of the brain.

Countermeasure: An action taken to offset another action.

Critical band: A frequency band (a) within which distribution of sound energy has no impact on loudness of sound and (b) an extension of the continuous masking noise outside of which has no impact on hearing threshold of a tone located in the center of the band.

Critical distance: The distance from the sound source at which the intensity of the direct and reflected sound fields are equal.

Critical flicker frequency (CFF): The frequency at which a flickering light appears to no longer flicker; that is, when the flicker "fuses" into an apparently continuous light. This is also sometimes referred to as the critical flicker fusion frequency or the critical flicker frequency.

Critical ratio: The level of pure tone at threshold (in dB) minus the spectrum level (dB per Hz) of the noise.

Cross-talk: A leakage of unwanted energy or message into a communication channel from another channel.

Crystalline lens: The transparent lens within the eye that provides additional focusing power to the eye and, in the young eye, through its ability to change shape provides accommodation to view near objects.

D

Dark adaptation: The physiological process by which the retinal photoreceptors re-adjusts sensitivity to allow vision in darker conditions. That is, when the eyes "adjust to the dark."

Dark focus: The point of accommodation of the eye in the absence of visual stimuli.

Decay time: The time taken by a quantity to decrease its level by a specified amount from its peak.

Decibel (dB): A logarithmic unit of the ratio of two powers (P) expressed as 10 log (P_1/P_2).

Declarative memory: Memory experiences that can be explicitly recollected or declared.

Dehydration: Depletion of bodily fluids; the loss of too much body fluid through frequent urinating, sweating, diarrhea, or vomiting.

Demographics: The physical characteristics of a population such as age, sex, marital status, family size, education, geographic location, and occupation.

Denotation: An explicit or direct indication of the meaning of a word or expression.

Depth perception: The visual discrimination of absolute and relative distance using monocular and binocular cues.
Descemet's membrane: The fourth layer of the cornea, just anterior to the endothelium. [See **Cornea**]
Design eye position: The midpoint of the line segment of the open nosed vision line connecting two points which represents the predicted eye positions of the extremes of the aircrew population.
Detection: Determination of the presence of a sensory stimulus.
Deutan: A type of hereditary color vision anomaly in which the patient is missing or has defective M-cones. Since M-cones have peak sensitivity in the middle wavelengths range of the visible light spectrum, deutans are sometimes called, "green weak, or green color blind."
Diatonic scale: An ascending or descending scale of 7 music tones separated by 5 tones and 2 semitones. Diatonic scale corresponds to playing only the white keys on a piano.
Dichotic: A condition in which the signal presented at the left sensor differs from the signal presented at the right sensor (ears or eyes).
Dichotic display: An earphone display presenting different acoustic signals to the left and right ear of the listener.
Dichotic mode: An information delivery mode in which different signals are delivered to the left and right sense organs (ears or eyes).
Difference limen (DL): A smallest perceived change in a physical variable.
Differential threshold: The smallest detectable difference in a specified modality of sensory input.
Diffraction: The spreading out of waves when they encounter a small obstacle or pass through a narrow opening
Diffuse field: An acoustic space where sound waves have an equal probability of coming from any direction at any given moment due to their reflection from multiple surfaces.
Diffusion: A scattering of sound waves from irregular objects and space boundaries.
Digital micromirror device (DMD): A matrix display where each pixel is a very small square mirror on the order of ten to twenty microns. Each mirror pixel is suspended above two electrodes driven by complementary drive signals.
Diopter: A unit expressing the refractive power of an optical system/component as the reciprocal of the focal length in meters.
Diotic: A condition in which the signal presented at each ear or eye is identical.
Diotic display: A display presenting the same signal to both the right and left sense organs (ears or eyes).
Diotic mode: A stimulus delivery mode in which the same signal is delivered to both he right and left sense organs (ears or eyes).
Diplopia: A condition in which a single object appears as two objects; double vision. Normally this occurs when the two eyes are pointed in different directions, such as with crossed eyes. This causes a single object to appear in a different location for each eye, so when input from the two eyes are combined, the object is seen in two different locations, that is, double. In some unusual conditions, it is possible to experience diplopia with one eye (monocular diplopia).
Dipvergence: The shifting of the eyes vertically, one up and one down.
Directional device: A device in which the received or radiated signal is dependent on the direction of observation.
Discrimination: Determination that two specific sensory stimuli are different.
Disparity: Difference or misalignment.
Display: A unique device or assemblage of devices used to systematically present specific information capable of being perceived by the human senses; a structured presentation of information to the senses.
Display lag: The time delay in a display measured from the time when the imaging data are received and the time they are presented.
Dissonance: A music interval or chord that sounds unpleasant or rough. The frequency components resulting in dissonant sound do not have simple frequency ratios.

Distortion: An unwanted variation in magnification or a prismatic deviation with angular distance from the center of an optical component or system; any undesired change in the frequency or amplitude of an acoustical signal.
Divergence: The shifting of the eyes outward.
Divided attention: An intentional effort to be aware of two or more things simultaneously.
Duration: The time during which something exists.
Dynamic retention: When pertaining to helmets, the condition of preventing the loss of a helmet during a crash sequence.
Dynamic range: In a system or a transducer, the difference, measured in decibels, between the overload level and the minimum acceptable level. The minimum level is commonly fixed by any or all of the following: noise level, low-level distortion, interference, or resolution level.
Dysbarism: A medical conditions resulting from exposure to decreased or changing barometric pressure.

E

Ear canal: A part of the external ear that directs sound from the pinna to the tympanic membrane.
Ear simulator: A device simulating the acoustic characteristics of the human ear upon the sound radiated by an external sound source such as an earphone.
Earbud: A small earphone intended to be placed in the pinna at the entrance to the ear canal.
Earcon: An abstract, non-speech sound (e.g., synthetic sound, music sound) used as a communication signal.
Earcup: An enclosure surrounding the pinna.
Eardrum: [See **Tympanic membrane**]
Earmuff: [See **Earcup**]
Earphone: An electroacoustic transducer converting electric current into sound and directly coupled to the ear of the listener.
Earphone display: An audio display created by earphones.
Earplug: A device intended to be placed in the ear canal.
Effector: Any biological organ or system that becomes active in response to neural stimulation.
Efferent: Leading outwards, toward the periphery.
Egocentric: Using one's self as the reference frame.
Egress: The process of exiting an enclosed area (e.g., cockpit, tank interior).
Electroacoustic transducer: A transducer designed to receive an electrical signal and convert it into an acoustic signal or vice versa.
Electroencephalography (EEG): The measurement of brain wave rhythms of different frequencies. Brain waves are divided into: Delta (0.5 to 3 Hz), Theta (4 to 7 Hz), Alpha (8 to 12 Hz), Beta (13 to 30 Hz), and Gamma (>30 Hz).
Electroluminescence (EL): A flat panel display technology in which a layer of phosphor is sandwiched between two layers of a transparent dielectric (insulator) material which is activated by an electric field.
Electromagnetic spectrum: The entire range of radiation extending in frequency from approximately 10^{23} Hz to 0 Hz or, in corresponding wavelengths, from 10^{-13} cm to infinity and including, in order of decreasing frequency, cosmic-ray photons, gamma rays, x-rays, ultraviolet radiation, visible light, infrared radiation, microwaves, and radio waves.
Electrophoresis (EP): A nonemissive flat panel technology based on the movement of charged particles (of one color) in a colloidal suspension (of a second color) under the influence of an electric field. The application of the electric field changes the absorption or transmission of light through the solution.
Electrostatic transducer: A transducer consisting of a fixed electrode and a movable electrode, charged electrostatically in opposite polarity; the motion of the movable electrode changes the capacitance between the

electrodes and thereby makes the applied voltage change in proportion to the amplitude of the electrode's motion; also known as condenser transducer.

Elevation angle: An angle at which the specific sound source is situated in the vertical plane in respect to the horizontal reference plane of the listener. Elevation is measured in angle degrees.

Emitter: [See **Transmitter**]

Emmert's law: A law used in vision science that states that objects that generate retinal images of the same size will look different in physical size (linear size) if they appear to be located at different distances. Specifically, the *perceived linear size* of an object increases as its *perceived* distance from the observer increases.

Emmetropia: The condition of an eye with perfect optics, that is, no refractive error. When an emmetropic eye views a distant object, its image correctly focuses onto the retina.

Endotracheal intubation: The placement of a tube into the trachea (windpipe) in order to maintain an open airway in patients who are unconscious or unable to breathe on their own.

Energy: The ability or capacity of an object to do work.

Energetic masking: The type of masking that physically affects the audibility of the target sound through the presence of acoustic energy in the same spectral region as the target sound.

Envelope: An imaginary line connecting sequential peaks of a sound.

Environment: A set of circumstances and conditions that are extraneous to a given process but affect its nature or effectiveness.

Equally masking noise: Noise that equally masks tones of all frequencies.

Equivalent rectangular bandwidth (ERB): The bandwidth of a rectangular filter that has the same peak transmission as a given filter and that passes the same total power for a white noise input.

Ergonomics: [See **Human factors**]

Eustachian tube: The air channel that connects the middle ear cavity with nasopharyngeal cavity.

Event related potentials (ERP): Non-volitional EEG responses that generate a voltage – either negative (N) or positive (P) occurring within a specific timeframe – after an observed event. The P3 (also called the P300) is a positive voltage that occurs roughly 300 milliseconds after a sensory stimulus and the N1 is a negative voltage that occurs roughly 100 milliseconds after a stimulus.

Executive control: That part of our brain which allows us to direct attention or filter out unwanted stimuli.

Exit pupil: The region where the observer's eye(s) must be located in order to view the total field of view. In optics, it is the image of the aperture stop as formed from the image side of the optics.

Exit pupil expander (EPE): An optical device that increases the exit pupil.

External auditory meatus: [See **Ear canal**]

Experimental psychology: A field of study that investigates human behavior through scientific measurements.

