

AD-A073 598

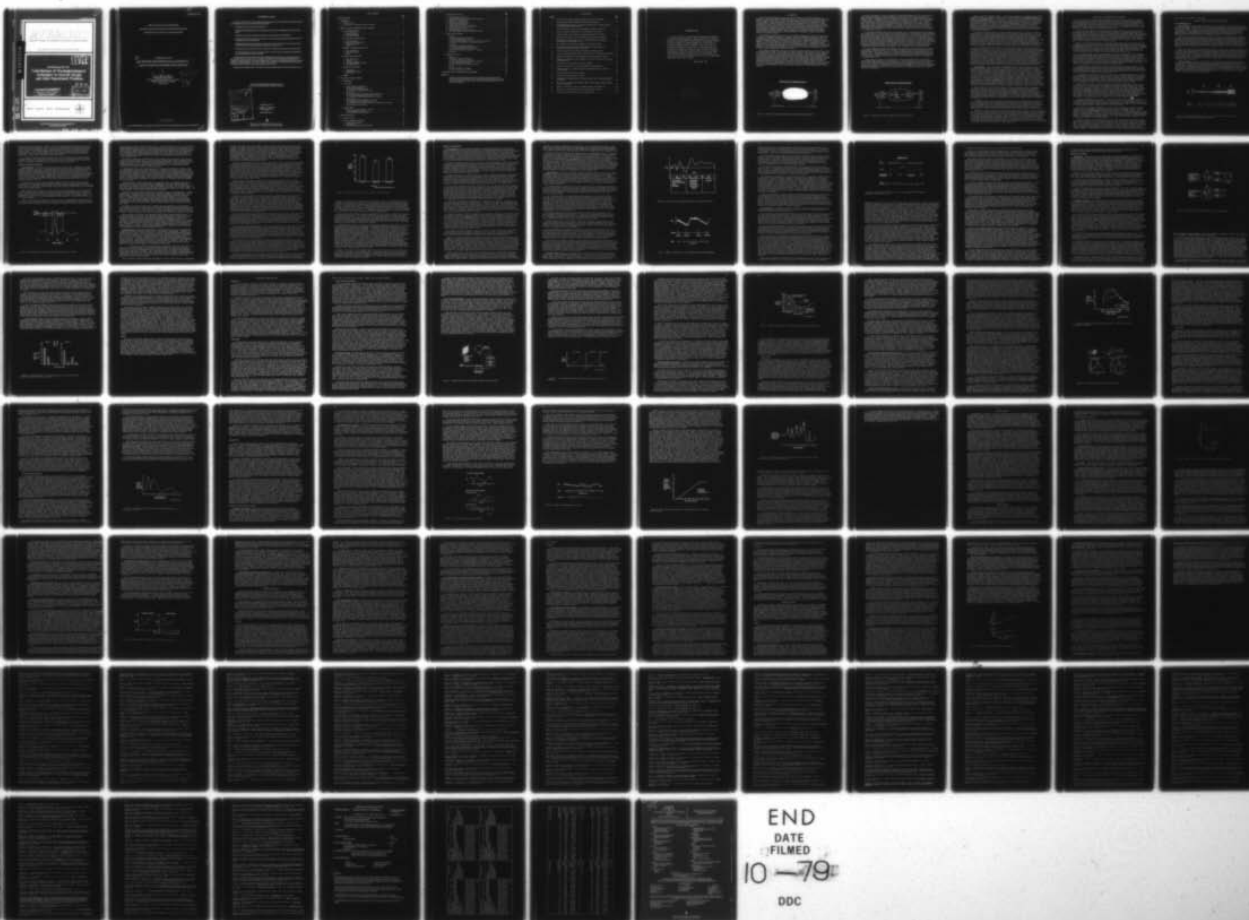
ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 5/10
CONTRIBUTIONS OF PSYCHOPHYSIOLOGICAL TECHNIQUES TO AIRCRAFT DES--ETC(U)
JUL 79 R D O'DONNELL

UNCLASSIFIED

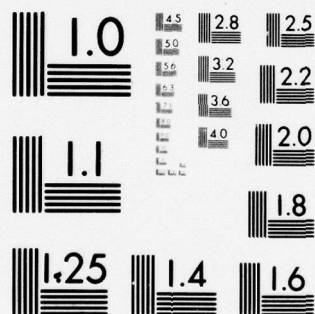
AGARD-A6-244

NL

| OF |
ADA
073598



END
DATE
FILMED
10-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AG 244

CONTRIBUTIONS OF PSYCHOPHYSIOLOGICAL TECHNIQUES TO
AIRCRAFT DESIGN AND OTHER OPERATIONAL PROBLEMS

DDC FILE COPY

AD A 073598

AGARD-AG-244

AGARD-AG-244

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

LEVEL II

AGARDograph No. 244

Contributions of Psychophysiological Techniques to Aircraft Design and other Operational Problems