Externalization: The sensation that a sound source is located away from the head.

Extraocular: Refers to structures outside of the eye generally associated with or connected to the eye (e.g., extraocular muscles).

Eye clearance distance (ECD): The minimum clearance from the closest display system component to the cornea of the eye. This parameter is important in determining system compatibility with add on devices, e.g. corrective lenses, protective masks, etc; also referred to as *physical eye relief*.

Eye dominance: The tendency of clusters of nerve cells in the visual system to respond primarily to one eye rather than to the other.

Eye relief: The distance between the last surface of the optical elements and the cornea of the eye.

Eyelids: The portion of moveable thin skin which serves to protect the front of the eye. The human eye is protected by an upper and lower eyelid; also referred to as "lids."

Eyelashes: Small hairs, also known as "cilia," which grow along the edge of each eyelid.

F

f number (f/#): The expression denoting the ratio of the equivalent focal length of a lens to the diameter of its entrance pupil.
Factor analysis: A statistical method that reduces a larger set of variables to a smaller set of dominant factors based on correlations between those variables.
Fatigue: A condition of weariness, exhaustion, or decreased sensory sensitivity from labor, exertion, or prolonged stimulation.
Fast Fourier transform (FFT): An algorithm that allows quick, economical application of Fourier techniques to a wide variety of analyses.
Fidelity: Similarity of a given auditory image to the specific auditory standard or to another auditory image.
Field emission display (FED): An emissive flat panel display technology which consists of a matrix of miniature electron sources which emit the electrons through the process of field emission. Field emission is the emission of electrons from the surface of a metallic conductor into a vacuum under the influence of a strong electric field.
Field-of-view (FOV): The maximum image angle of view that can be seen through an optical device.
Figure-of-merit (FOM): A metric which quantifies some aspect of image quality.
Filter: A device or material that passes signals (waves) of certain frequencies while stopping others.
Fiscal year (FY): A 12-month period over which the military budget allocates funding. It runs from October 1 of the prior year through September 30 of the next year.
Fixed-wing aircraft: A powered aircraft that has wings attached to the fuselage so that they are either rigidly fixed in place or adjustable, as distinguished from rotary-wing aircraft, like a helicopter.
Flashblindness: A temporary loss of vision as a result of sudden high level of luminance, e.g., nuclear explosion.
Flat panel display (FPD): A wide-encompassing category of display technologies characterized by significantly lower depth compared to the height and width.
Flicker: A perceived rapid variation in brightness (intensity).
Fluctuation strength: Intensity of perceptual impression created by amplitude and frequency modulations of sound at low modulation rates, up to about 20 Hz.
Focal visual mode: Generally located in the central portion of our vision, it is that portion of our perception that provides us detailed information. It requires that we direct our attention and may narrow under cognitive load.
Foot-Lambert (fL or ft-L): A unit of luminance (photometric brightness), equal to $1/\pi$ candela per square foot, or to the uniform luminance of a perfectly diffusing surface emitting or reflecting light at the rate of 1 lumen per square foot.
Formant: A significant peak in the complex sound spectrum of a given auditory stimulus.
Forward masking: Masking observed when a masking stimulus occurs before the test signal.
Forward looking infrared (FLIR): A thermal imaging sensor, where sensor output is based on infrared radiation (usually between 3 to 5 or 8 to 12 micron spectral range) generated by the external scene.
Fourier analysis: Data series analysis based on the concept that each shape of a waveform is a sum of several sinusoidal functions; a mathematical decomposition of a complex signal into elementary sine waves.
Fovea: A small microscopic depression at the center of the retina, which has the greatest density of cone photoreceptor cells, and therefore the best visual acuity. The center of an object being viewed is imaged onto the fovea, therefore this point corresponds to the straight-ahead visual direction; also referred to as the “foveola” or “fovea centralis.”
Frame rate: The frequency of frames produced per second (expressed in Hertz [Hz]).
Frangibility: The ability of a subsystem or component to separate from the major system. Some helmet and display system designs may employ helmet mounted displays, eye protection devices, etc., which actively or passively separate from the helmet under crash conditions.

Frankfurt plane: The eye-ear plane in which the human skull is placed in a position so that the lower margins of the eye socket and the upper margins of the auditory opening are on the same horizontal plane.
Free field: An acoustic reflection-free environment in which sound pressure level is inversely proportional to the distance from a sound source (a 6 dB decrease for each doubling of distance).
Frequency: Number of complete oscillation cycles per unit of time. The unit of frequency is the Hertz (Hz).
Frequency modulation (FM): A systematic variation of the frequency of one signal (carrier) in proportion to the magnitude changes of another signal (modulating signal).
Frequency response: The measure of any system's spectrum response at the output to a signal of varying frequency (but constant amplitude) at its input; in the audible range it is usually referred to in connection with electronic amplifiers, microphones and loudspeakers.
Frequency resolution: A precision with which a person or a system can differentiate between fundamental frequencies of two waveforms.
Frontal plane: An imaginary plane dividing human body in front (anterior) and back (posterior) parts.
Fundamental frequency: The lowest frequency in a harmonic series; the lowest common factor in a harmonic series.

G

G-loading: An effect of the gravitational force at the earth's surface (force due to gravity); usually expressed as the numerical ratio.
Gabor patch: A luminance profile where the intensity at the center is the maximum grayscale value and the intensity at the edge of the diameter is one grayscale step above the background.
Ganglion: A group of cell bodies in the peripheral nervous system.
Gestalt laws: Gestalt is a German word meaning *form* or *pattern.* Gestalt laws refer to a set of principles used by the perceptual system to organize sensory information into patterns that are regular, orderly, symmetric, and simple. They include the *laws of proximity, similarity, good continuation, closure, simplicity* and *common fate*.
Golay code: A type of error-correcting code used in digital communications.
Ghost image: A spurious image produced as a result of an echo or reflection in the transmission of an image or signal.
Glare: A condition in which a bright light interferes with vision. One example is intraocular light scatter with cataracts that reduces contrast and degrades vision.
Glaucoma: A condition in which the intraocular pressure exceeds the eye’s ability to maintain normal function. Glaucoma results in damage to the nerve cells within the eye which leads to loss of vision in the mid-peripheral visual field.
Globe: A protective structures of the eye consisting of the sclera and cornea which maintain the shape of the eye.
Granit-Harper law: A temporal vision phenomenon in which flicker is more easily seen if the light is larger.
Grayout: A transient loss of vision characterized by a perceived dimming of light accompanied and loss of peripheral vision.

H

Hair cells: Sensory cells of the hearing and balance senses with tiny hair-like projections, called stereocilia or cilia, extending from the top of a cell and giving the cell its name.
Halation: A halo or glow surrounding a bright spot on a fluorescent screen or a photographic image.
Hallucination: A sensory perception (auditory, visual, etc) appearing without an actual physical stimulus. Unlike perceptual illusions, hallucinations are usually individual to the perceiver and may signal abnormal circumstances.

Hand-arm vibration (HAV): Vibration that is transmitted from vibrating surfaces of objects, such as hand tools, through the hands and arms.

Haptic: Refers to all the physical sensors that provide a sense of touch at the skin level and force feedback information from muscles and joints.

Haptics: The design of clothing or exoskeletons that not only sense motions of body parts (e.g., fingers) but also provide tactile and force feedback for haptic perception of a virtual world.

Harmonic: A pure tone component of a complex tone which frequency is an integral multiple of the fundamental frequency.

Head-related transfer function (HRTF): A frequency domain representation of the changes in magnitude and phase of the auditory signal at the entrance of the ear canal in relation to the signal at the source. The HRTF represents a linear transformation that occurs as a sound generated by a point source propagates to the left and right ears of a listener. The HRTF includes diffraction effects by the head and torso, as well as the directional spectral shaping effects of the outer ear or pinna. Unless otherwise specified, the HRTF is assumed to be the free-field HRTF.

Head-supported weight (HSW): the added weight required to be supported by the neck muscles as a result of the HMD system; used interchangeably with head-supported mass (HSM).

Head-up display (HUD): Any transparent display that presents data without obstructing the user's view.

Headgear: A system that covers the head or a part of it.

Headphones: Earphones applied outside the ear and supported by a headband.

Health hazard assessment (HHA): Assessment of risk to the health and effectiveness of personnel who test, use, and maintain the system. Hazards can arise from characteristics of the system itself or from the environment in which it operates.

Hearing: A sense by which biological systems are aware of the surrounding acoustic environment and perceive sound; an ability to perceive a sound.

Hearing level (HL): An amount by which a specific sound pressure level or force level exceeds a reference hearing threshold level.

Hearing loss: Any degree of impairment of the ability to hear sound.

Hearing protection device (HPD): A device designed or used to reduce the noise level reaching the auditory system.

Hearing threshold: A minimum (a) sound pressure level or (b) force level of a signal that is capable of evoking an auditory sensation in a specified fraction of the trials. The hearing threshold is defined for a given listener and a specified signal.

Heat stress: A group of conditions due to overexposure to or overexertion in excess environmental temperature. It includes heat cramps, heat exhaustion, which is more serious, and heatstroke.

Helicotrema: A narrow passage at the apex of the cochlea at which scala tympani and scala vestibuli are connected.

Helix: The cartilaginous fold of the pinna that curves around the outside edge of the pinna.

Helmet: Device covering the head and used for protecting the user from hazard to the head. A modern helmet serves as both the head protector and the supporting element for the communication system.

Helmet Integrated Display Sight System (HIDSS): A partially-overlapped biocular helmet-mounted display system under development for the RAH-66 Comanche helicopter consisting of two components: pilot retained unit (PRU) and an aircraft retained unit (ARU). The PRU is the basic helmet with visor assembly; the ARU is a front piece consisting of two image sources and optical relays attached to a mounting bracket.