DISTRIBUTION STATEMENT A

Approved for public release
Distribution Unlimited

DDC

SEP 11 1979

NORTH ATLANTIC TREATY ORGANIZATION



DISTRIBUTION AND AVAILABILITY
ON BACK COVER

20 00 10 104

NORTH ATLANTIC TREATY ORGANIZATION
ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT
(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

6 AGARDograph No.244

CONTRIBUTIONS OF PSYCHOPHYSIOLOGICAL TECHNIQUES TO
AIRCRAFT DESIGN AND OTHER OPERATIONAL PROBLEMS

by

10 Robert D. O'Donnell Ph.D
Lieutenant Colonel, USAF

6570 Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base
Ohio 45433

11 Jul 49

12 94p.

400 443

AB

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced directly from material supplied by AGARD or the author.

Accession For	
NTIS GAKI	☒
DDC TAB	☐
Unannounced	☐
Justification	
By _____	
Distribution/	
Availability Codes	
Dist.	Avail and/or special
A	

Published July 1979

Copyright © AGARD 1979

All Rights Reserved

ISBN 92-835-1325-8



Printed by Technical Editing and Reproduction Ltd
Harford House, 7-9 Charlotte St, London, W1P 1HD

TABLE OF CONTENTS

	<u>Page</u>
INTRODUCTORY QUOTE.	1
I INTRODUCTION.	2
A. Purpose and Plan	4
B. Brief History of Psychophysiological Techniques	5
II SURVEY OF SPECIFIC PSYCHOPHYSIOLOGICAL TECHNIQUES	6
A. THE ELECTROMYOGRAM (EMG).	6
1. Basic Methodology	6
2. Analysis Techniques	6
3. Early Applications.	7
B. THE GALVANIC SKIN RESPONSE (GSR).	7
1. Basic Methodology	7
2. Analysis and Interpretation	8
3. Early Applications.	8
4. Summary	8
C. MEASURES OF CARDIOVASCULAR FUNCTION	9
1. Basic Methodology	9
2. Early Applications.	10
3. Summary	12
D. MEASURES OF BRAIN FUNCTION.	13
1. Basic Methodology	13
2. Analysis.	13
3. "Evoked" Responses.	14
4. Contingent Negative Variation	14
5. Early Applications.	16
E. MEASURES OF EYE FUNCTION.	19
1. Introduction.	19
2. Measurement Techniques.	19
3. Applications.	21
F. SUMMARY	24
III EVALUATION OF SENSORY FUNCTION.	25
A. INTRODUCTION.	25
B. VISUAL INPUT.	25
1. Visual Sensitivity Measurement.	26
(a) The Electroretinogram	26
(b) The Cortical Evoked Response.	27
(c) Modulation Transfer Function Area	29
2. Visual Acuity	30
(a) Acuity and Spatial Frequency Contrast Sensitivity	31
(b) Transient Evoked Response Measures of Acuity.	31
(c) Steady-State Evoked Response Measures of Acuity	32
3. Color Vision.	34
(a) Transient Visual Evoked Response to Color	35
(b) Steady-State Evoked Response to Color	35
(c) Determination of Color Responses with Evoked Response Feedback.	36
4. Critical Flicker Fusion (CFF)	37
C. AUDITORY INPUT.	39
1. Measures of Absolute Auditory Threshold	39
(a) Psychophysical Techniques	39
(b) Psychophysiological Techniques.	40
IV COGNITIVE FUNCTION.	46
A. THOUGHT PROCESSES	46
B. STIMULUS MEANING AND RELEVANCE.	47
1. Galvanic Skin Response.	47
2. Pupilometry	47
3. Electroencephalographic (EEG) Measures.	47