Helmet-mounted display (HMD): A multimodal display systems used to enhance the user's situational awareness; a device, worn on the head or as part of a helmet, which has a small display optic in front of one (monocular HMD) or each eye (binocular HMD). HMDs can present both audio and visual information.

Hemorrhage: The act of bleeding or a collection of blood generally within a tissue (e.g., retinal hemorrhage).

Homeostasis: The ability or tendency of an organism or cell to maintain internal equilibrium by adjusting its physiological processes.
Homeotherm: An organism that is capable of keeping its core temperature within a relatively narrow range.
Horizontal plane: An imaginary plane dividing human body into zenith (superior) and nadir (inferior) parts.
Horopter: The region in space where the two images from an object falls on corresponding locations on the two retinas.
Hot spot: Pressure points that develop over time during the wearing of headgear.
Human factors: The science of human-machine relationships and interactions including all biomedical and psychological considerations; the science of designing the objects and environments according to human needs and capabilities.
Human factors engineering assessment (HFEA): Analysis of acceptable human engineering design criteria, principles and practices.
Human-in-the-loop (HITL): A model that requires active human interaction.
Human-machine interface (HMI): Any device that serves as a "*bridge*" connecting the (human) user and the machine. Common examples include keypad, mouse, touch screen, or keyboard.
Hue: The quality of color is most closely associated with a particular wavelength. Examples of hues include red, orange, yellow, green, blue and violet. To fully describe a color you must mention not only its hue, but its saturation and brightness as well.
Hypnotics: Drugs classified as central nervous system depressants used to induce sleep.
Hyperacuity: Usually refers to vernier acuity, or another visual task in which the threshold is significantly better than one arc minute, which is minimum angle of resolution expected for a standard Snellen visual acuity test.
Hyperopia: Farsightedness; a kind of refractive error in which an eye, viewing a distant object, focuses the image beyond the retina. With hyperopia, distant objects usually appear clearer than near objects, however some young patients can compensate for hyperopia without glasses by over accommodating.
Hyperstereopsis: A condition of exaggerated depth perception which occurs as a result of separation of the sensors greater than the eyes of the user.
Hypoxia: A condition resulting from a deficiency in the amount of oxygen reaching body tissues.

I

Identification: An act of assigning a unique name to a given stimulus or object.
Idiopathic disease: A disease having no known cause.
Illuminance: A measure of visible energy falling on a surface.
Illusion: An erroneous mental representation.
Infrasound: An acoustic wave of a frequency lower than the lower limit of human hearing; usually considered to be a sound having frequency lower than 20 Hz.
Image intensification (I^2): Sensor technology based on amplification of ambient light. Photons are imaged onto a photocathode which converts them into electrons. The number of electrons is multiplied and channeled onto a phosphor screen.
Image overlap: The portion (usually expressed as a percentage) of the total field of view of a biocular/binocular system that can be viewed simultaneously by both eyes.
Image smear: An image artifact resulting from relative motion between scene and sensor. This is caused by insufficient temporal characteristics within the imaging system, e.g., phosphor persistence, scan rate, etc.
Immersion: The feeling of being integrated in a computer-generated world.
Impact attenuation: The reduction of the mechanical force (and energy) through a protective material or device.
Impedance: The ratio of sound pressure to particle velocity of the sound wave; an opposition to the flow of energy through a system.

Impulse noise: A category of (acoustic) noise that is very intense and of short duration, usually less than a second, such as backfires from motor vehicles, sonic booms and weapons fire.

In-the-head localization: The sensation that all sound sources are located in the listener's head. Stereophonic sounds are typically considered lateralized in-the-head while spatial sounds are considered to be localized outside-the-head.

Index of refraction: The ratio of the speed of light in a vacuum to the speed of light in a substance; a relative measure of a lens material's ability to refract (bend) light.

Inferior colliculus: A part of the central auditory nervous system located in the dorsal part of the midbrain.

Information: A temporary change in a state of an object or matter. [See **Message**]

Information superiority: The capability to collect, process, and disseminate an uninterrupted flow of information while denying an adversary's ability to do the same.

Informational masker: A form of perceptual masking or interference that cannot be construed as energetic masking

Infrared: A portion of the electromagnetic spectrum; **a**n invisible band of radiation with wavelengths from 750 nanometers to 1 millimeter, infrared starts at the end of the microwave portion of the spectrum and ends at the beginning of visible light portion.

Inner ear: A complex system of interconnecting cavities, consisting of cochlea (which contains the nerves for hearing), the vestibule (which contains receptors for balance), and the semicircular canals (which contain receptors for balance).

Insert earphone: An earphone that is inserted into the ear canal or is coupled to the ear canal by a tube, earmold, or other device.

Intelligibility: [See **Speech intelligibility**]

Intensity resolution: A precision with which a person or a system can differentiate between two levels of the same signal.

Interaural cross correlation (ICC): A measure of the difference in a signal received by the two ears. Its value varies from -1, meaning the signals are equal and out of phase, through 0, meaning the two signals have nothing in common, to +1, meaning the signals are equal and in phase.

Interaural intensity difference (IID): The difference between the intensity of the sound reaching the right ear and the left ear of a listener. IID depends on the location of the sound source and the frequency of the sound.

Interaural level difference (LD): [See **Interaural intensity difference (IID)**]

Interaural phase difference (IPD): A difference in the phase of the continuous periodic sound reaching the right ear and the left ear of a listener. IPD depends on the location of the sound source and the frequency of the sound.

Interaural time difference (ITD): The difference in the time of arrival of a sound reaching the right ear and the left ear of a listener. ITD is independent of the sound frequency but depends on sound source location.

Interlace ratio: The number of fields per frame pertaining to displays.

Interface: A boundary or connection between systems, equipment, concepts, or humans beings; a special device or system providing operative compatibility between two or more different devices or systems.

Interference: Any process in the same medium or channel other than the given process or signal itself.

Internalization: The sensation that a sound source is located inside the listener's head. Sounds presented through earphones without spatial processing appear as internalized.

Interpupillary distance (IPD): The distance between the centers of the pupils of the two eyes.

Interval: A distance between two notes corresponding to a ratio of two frequencies. Intervals can be measured by Hertz, cents, scale steps or semitones.

Intraocular: Refers to structures or conditions inside the eye (e.g., intraocular hemorrhage).

Impact attenuation*:* The reduction in mechanical force through the protective helmet.

Ipsilateral: A term generally used to refer to anatomical structures that are "on the same side" of the body as another structure (e.g., the ipsilateral optic nerve for the right eye would be the optic nerve on the right side of the body/brain).
Iris: The iris is forms the aperture of the eye, the "pupil." The iris consists of two opposing muscles which either constrict (sphincter muscle) or dilate (dilator muscle) the iris in response to light or neurological stimuli.

J

Jitter: Small, rapid variations in a signal due to vibrations, voltage fluctuations, control system instability, and other causes.
Just noticeable difference (jnd): [See **Difference limen**]

K

Keratorefractive: Describes any method that changes the eye's focal power by changing the shape of the cornea. Kerato refers to the cornea and refractive refers to the focusing of light. The most common use of this word is with keratorefractive surgery, which alters the cornea's focal power, usually with laser sculpting of the cornea.
Key: A first tone of a diatonic (major or minor) scale that a piece of Western music is based on. The key's pitch is not absolute, but can be of any of several notes sharing the same pitch class e.g., the key of "C" refers to the note "C" in any octave
Knot: A unit of speed of one nautical mile (6,076.12 feet or 1,852 meters) per hour.
Knot-hole effect: The apparent limitation of the field-of-view due to the exit aperture.

L

Lambertian emitter: An optical source with a luminous distribution that is uniform in all directions.
Lamina cribrosa: The mesh-like structure at the posterior portion of the globe of the eye through which the optic nerve passes.
Language: A system of symbols, signs, or signals used to convey information.
Laser: Any of several devices that emit highly amplified and coherent radiation of one or more discrete frequencies. The term "laser" is an acronym for "light amplification by stimulated emission of radiation."
Lasik: An acronym for **l**aser-**a**ssisted **i**n **s**itu **k**eratomileusis, a type of laser eye surgery designed to change the shape of the cornea to eliminate or reduce the need for glasses and contact lenses in cases of severe myopia (nearsightedness).
Latency: The period between the initiation of something and the occurrence.
Lateral geniculate nucleus (LGN): A structure within the dorsal thalamus which regulates visual information received via the optic tracts from each eye.
Lateralization: The process by which a person determines location of a specific activity or mental event in one side of the body. The process of lateralization applies to the sound sources perceived as being located inside the listener's head.
Lead, lanthanum, zirconate, and titanate (PLZT): A material that can be electronically switched rapidly in polarity such that when sandwiched with a near infrared blocking material and a fixed polarizing material, the visual transmittance can be varied from full open state (approximately 20%) to a full off (optical density (OD) is greater than 3.0) in approximately 150 microseconds.
Lens: An object made of transparent material, usually with two curves surfaces, that bends (refract) or focus light rays passing through it; the transparent structure inside the eye that focuses light rays onto the retina.

Light emitting diode (LED) display: Emissive display composed of multiple light emitting diodes arranged in various configurations which can range from a single status indicator lamp to large area x-y addressable arrays.
Lightness: The perceptual correlate of reflectance; perceived reflectance.
Liquid crystal display (LCD): A type of nonemissive flat panel display technology which produces images by modulating ambient light. The ambient light can be reflected or transmitted light from a secondary, external source (e.g., a backlight).
Line of sight (LOS): The line between the pupil of the eye to the object of interest.
Line replaceable unit (LRU): A maintenance term referring to a systems or module that can be replaced in the field; usually requires no special alignment.
Listening: An act of attentive audition.
Lombard effect: An involuntary tendency of the talker to increase voice intensity in noise.
Localization: The process by which a person determines the direction of an incoming stimulus or the direction to a specific object in space. The process of localization applies to the sound sources perceived as being located outside the listener's head.
Long-term memory: A hypothesized system of human memory that holds information for durations ranging from several seconds to years.
Loss aversion: A property of human decision-making. People are usually more sensitive to perceived losses than perceived gains.
Loudness: A perceptual attribute of sound in terms of which sounds may be ordered on a scale extending from soft to loud; a perceived impression of the intensity of sound. Loudness depends on the actual intensity of sound and its spectral content (frequency) and duration.
Loudness adaptation: [See **Auditory adaptation**]
Loudness level: A median sound pressure of a 1000 Hz tone that is judged equally loud as a given sound.
Loudspeaker: An electroacoustic transducer that converts electric current into sound radiating into open space.
Loudspeaker display: An audio display using loudspeakers.
Law of the first wavefront: (See **Precedence effect**)
Luminance: Luminous flux per unit of projected area per unit solid angle leaving a surface at a given point and in a given direction; measured in foot-Lamberts (fL).
Luminance disparity: In biocular/binocular helmet-mounted displays, the difference in the image luminance between the two channels.
Luminance transmittance: The fraction of luminance of the outside world seen through an optical component or system; usually expressed as a percentage.
Luminous efficiency: The ratio of the energy of the visible light output, such as the energy emitted by a phosphor, to the electron energy of the input signal.
Luning: The subjective darkening that can occur in the monocular side regions near the boundaries of the partially overlapped region in a binocular display.

M

Macula: The central region at the posterior aspect of the retina which includes the fovea at its center. The macula has a denser distribution of cones than rods and is responsible for defined vision and color perception; also referred to as "macula lutea."
Macular degeneration: A common cause of vision loss in the elderly due to a degeneration of the central portion of the retina known as the macula. The degeneration is due to a build up of waste material in the macular region.
Magnetic resonance imaging (MRI): A method of scanning that produces detailed maps of the tissue relying on the difference in the magnetic resonance of certain atomic nuclei.

Manpower and personnel integration (MANPRINT) program: An Army system analysis which addresses manpower, training, personnel requirements: health and safety issues; and human factors issues.
Masking: A reduction of sensitivity to one stimulus resulting from the presence of another stimulus.
Masking margin: The additional amplification of the masker needed to completely mask the target sound.
Mass moment of inertia (MOI): The sum of the products formed by multiplying the mass of each component of a system by the square of its distance from a specified point.
Mastoid (bone): A hard, boney structure behind the ear in which the ear mechanism is housed.
Maximum-length sequence (MLS): A pseudorandom binary sequence that is used to measure impulse response of the transmission system.
McGurk effect: Auditory-visual illusion causing the misperception of a spoken phoneme. Occurs when the visual and auditory information disagree causing the speaker's mouth to appear to be uttering a different phoneme.
Mean time between failure (MTBF): For any device, a measure of the reliability of a component or system.
Mechanical impedance: The ratio of the effective pressure (force acting on a specific area of an acoustic medium or mechanical system) to he resulting effective velocity through or of this area. The units for mechanic impedance are Pa-s-m or dyne-sec/m, which are called the mechanic ohm (Ω).
Medial geniculate nucleus: A part of the thalamus that receives afferent auditory projections and from which they project to various parts of the cortex and cerebrum.
Median plane: The sagittal plane running through the midline and dividing human body into right and left parts.
Mel: A unit of pitch. A tone of frequency 1000 Hz and sound intensity of 40 dB (re 20 μPa) produces a pitch of 1000 mels. A tone of frequency 1000 Hz and sound intensity of 40 dB (re 20 μPa) has a pitch of 1000 mels.
Melatonin: A naturally occurring hormone found in most animals, including humans, which is important in the regulation of the circadian rhythms of several biological functions.
Memory: A hypothetical storage system. The ability or process of retaining and recalling what has been experienced and learned. Memory is frequently interpreted as an associative mechanism within the brain that relates present and past stimulations.
Mesopic: A state of visual adaptation which is between photopic (daylight) and scotopic (dark) conditions. Under mesopic conditions both the rod and cone photoreceptors are working.
Message: Meaningful information.
Meta-knowledge: Describe information that is one step removed from the actual knowledge itself, because it is derived primarily from sensors or displays. It is knowledge about knowledge and requires cognitive processing to convert it to useful knowledge.
Metacontrast: A type of backward masking in which the test stimulus and masking stimulus do not overlap spatially in the visual field.
Michelson contrast: One mathematical definition for contrast. It can have a maximum value of 1.0, which is the contrast of pure black stripes on a pure white background. It can have a minimum value of 0, which is the contrast of a neutral gray stripes on a neutral gray background; that is, a uniform gray field with no visible pattern.
Microdisplay: A small, usually 1-inch diagonal or less, electronic display device that can be suspended near the eye and viewed through magnifying optics or used with higher magnification optics to project an image.
Microphone: An electroacoustic transducer converting sound into electric current.
Microsleep: A brief, unintended episode of loss of attention associated with events such as blank stare, head snapping, and prolonged eye closure that may occur when a person is fatigued but trying to stay awake.
Middle ear: The main cavity of the ear; between the eardrum and the inner ear, containing the ossicles - three small bones that are connected and transmit the sound waves to the inner ear.
Mid-sagittal plane: [See **Median plane**]
Military occupational specialty (MOS): A job classification used by the U.S. Army and Marine Corps; the occupational specialty system uses a system of letters and numbers to identify general and specific jobs of

military personnel. The U.S. Air Force uses a system of Air Force Specialty Codes (AFSC). In the Navy, a system of naval ratings and designators is used along with Navy Enlisted Classification (NEC) system.

Minimum angle of resolution (MAR): A parameter used to describe visual acuity. It is the smallest angle between two objects for which they can be seen as two. For standard Snellen letters it refers to the width of one stroke of the letter.

Minimum audible field: Minimum audible sound pressure heard in a sound field.

Minimum audible pressure: Minimum audible sound pressure heard through the earphones.

Mistakes: A type of human error that involves incorrect intentions or plans.

Modified rhyme test (MRT): The accepted speech material used for determining speech intelligibility of a communication device.

Modiolus: The central bony pillar around which the spiral of cochlea winds.

Modulation: The systematic variation of one signal (carrier) caused by another lower frequency signal (modulating signal).

Modulation rate: A frequency of changes in a carrier caused by a modulating signal.

Modulation transfer function (MTF): The sine-wave spatial-frequency amplitude response used as a measure of the resolution and contrast transfer of an imaging system; a plot that describes the optical quality of an image-forming system, such as a camera or the human eye. This is not to be confused with the contrast sensitivity function (CSF), which describes visual performance, and includes neural image processing.

Monaural: Pertaining to, using, or involving the function of a single ear.

Monaural listening: Listening with a single ear.

Monaural mode: [See **Monotic mode**]

Monaural signal: An audio signal recorded with a single microphone located at a single ear of the listener or of the binaural dummy head.

Monochromatic: Description of a light that contains a single wavelength and therefore appears to have one particular colored hue. White light contains a mixture of many wavelengths (colors), therefore it is not monochromatic.

Monophonic signal: Audio signal that does not contain information about spatial distribution of sound sources.

Monophonic system: Means to record, transmit, or deliver a monophonic signal.

Montonic: Condition in which a sound stimulus is presented to only one ear.

Monotic display: An earphone display presenting acoustic signals to a single ear of the listener.

Monotic mode: A sound delivery mode in which auditory stimuli are delivered to a single ear of the listener.

Monovision: A vision correction method for people with presbyopia in which one eye is corrected for near vision and the other for far vision; the purposeful adjustment of one eye for near vision and the other eye for distance vision.

Most comfortable loudness (MCL): A loudness level of a specific auditory stimulus that is the most comfortable for the listener.

Motion aftereffect: The illusory impression, after prolonged viewing of movement in one direction, that a stationary object is moving in the opposite direction.

Motion box: The volume space in the cockpit within which the head-tracking sensors accurately can determine head position.

Motion parallax: A monocular depth perception cue based on the relative motion of object images that are at different distances from the observer.

Mouth simulator: A device simulating the acoustic characteristics of the head and mouth upon the radiated sound.

Multidimensional scaling (MDS): A statistical mapping technique in which the differences between N items are represented as points on n-dimensional map, where $n \ll N$. MDS technique is used to uncover dominant variables differentiating a given set of items.

Multiple resources theory (MRT): A theory which states that there are separate resources used for cognitive and perceptual activities based upon their different physical locations within the brain. These are: 1) Input perceptual or sensory modalities, 2) Central processing stages, 3) Response codes and 4) Channels of vision. Activities which share or overload these resources will cause greater interference and therefore poorer performance.
Multisensory: Refers to the use of more than one of the five human senses – vision, hearing, touch, smell, and taste.
Multistable perception: A perceptual phenomenon in which multiple perceptual interpretations are formed from a single sensory pattern. Multistable perception results from ambiguity in the sensory information that allows for more than one valid interpretation.
Multitalker noise (MTN): A noise made be multiple talkers speaking simultaneously.
Myelin: A fatty segmental covering on nerve fibers interrupted at the nodes of Ranvier. It accelerates the rate of propagation of the action potential along the nerve.
Myopia: Nearsightedness. A kind of refractive error in which an eye, viewing a distant object, focuses the image in front of the retina; near objects are seen more clearly than distant objects.

N

Nadir: The direction pointing directly below a particular location.
Naturalness: A sensation of an agreement of a given auditory image with expectations of the listener.
Nerve: A collection of neurons that are bundled together forming a communication pathway.
Network centric warfare: A military doctrine or theory of war pioneered by the U.S. Department of Defense that seeks to translate an information advantage, enabled in part by information technology, into a competitive Warfighting advantage through the networking of well-informed geographically dispersed forces; also called network-centric operations (NCO).
Neuron: A cell that is capable of transmitting electrochemical information within the nervous system of the body.
Neuroergonomics: A relatively new field that integrates neuroscience and ergonomics with the goal of improving human performance through an understanding of how humans process visual, auditory and tactile information in the real world.
Neutrality: The characteristic of an optical medium which denotes reasonably flat transmittance over the visible spectrum (e.g. gray tint).
Night myopia: A condition that can occur in the dark, when an eye incorrectly focuses too close (over-accommodates). This causes blurred vision that is optically similar to myopia.
Night vision goggle (NVG): While strictly defined as second generation I^2 light amplification devices, the term often is used for all I^2 systems.
Nit: A metric unit for luminance, which is equal to 1 candela per meter squared.
Noise: Any unwanted, meaningless, or interfering information.
Noise induced hearing loss: A hearing loss that is caused either by a one-time or repeated exposure to very loud sounds.
Nondeclarative memory: Nonconscious memories that influence behavior but are not explicitly recalled.
Nonsense syllable: A pronounceable combination of phonemes that do not make a word used to test speech articulation.
Note: A tone having a specific pitch and duration of which musical pieces are composed.
Numerical aperture (NA): The sine of the vertex angle of the largest cone of meridional rays that can enter or leave an optical system or element, multiplied by the refractive index of the medium in which the vertex of the cone is located.

O

Obscurant: Natural or made-made materials in the atmosphere that reduce or block visibility, e.g., smoke, fog, dust cloud, etc.
Occlusion (vision): A relative visual depth perception cue based on one or more objects blocking the view of one or more other objects.
Occlusion effect (audition): The perception of one's own voice as "hollow" or "booming" when the talker's ear canal is closed (covered). Occlusion effect is due to the amplification of bone conducted speech by the closed cavity of the outer ear.
Octave: A music interval produced by halving or doubling frequency.
Octave band: A band of frequencies where the highest frequency is the double of the lowest frequency.
Oculomotor nerve: The nerve that controls the movement of the muscles that move the eyeball.
Omidirectional device: A device in which the received or radiated signal is independent of the direction of observation.
Operational memory: [See **Working memory**]
Ophthalmoscope: A hand-held instrument used to inspect the internal parts of the eye.
Optic chiasm: The optic chiasm is where the optic nerves from the two eyes come together and retinal ganglion cell fibers from specific parts of the retina cross to the contralateral optic tract.
Optic disc: The portion of the optic nerve that is visible inside the eye; sometimes referred to as the "optic nerve head."
Optic nerve: The optic nerve is the third cranial nerve. It consists of a bundle of approximately 1 million retinal ganglion cell axons. The optic nerve exits the eye (or globe) posteriorly through the sclera at the lamina cribrosa.
Optic relay: A lens or lens system used to transfer a real image from one point within an optical system to another, with or without magnification.
Optic tract: The bundle of nerve fibers from the optic chiasm to the lateral geniculate nucleus.
Optical axis: The axis of symmetry of an optical system.
Optical resolution: The ability of an optical system to display all images as separate entities.
Optimum sighting alignment point (OSAP): Maximum eye clearance distance to obtain a full display field of view.
Orbit: The portion of the bony skull that surrounds and protects the eye and its supporting structures.
Organ of Corti: The sense organ of hearing located along the basilar membrane in the cochlea of the inner ear.
Organic LED (OLED): A thin film light-emitting technology that consists of a series of organic layers between two electrical contacts (electrodes) the acronym is derived from **O**rganic **L**ight **E**mitting **D**evice, **O**rganic **L**ight **E**mitting **D**iode.
Ossicles: Three small bones in the middle ear that transmit vibrations from the tympanic membrane to the cochlea.
Otitis media: An infection of the middle ear.
Otologically normal person: A person without any sign of disease of the ear.
Otosclerosis: A condition in which bone grows around the oval window and stirrup, causing the stirrup to become immobile, and resulting in conductive hearing loss.
Ototoxic substance: A substances that have a toxic effect on the structures of the ear causing temporary or permanent damage to organs of hearing and balance.
Outer ear: The visible part of the ear, consisting of the pinna or auricle and is made of skin and cartilage.
Over-the-counter (OTC): Refers to be able to purchase without a prescription.
Overlap: The lateral angle subtended by the intersecting individual binocular fields-of-view.
Overtone: A component of sound with a frequency higher than the fundamental frequency. In a harmonic sound, N^{th} overtone is (N+1) harmonic.

P

Panoramic NVG: A night vision system that provides a horizontal field-of-view in excess of 100 degrees.
Paracontrast: A type of forward masking in which the test stimulus and masking stimulus do not overlap spatially.
Parallax: The apparent displacement or change of position of an object when viewed from different places, such as with the alternate use of the right and left eye.
Partial: A pure tone component of a complex tone.
Pascal: A unit of sound pressure equal to one Newton per square meter,
Pentatonic scale: A music scale using only five tones, usually the first, second, third, fifth, and sixth tones of a diatonic scale. Pentatonic scale corresponds to playing only the black keys on a piano.
Perceived duration: A perceptual assessment of the duration of the sensory stimulus.
Perceived sound quality (PSQ): A degree of the listener's satisfaction with perceived auditory image; an esthetic (beauty) or utilitarian (utility) value of an auditory stimulus.
Perceptual conflict: Situation that occurs when information from various sensory modalities or from within a modality is ambiguous. Some examples of perceptual conflict are when a visual object and a sound event are not co-located or when two sounds are arriving from different directions but seem to be produced by the same sound source. Depending on expectations and motivation, the brain can interpret conflicting stimulation in one or another way and the interpretation may change in time.
Perceptual illusion: A distorted perception of reality caused by misinterpretation of the stimulation pattern by the brain. Perceptual illusions are stable and generally shared by most people. An example of perceptual illusion is a pitch of sound that does not correspond to any frequency component of sound. Perceptual illusions reveal how the brain normally organizes and interprets sensory stimulation.
Periodicity pitch: Pitch determined on the basis of the period of the waveform of a stimulus.
Periodicity theory of hearing: A theory of hearing stating that differences in sound frequency are coded in time and resolved by the central nervous system.
Peripheral masking: Masking that occurs when a masking stimulus is present in one ear and its masking effect is observed in the same ear.
Peripheral vision: Vision near the edges of the visual field. That is vision in the side of the visual field, far from straight ahead.
Percept: Something what the perceiver sees or hears as a result of stimulation, as opposed to the physical reality of the stimulation; a perceptual image of the reality; the mental construct build up from sensory data by a biologic organism.
Perception: A mental analysis of sensations based on prior experience and world knowledge to form a mental representation of the surrounding environment; awareness of the surrounding environment through sensory stimulation; the conscious mental registration of a sensory stimulus.
Permanent hearing loss: [See **Permanent threshold shift**]
Permanent threshold shift: A non-reversible hearing loss due to chronic, sudden, or extended exposure to intense noise.
Permanent memory: [See **Long-term memory**]
Personal space: The area that a person reserves for themselves during business interaction with other people (within a 1-meter (3.28-foot) radius).
Phase: The fractional part of the wave period. Phase is frequently expressed as an angle that is an appropriate fraction of 360°.
Phase difference: The difference in phase angle between two waveforms.
Phase locking: The tendency for nerve firings to occur at a particular phase of the stimulating waveform on the basilar membrane.

Phon: Unit of loudness level. A tone of frequency 1000 Hz and sound intensity of 40 dB (re 20 µPa) presented frontally to the listener has loudness level of 40 phons.

Phoneme: The smallest unit of speech.

Photopic: Referring to the spectral sensitivity of the human eye due to the activity of the cones of the retina; exhibited under moderate to high light levels of illumination.

Photoreceptor: The specialized cells in the retina designed to capture photons of light. The two types of photoreceptors are rods, which are more sensitive to low luminance conditions and motion, and cones, which are more sensitive to high luminance conditions and color.

Photorefractive keratectomy (PRK): A kind of refractive surgery in which the superficial cell layer of the cornea is removed to expose the underlying stroma, which is then ablated with a laser to reshape the cornea and change its refractive power. The superficial cells grow back within a few days following the procedure. PRK has largely been supplanted by LASIK refractive surgery.

Phototransduction: A complex biochemical process that occurs within the photoreceptors (rods and cones). It begins with the adsorption of light and proceeds through a series of complex steps to produce an electrical signal that is relayed to the next neuron along the visual pathway.

Physical-ear attenuation test (PEAT): An acoustical test used to establish baseline sound attenuation data for evaluating the level of hearing protection provided by a system. An alternative test is the Microphone in Real Ear (MIRE).

Piezoelectric transducer: A device that uses the piezoelectric effect to measure pressure, acceleration, strain or force by converting them to an electrical signal.

Pilot retained unit (PRU): The helmet part of the RAH-66 Comanche Helmet Integrated Display and Sight System (HIDSS).

Pilot's night vision system (PNVS): A forward-looking infrared sensor mounted on the nose of the AH-64 Apache aircraft which serves as an imagery source for pilotage and/or targeting.

Pincushion distortion: A type of optical distortion that causes images to bow inwards on the horizontal and vertical planes; an image aberration that compresses the centre of the field.

Pinna: The external part of the human ear attached to the head around the opening of the external auditory meatus; the most visible part of the ear.

Pitch: A perceptual attribute of sound that is described by pitch height, pitch chroma, and pitch strength. Pitch depends primarily upon the fundamental frequency and spectral content of sound, but it also depends to some degree of sound intensity and duration of sound. Pitch is one of the three major auditory attributes of sounds along with loudness and timbre.

Pitch class: The set of all pitches that are a whole number of octaves apart. (e.g., the pitch class C consists of the Cs in all octaves).

Pitch coding: Term referring to the peripheral mechanisms used to represent frequency information in the auditory system.

Pitch height: A perceptual attribute of sound in terms of which sounds may be ordered on a scale extending from low to high; the perceived dominant frequency of a sound.

Pitch strength: A degree to which a sound has a definable pitch. Noise has low pitch strength. Pure tones, narrow bands, and complex tones with harmonic frequency components have stronger pitch strength.

Pixel: Short for "picture element;" represents the smallest individually addressable image element.

Place theory of hearing: A theory of hearing assuming that the basilar membrane is high resolution frequency analyzer. According to the place theory of hearing pitch is determined by sensing the place on the basilar membrane that has maximum excitation.

Plasma display: Emissive gas discharge flat panel display technology which produces light when an electric field is applied across an envelope containing a gas.

Pleasantness: A degree of the listener's satisfaction with the auditory image caused by a given auditory stimulus.

Antonym of pleasance is annoyance.

Pointing accuracy: A measure of the angular error between the pilot's line-of-sight (when aligned with the sighting reticle) and the sensor's and/or weapon system's line-of-sight.

Polarity: The condition of being positive or negative with respect to some reference point. For a sinusoid, reversing polarity essentially shifts the phase by 180 degrees.

Posterior chamber: The back chamber of the eye formed by the back surface of the crystalline lens, the ciliary body and the inside of the globe.

Power spectrum: The distribution of energy emitted by a source in a unit of time along the frequency scale.

Precedence effect: The ability of the auditory system to determine the actual position of the sound source without being confused by early sound reflections. When two identical sounds, originating from two locations arrive within 5 to 40 ms of each other, only the first sound is heard and the location information of the second sound is suppressed.

Presbycusis/presbacusis: A hearing loss associated with aging that develops when hair cells within the cochlea wear out, causing a loss of sensitivity to sound.

Prominence ratio: A ratio of the power in the critical band centered on the tone of interest to the mean power of the two adjacent critical bands (ANSI S1.13, 2005).

Presbyopia: The age-related condition in which the eye looses it ability to accommodate, that is, focus on near objects. Usually by about age 40, a patient with perfect distance vision begins to have difficult focusing on fine print at a normal reading distance. Presbyopia can be compensated by reading glasses, bifocals or progressive lenses.

Prismatic deviation: A measure of the angular deviation in light rays that occurs when light rays pass through an optical medium, whose boundaries are nonparallel.

Process: A sequence of actions or events leading to a result.

Proprioception: The sense of body position

Protan: A type of hereditary color vision anomaly in which the patient is missing or has defective L-cones. Since L-cones have peak sensitivity in the long wavelengths range of the visible light spectrum, protans are sometimes called, "red weak, or red color blind."

Prototype: A standard representation of items in long term memory that correspond to a concept or category.

Psychoacoustics: A science of the relationships between auditory stimuli and auditory sensations.

Psychomotor: Relating to movement or muscular activity associated with mental processes.

Psychophysics: A science of the relations between external stimuli and sensory responses.

Psychophysiological measures: Real-time measurements of an individual that can give us an understanding of their physical and cognitive state. These include eye behavior (pupil diameter, blink and gaze), electroencephalography (EEG), heart rate, galvanic skin response (GSR), and functional near infrared imaging (fNIR).

Pulfrich phenomenon: A binocular visual effect in which a pendulum swinging in a plane parallel to the face appears to be swinging in an elliptical orbit. This occurs when a dark lens is placed over one eye while observing the pendulum. The brain receives a slightly delayed image from the covered eye, which causes the illusion of stereoscopic depth that varies with the pendulum's position.

Pupil: The hole or aperture in the center of the iris, which automatically adjust in size in response to light. The pupil plays an important role in the formation of the retinal image, directly controlling its illumination and quality of focus.

Pupil forming optical design: A system in which the eyepieces collimate virtual images that are formed using relay optics.

Pure tone: A sound consisting of only one sinusoidal component and no harmonics.

Purkinje shift: The shift in peak sensitivity from photopic to scotopic vision.

R

Rarefaction: In the physics of sound, the segment of the longitudinal wave where pressure is reduced, the other segment being compression.
Real image: An optical image formed when light rays converge such that the image can be projected onto a screen.
Receiver: Any device, system, or agent that responds to a specific signal.
Receptor: A specialized sensory cell that responds to a unique type of stimulus such as light, sound, or smell, and transmits this information to the central nervous system; a biological receiver.
Recognition: An act of assigning a stimulus or an object to a specific class or category.
Recruitment: An increase in loudness with increasing sound intensity at a rate greater than for normally hearing person.
Redundant signal effect (RSE): The speeding of reaction time (RT) with two rather than one stimulus.
Reference hearing threshold level: A mean standardized value of hearing threshold, expressed in dB (re 20 μPa), obtained under specific listening conditions for an adequately large number of ears of otologically normal listeners between the ages of 18 and 25 years.
Reference equivalent threshold force level (RETFL): A force level causing threshold sensation during bone conduction stimulation measured with the help of an acoustic couple or ear simulator.
Reference equivalent threshold sound pressure level (RETSPL): A sound pressure level causing threshold sensation during air conduction stimulation measured with the help of an acoustic couple or ear simulator.
Reflection: Return of radiation by a surface, without change in wavelength. The reflection may be specular, from a smooth surface; diffuse, from a rough surface or a combination of the two.
Refraction: The bending effect of incident rays as they pass from a medium having one refractive index into a medium with a different refractive index.
Refractive error: An optical aberration in which the eye has too much or too little focusing power. This causes blurred vision. The three most common and familiar refractive errors are myopia (nearsightedness), hyperopia (farsightedness) and astigmatism.
Refractive index: The ratio of the velocity of light in one medium to the velocity of light in the next medium.
Refractive power: The focusing effect of an optical component or system.
Refresh rate: The rate at which the picture on a display is redrawn.
Relay optics: An optical system which relays a real image from one plane within the system to another plane, usually for the purpose of magnification.
Retention: The act of keeping something in place (e.g., retaining a helmet on the head).
Residual pitch: [See **Periodicity pitch**]
Resolution: [See **Angular resolution**, **Frequency resolution**, **Intensity resolution**, **Optical resolution**, S**patial resolution**, and **Spectral resolution**]
Resonance: A tendency of a mechanic or electric system to oscillate at a certain frequency characteristic for this system.
Resonator: A system that stores energy at a specific frequency that depends on the resonator properties.
Response time experiment: A method of experimental psychology that measures how long it takes a person to complete a task.
Reticle: A fine line pattern which is located in one of the focal planes of an optical device.
Retina: The thin neural layer at the back of the eye responsible for the initial capture and neural processing of light entering the eye. The retina consists of 10 layers, including the neural components of ganglion cells, bipolar cells, amacrine cells and photoreceptor cells, some dividing membranes, and the retinal pigment epithelium. In the very center of the retina at the posterior pole of the eye is a small area called the macula, which contains a pit or indentation called the fovea where the most defined vision occurs. [See **Fovea** and **Macula**]

Retinal disparity: Misalignment of the two retinal images.
Retinal scanning display (RSD): A system which employs the use of a laser which scans the image directly onto the retina of the user's eye.
Reverberation: Multiple reflections of sound off a hard surface.
Rhyme test: A speech intelligibility test where the listener must choose the answer from a multiple options, all differing only by a consonant, (i.e., items that rhyme).
Risk-avoiding behavior: In human decision making people tend to prefer certain choices when all choices have gains.
Risk-seeking behavior: In human decision making people tend to prefer uncertain choices when all choices have some loss.
Rhodopsin: The light-sensitive receptor protein in the retina. When rhodopsin absorbs a photon of light it releases energy, leading ultimately to an electrical signal.
Rod: One of the two principal light receptors of the retina; highly sensitive to low variations in illumination but relatively insensitive to color differences. [See **Photoreceptor**]
Roll compensation: In HMDs, the capability of keeping the imagery aligned about the roll axis.
Roughness: A sensation of the amount of harshness in sound. Perceptual impression created by amplitude and frequency modulations in sound at high modulation rates, above about 20 Hz.

S

Saccadic eye movement: The sudden simultaneous movement of both eyes from one fixation point to another. The peak angular speed of the eye during a saccade reaches up to 1000 degrees per second. Saccades last from approximately 20 to 200 milliseconds.
Saccule: One of the two organs of balance (the other one is utricle) that responds to linear acceleration and head position relative to gravity.
Safety of flight (SOF): Refers to a process ensure that equipment is safe for air vehicle operation.
Sagittal plane: An imaginary plane passing through human body and dividing it into left and right parts.
Saturation: The purity or richness of a particular colored hue. For example, crimson is a highly saturated red hue, while pastel pink is a desaturated version of the same hue.
Saccade: A rapid shift in gaze that occurs when looking from one point to another.
Scan line: A single continuous narrow strip created by the scanning beam as it passes over the elements of a given area.
Schema: A cognitive structure that contains a mental model of how the world operates.
Sclera: The thick outer shell of the eye. The sclera is a thick collagen structure that protects the internal structures of the eye and serves an attachment point for the extraocular muscles of the eye. It covers 95% of the eye and connects to the cornea at the limbus at the front of the eye.
Scotopic vision: A state of visual adaptation under low illumination, such as during nighttime. Under scotopic conditions, light levels are below the working range for the cone, so only the rod photoreceptors are working.
See-through display: A display that presents imagery/symbology as a virtual image, allowing the viewer to look through the imagery (in varying degrees).
Selective attention: An act of purposely focusing conscious awareness onto a specific stimulus.
Semicircular canals: Three canals of the vestibular system that respond to angular acceleration of the body.
Semitone: A music interval equal 1/12 of an octave. On a piano a semitone is the interval between two adjacent keys.
Sensation: An awareness of external stimulation; an immediate reaction to external stimulation of a sense organ.
Sensation level (SL): An amount by which a specific sound pressure level or force level exceeds hearing threshold of a given listener for a specific sound.

Sense: A mechanism by which living organisms acquire information about the surrounding environment**.** The five human senses are vision, hearing, smell, touch, and taste.

Sensitivity: The capacity of a system or sensory organ to respond to stimulation; the smallest value of the stimulus that causes a specific reaction.

Sensorineural hearing loss: A hearing loss caused by damage to the sensory cells and/or nerve fibers of the inner ear.

Serial position effect: A memory-related term that refers to the tendency to recall information that is presented first and last (like in a list) better than information presented in the middle.

Shades of gray (SOG): Progressive steps in luminance where each step differs from continuous steps by a prescribed ratio, typically the square root of two.

Shape constancy: The recognition (visual perception) that the same object viewed at different distances, visual angles, and/or perspectives is the same objective shape.

Sharpness: An auditory sensation caused by acoustic energy concentrated in a narrow band around relatively high center frequency of sound.

Shell tear resistance: The property of the helmet shell to resist projectile damage.

Shift lag: A series of symptoms, to include excessive sleepiness, poor concentration, low productivity, and insomnia, associated with working and sleeping outside the normal circadian period for the activity. The clinical term is *shift-work sleep disorder*.

Shop replaceable unit (SRU): A maintenance term referring to a systems or module that cannot be replaced in the field; usually requires special tooling or fixturing for installation and alignment.

Short-term memory: A hypothesized system of human memory that holds information for durations ranging from one to thirty seconds.

Sight: [See **Vision**]

Signal: A change in the form or amount of energy intended to transmit information and by which information is transmitted.

Signal-to-noise ratio (SNR): The ratio of some measured aspect of a signal to a similar measure of concurrent noise expressed usually in a logarithmic form. The measured aspect, frequency range, and statistical properties of the signal and the noise should be stated explicitly.

Simulator sickness: Also referred to as cybersickness, a series of conditions which may include nausea, dizziness, and overall disorientation experienced during or after simulator training.

Simultaneous masking: Masking observed when a masking stimulus and a test signal occur at the same time.

Situation awareness (SA): A dynamic understanding of the individual (and vehicle or aircraft), environment, and status surrounding the individual. It is commonly divided into three levels: 1) the perception of the elements in the environment within a volume of time and space, 2) the comprehension of their meaning, and 3) the projection of their status in the near future. Lacking SA or having inadequate SA has been identified as one of the primary factors in accidents attributed to human error

Size constancy: The recognition (visual perception) that the same object viewed at different distances, visual angles, and/or perspectives is the same size.

Slaving lag: The latency of the sensor/weapon line-of-sight relative to the helmet line-of-sight. This includes the tracker computational time, data bus rate, and physical slaving time of the sensor/weapon.

Sleep deprivation: Refreshing sleep quality or quantity insufficient to support optimal daily functions; a wake-state-associated physiologic and/or psychological condition characterized by persistent sleepiness and/or impaired cognitive functioning.

Slips: Human errors in execution and/or storage of an action sequence.

Snellen visual acuity: A test of visual acuity commonly used and expressed as a comparison of the distance at which a given set of letters are read correctly to the distance at which the letters would be read by someone with clinically normal vision. Normal visual acuity is 20/20, which is equivalent to 0.29 milliradians (1 arcminute) of resolution.

Sone: Unit of loudness. One sone is the loudness of a pure tone of frequency 1000 Hz and a sound pressure level of 40 dB (re 20 µPa) presented frontally to the listener. The loudness of sound that is judged by the listener to ne N times that of 1 sone is N sones.

Sonification**:** A mapping of numerically represented relations in a non-acoustic domain to relations in an acoustic domain to facilitate interpretation of the relations in the non-acoustic domain; an interpretation of data sets by representing the data with sound; data-controlled sound generation.

Sound: The presence of a sound wave; an auditory sensation caused by a sound wave.

Sound field: [See **Acoustic fiel**d]

Sound intensity: The amount of sound power that travels through a certain area (W/m^2).

Sound intensity level (SIL): Ten times the logarithm to the base ten of the ratio of the time-mean sound intensity in a stated frequency band to the reference sound intensity of 10^{-12} W/m^2.

Sound pressure: The magnitude of change in the local pressure caused by the propagating sound wave.

Sound pressure level (SPL): Ten times the logarithm to the base ten of the ratio of the time-mean-square pressure of sound in a stated frequency band to the square of the reference pressure of 20 micropascals (µPa).

Sound quality: An objective or perceptual assessment of value of the auditory stimulus according to specific criteria. [See **Perceived sound quality**]

Sound wave: An acoustic wave at a frequency that is capable of being heard by a human listener. The nominal frequency range of acoustic waves that can be heard extends from 20 Hz to 20,000 Hz.

Soundscape: An acoustic environment. An environment created with sound.

Spaciousness: An auditory overall impression made by the surrounding acoustic space.

Spatial disorientation (SD): When the aviator experiences loss of situational awareness with regard to the position and motion of his aircraft or himself.

Spatial frequency: A parameter that corresponds with the size of black/white stripes in a pattern used to test vision. It is expressed as the number of cycles (black/white pairs) contained in one degree of visual angle. A high spatial frequency pattern contains many narrow bars while a low spatial frequency pattern contains a few broad bars.

Spatial resolution: A precision with which a person or a system can differentiate between two signals or objects presented from two different locations in space.

Spatial signal: [See **Stereophonic signal**]

Spatial vision: The aspect of vision concerned with how well we see images, without regard to color, motion, time, etc. The study of spatial vision considers how images are formed by the eye's optical system, and include subtopics such as visual acuity and contrast sensitivity.

Specific acoustic impedance: The ratio of the effective sound pressure to the effective particle velocity at a point of an acoustic medium or mechanical system. The units for specific acoustic impedance are Pa-s/m or dyne-s/cm^3; which are called the *rayl* in honor of Lord Rayleigh. If specific acoustic impedance is measured at a given point in a free progressive sound wave (free field) it is called the *characteristic impedance* and is equal to the product of the density of the medium and the speed of sound in this medium ($\rho_o c$).

Spectral density: [See **Spectral power density**].

Spectral envelope: The imaginary line connecting the maxima of the sound spectrum.

Spectral resolution: A precision with which a person or a system can differentiate between frequencies of two simultaneously presented sine waves.

Spectral transmittance: That amount of radiant energy passing through an optical component or system as a function of wavelength.

Spectral power density: The amount of the total power available in the specific bandwidth divided by the width of the bandwidth (W/Hz). The reference bandwidth is 1 Hz.

Spectrum: The distribution of amplitudes of the sinusoidal components of the complex waveform along the frequency scale. The spectrum can be instantaneous or averaged over time.

Spherical aberration: The failure of an optical component or system to focus all monochromatic paraxial and peripheral light rays at a single point; a rotationally symmetric higher-order optical aberration that causes a point of light to be imaged as a blurred circle. Within the Zernike system for classifying ocular aberrations, spherical aberration is labeled Z(4,0).

Speech: An expression of thoughts in spoken words.

Speech articulation: The act or process of producing speech.

Speech articulation (metric): A percentage of spoken phonemes or meaningless syllables correctly received by the ideal listener. A metric typically used to assess speech production by talkers or speech synthesizers.

Speech audibility (metric): A percentage of words or other meaningful units of the ideal speech transmitted through a given transmission system and correctly received by a listener.

Speech awareness threshold: The lowest level at which one detects speech fifty percent of the time.

Speech communicability (metric): A percentage of words or other meaningful units of a speech signal correctly received by a listener under a given set of conditions.

Speech detection threshold: [See **Speech awareness threshold**]

Speech intelligibility: Property of speech leading to its recognition.

Speech intelligibility (metric): A percentage of words or other meaningful units of speech correctly received by the ideal listener.

Speech intelligibility index (SII): A general term for objective measures if speech intelligibility used in ANSI S3.5-1997 (R2007).

Speech reception threshold: The lowest level at which one correctly identifies 50% of the words from a list of words (usually spondees).

Speech recognition: Ability of the listener to understand speech.

Speech recognition metric: A percentage of words or other meaningful units of the ideal speech transmitted over the ideal transmission system and correctly received by a listener.

Speech recognition threshold: [See **Speech reception threshold**]

Speech transmission index (STI): A measure of speech intelligibility (one of two described by ANSI S3.5-1997(R2007)) where speech is modeled by a special test waveform that is modulated by low-frequency signals. The depth of modulation of the received signal is compared with that of the test signal in each of a number of frequency bands and reductions in the modulation depth are associated with loss of intelligibility.

Speech transmissibility metric: A percentage of words or other meaningful units of the ideal speech transmitted through a given transmission system and correctly received by the ideal listener.

Spondee: A two-syllable word with equal emphasis on each syllable (e.g., ice cream, northwest, and airplane) used in determination of a speech reception threshold.

Spot size: The diameter in millimeters of a spot typically at 50 percent of its normal intensity level.

Steady-state sound: Sound with negligible fluctuations of level within the period of observation.

Stereocilia: Hair-like projections extending from the top of a hair cell.

Stereophonic signal: An audio signal that contains information about the spatial distribution of sound sources.

Stereophonic system: Means to record, transmit, or deliver a stereophonic signal.

Stereopsis: A very high-quality sense of depth perception that is unique to binocular vision. Stereopsis is stimulated when the brain detects slight differences (disparity) in the positions of objects seen by the two eyes. The image that the brain receives from each eye is slightly different because each eye views objects from a slightly different position.

Stiles-Crawford effect: The phenomenon by which light seems brighter if it enters the eye through the center of the pupil rather than the peripheral pupil.

Stimulus: An agent, action, or environmental change that causes or is intended to cause a reaction. A stimulus is a physical realization of a signal. [See **Signal**]
Streaming: The task of analyzing complex sounds and partitioning them into auditory streams; the process of separating sound elements into different auditory objects is called *auditory stream segregation*. Conversely, the process of assigning different sound elements to a single object is known as *auditory stream integration*.
Stress: A nonspecific response of the body to a demand, which can physical, environment, psychological, etc.
Stressor: Any agent that causes stress to an organism; any stimulus or condition that causes physiological arousal beyond what is necessary to accomplish an action.
Stroma: The thick central layer of the cornea consisting of lamellar sheets of collagen. The arrangement of the collagen lamellae provides strength as well as transparency to the cornea.
Superior olivary complex: A group of nuclei of the central auditory nervous system located in the pons and playing a major role in coding sound localization information.
Suppression: The unconscious inhibition of an eye's retinal image. The condition in which sensations from one or both eyes is voluntarily or involuntarily ignored.
Supra-aural earphone: Earphone that rests on the external ear against the pinna.
Symbol: An individual representation of information.
Symbology: A set of symbols.
Synchrony: The state of two or more events occurring at the same time.
Synthetic vision: A system that uses various sensors to augment the viewer's view of the outside world.
System: A structure of elements operating together to accomplish a predescribed end result.
Systems safety assessment (SSA): A system analysis which addresses safety and health issues.

T

Tarsal plate: The cartilage-like plate within the upper and lower eyelids that provide rigidity and shape to each eyelid.
Technology readiness level (TRL): A measure used by some U.S. government agencies and many major world's companies (and agencies) to assess the maturity of evolving technologies (materials, components, devices, etc.) prior to incorporating that technology into a system or subsystem.
Tectorial membrane: The gelatinous membrane that lies over the hair cells of the organ of Corti.
Telepresence: Enables the operator to participate in activities at remote locations.
Temporal envelope: The imaginary line connecting the maxima of the sound waveform.
Temporal integration: [See **Temporal summation**]
Temporal masking: A masking effects that occur when the masker reduces sensitivity to the sounds that is presented immediately preceding or following the masker.
Temporal resolution: The precision of sensation with respect to time; the ability to detect rapid changes in auditory or visual information.
Temporal summation: Sensory addition of the effects of a single stimulus or several stimuli over a short period of time.
Temporal vision: The time-related or time-dependent aspects of vision; it is closely related to motion perception.
Temporary hearing loss: [See **Temporary threshold shift**]
Temporary threshold shift: A temporary reduction in hearing sensitivity due to exposure to intense levels of noise.
Terminal threshold: A sensory threshold above which specific sensation does not exist or changes its character.
Thermoneutral zone (TNZ): The temperature range when metabolic heat production does not need to be increased to maintain thermostability.

Thermoplastic liners (TPL™): A liner developed by Gentex Corporation, Carbondale, PA, consisting of two to five plies of thermoplastic sheets covered with a cloth cover, designed to improve comfort and to alleviate helmet fitting problems.

Thermoregulation: The regulation of body temperature; the ability of an organism to keep its body temperature within certain boundaries.

Three-dimensional (3-D) audio: Variety of signal processing techniques that simulate sounds coming from all directions using a single pair of audio transducers.

Threshold of pain: A sound pressure level beyond which sound causes pain.

Timbre: A perceptual attribute of sound in terms of which a listener can judge that two sound that are similarly presented and have the same loudness, pitch, and subjective duration are dissimilar. Timbre is the main perceptual property of sound that guides in sound source recognition.

Time error: An error in a sensory judgment resulting from sequential presentation of stimuli.

Tonality: In music, tonality refers to the tonic, or the key in which a piece was written. In psychoacoustics, tonality refers to the degree to which a sound has a particular pitch.

Tone: An audible sound of specific pitch and periodic waveform; an interval of two semitones.

Tone chroma: [See **Pitch class**]

Tone-to-noise ratio: A ratio of the power of a specific pure tone to the power of the critical band centered on that tone. A method of specifying the audibility of a specific tonal component embedded in noisy background.

Tonic: [See **Key**]

Tonotopic: The one-to-one correspondence between specific sound frequency and its representation along the basilar membrane of within a specific neural structure of the auditory system.

Tracking: A helmet mounted display enhancement in which the line-of-sight-direction of the aviator is continuously monitored, and any change is replicated in the line-of-sight-direction of the aircraft-mounted sensor.

Tragion: An anthropometric point situated in the notch just above the tragus of the ear.

Tragus: A small cartilaginous part of pinna that is immediately anterior to the opening of the ear canal.

Transducer: A device for converting one form of energy into another (e.g., acoustic energy into electric energy).

Transfer function: The output versus input response characteristics of a device expressed either mathematically or graphically.

Transient sound: A state of motion that lasts only a very short time.

Transmeridian: Refers to crossing a number of time zones.

Transmission (T): An act or process of moving a certain quantity through a medium or a communication channel.

Transmissibility: The ratio of the magnitude of a certain transmitted quantity received after transmission to the magnitude that was sent. In optics, a ratio of the amount of the radiant flux received after propagating through a medium or a body to the amount that was sent; usually expressed as a percent.

Transmission coefficient: [See **Transmissibility**]

Transmission loss (TL): [See **Transmissibility**]

Transmitter: Any device, system, or agent that sends a signal out.

Transverse plane: [See **Horizontal plane**]

Tritan: A rare color vision anomaly in which the patient has abnormal sensitivity for short wavelengths. These patients are sometimes referred to as having a blue-yellow color vision defect.

Troland: A metric unit for retinal illumination. It describes the amount of light falling on the retina.

Tympanic membrane: A membrane separating the outer ear from the middle ear converting acoustic waves of the outer ear into mechanical vibration of the middle ear.

U

Ultrasound: An acoustic wave of a frequency higher than the upper limit of human hearing; usually considered to be a sound having frequency higher than 20 kHz.
Underload syndrome: A lack of stimulation (such as a boring job) can result in depression and health problems, e.g., headache, fatigue and recurrent infection.
Unmanned aerial vehicle (UAV): Remotely controlled or autonomous aircraft used for surveillance and strike missions.
Update rate: The rate at which the position of the helmet/head display or signal is sampled and used to provide drive inputs to the head-slaved sensor or display, usually expressed as a frequency (in Hz).
Utricle: One of the two organs of balance (the other one is saccule) that responds to linear acceleration and head position relative to gravity.

V

Vacuum fluorescent display (VFD): A flat vacuum tube emissive display device that uses a filament wire, control grid structure, and phosphor-coated anode.
Ventriloquism effect (VE): The result of the domination of visual localization over auditory location.
Vernier acuity: A type of visual acuity task in which the patient tries to detect a small offset of one line relative to another.
Vergence: The symmetric movement of the eyes toward or away from each other.
Vestibule: The part of the bony labyrinth that contains two organs of balance. The utricle and saccule are located within the vestibule and the semicircular canals begin and end at the vestibule.
Vestibulocochlear nerve: A nerve connecting inner ear with the brainstem.
Vestibulo-ocular reflex (VOR): A reflex that causes the eyes to rotate in the opposite direction as a head tilt. This helps to stabilize vision.
Vibration: An oscillation where the quantity is a parameter that defines the motion of a mechanical system.
Video: Pertaining to a visual signal encoded in electrical form and to the means of its transmission.
Virtual image: An optical image formed when light rays do not actually converge and cannot be projected upon a screen.
Virtual pitch: [See **Periodicity pitch**]
Virtual reality (VR): A synthetic (computer-generated) environment.
Vision: The act or power of sensing with the eyes.
Vista space: The viewable space around a person that is approximately 30 meters out and beyond.
Visual acuity: A measure of the ability of the eye to resolve spatial detail; a description of the sharpness or quality of spatial vision. [See **Snellen acuity**]
Visual angle: The angle subtended by an object at the eye or retina.
Visual capture: The phenomenon is which visual perception dominates when visual cues and other sensory cues – auditory, proprioceptive, haptic, etc. – are in direct conflict.
Visual cortex: Located at the posterior portion of the brain, this is the part of the brain where vision occurs; also referred to as the "occipital cortex."
Visual field: The extent of space that is visible to an eye while it is looking at one particular point; a plot of the remaining unaided field of vision available when wearing a helmet, helmet-mounted display, etc.
Visually coupled system (VCS): A system in which the line-of-sight of the user's eyes (or head0 is continuously monitored, and any change is replicated in the line-of-sight-direction of the sensor.
Visual adaptation: The automatic adjustment of the pupil in response to different levels of ambient illumination.

Visual search: An experimental method for measuring human behavior. The task is for an observer to find a designated target among a field with other information.
Vitreous humor: The fluid or gel body that fills the posterior chamber of the eye.
Vocal folds: A stretchable pair of bands of mucous membrane that project into laryngx. When air passes up from the lungs though stretched vocal folds it produces acoustic event that is the basis for all vocal (voiced) sounds of speech.
Vocal tract: The airway (tube) used in speech production. It consists of the upper part of the respiratory tube from larynx up including pharyngeal, mouth, and nasal cavities.

W

Warfighter: All military personnel trained to engage in combat operations.
Wave: A disturbance that travels through a medium by virtue of the elastic properties of that medium.
Weber fraction: A relation between the intensity of a standard stimulus and the intensity of a stimulus required to produce a just noticeable difference in perception.
Weber's law: A rule stating that a just-noticeable difference in a stimulus is proportional to the magnitude of the original stimulus.
Weber-Fechner law: Equal stimulus ratios correspond to equal sensation differences. An empirical law stating that sensation changes in equal artitmetic increments in response to geometric changes of the stimulus.
Whole-body vibration (WBV): Vibration that is transmitted to a workers body from vibrating surfaces on which a worker stands or sits.
Working memory: A term used for short-term memory that underscores its use as a working buffer for incoming information as well as information retrieved from long-term memory. [See **Short-term memory**]
Workload: The hypothetical relationship between a group or individual human operator and task demands.

Y

Yerkes-Dodson law: An empirical relationship between arousal and performance, stating that performance increases with physiological or mental arousal, but only up to a point, beyond which performance decreases.

Z

Zenith: The direction pointing directly above a particular location.
Zonule: The thin fiber-like structures that suspend the crystalline lens within the eye. These fibers are connected to the ciliary muscle, which controls tension on the fibers to allow for accommodation of the crystalline lens.

INDEX

D

G

H