	<u>Page</u>
C. PROCESSING AND DECISION.	49
1. Evoked Responses and Decision.	49
2. Interhemispheric Measures of Cognitive Processing.	49
3. Measures of Problem Solving.	50
4. Measures of Confidence Level	50
5. Signal Detection Performance	50
6. Reaction Time Performance.	51
7. The Respiratory Cycle and Reaction Times	51
8. Cardiac Activity and Reaction Time	51
9. The EEG and Reaction Time.	51
10. The Evoked Response and Reaction Time.	52
11. Contingent Negative Variation and Reaction Time.	53
D. ATTENTION AND VIGILANCE.	53
1. ATTENTION.	53
(a) Electroencephalographic Measures	54
(b) Evoked Response and Attention.	54
(c) The Contingent Negative Variation and Response Readiness	55
(d) Measures of Preparation for Voluntary Motor Acts	55
2. VIGILANCE.	56
(a) Cardiac Measures and Vigilance	56
(b) The Electroencephalogram and Vigilance	56
(c) The Evoked Response and Vigilance.	57
(d) Multiple Physiological Indices of Vigilance	57
3. MOTIVATION	58
E. WORKLOAD AND FATIGUE	58
1. WORKLOAD	58
(a) Cardiac Measures of Workload	58
(b) Pupillary Dilation and Workload.	59
(c) Electroencephalographic Measures of Workload	60
(d) Voice Analysis and Workload.	61
(e) Other Techniques for Assessing Workload.	61
2. FATIGUE.	61
(a) Specific Measures of Fatigue	61
(b) Field Measurement of Fatigue	61
V THE FUTURE ROLE OF PSYCHOPHYSIOLOGICAL MEASUREMENT	63
REFERENCES.	64

All references cited as being available from the National Technical Information Service (NTIS), Defense Documentation Center (DDC), NASA, or other US Military Installations should be requested from the National Technical Information Service, 5285 Port Royal Road, Springfield Va, 22161. All references are unclassified.

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	Traditional performance measurement paradigm using behavioral measures.	2
2	Performance measurement paradigm using physiological measures.	3
3	Electromyographic activity at low, medium and high levels of muscular tension, and typical integrated EMG patterns.	6
4	Principal peaks in a typical electrocardiogram, and sources of activity.	9
5	Heart rate results during aircraft carrier operations in Viet Nam.	12
6	Idealized visual evoked response showing major components and sections	15
7	Schematic representation of the contingent negative variation paradigm	15
8	EEG and state of consciousness of a pilot during high-G bomb run (Based on Sem-Jacobsen and Sem-Jacobsen, 1963)	17
9	Two techniques for recording eye movements for specific purposes	20
10	Eye fixation time on four flight instruments for two pilots during two successive ILS approaches (Based on Cox, 1975)	23
11	Schematic procedure for obtaining the steady-state evoked response	27
12	Plot of steady-state amplitude versus contrast ratio for three spatial frequencies	28
13	Schematic representation of the modulation transfer function area concept.	30
14	Amplitude of steady-state evoked response as a function of check size for normal and amblyopic eye.	33
15	Rapid refraction using steady-state evoked response.	33
16	Objective test for color vision using steady-state evoked response	36
17	Peak sensitivities for steady-state evoked response amplitude as a function of flicker rate.	38
18	Late and middle auditory evoked responses.	41
19	Frequency following response in the EEG.	42
20	Use of auditory evoked response to rapidly presented tones in assessing auditory thresholds	43
21	The auditory brain stem evoked response (BSR) to click stimuli, showing probable source of major peaks.	44
22	Amplitude of P300 as a function of stimulus relevance and memory load.	48
23	Reaction time and P300 latency under accuracy and speed criteria	52
24	Reaction time and P300 latency related to memory load.	60

INTRODUCTORY QUOTE

The key to the solution of many problems of human factors engineering is held by physiological psychology. The understanding of man's elemental capacities for the discrimination, identification, and processing of signals, for short-term memory, for patterned movements, for reaction time, etc., which are the foundation stones of performance theory, rest heavily on physiological psychology at a basic level. At a more applied level, the alleviation of environmental stresses, the design of protective equipment, the planning of optimal cycles of work and rest, and the solution of many special problems relating to the design of seats, lighting systems, and safety devices often require extensive data of a sort which the physiological psychologist is most likely to possess.

PAUL M. FITTS, 1963

INTRODUCTION

The field of human engineering is undergoing radical changes in orientation, methodology, and philosophy. Historically, there has always been somewhat of a discrepancy between the needs of the design engineer, expressed in questions concerning the system being developed, and the techniques available to the psychologist to answer those questions. Behavioral research techniques, utilizing reaction time, accuracy measures, tracking describing functions, and other dependent variables, were able to supply enough answers as long as the questions were reasonably straightforward. Thus, when interest centered in such things as optimal display and control location, sensory acuity and thresholds, ergonomic capacities, and other single parameters, the questions could be attacked successfully using relatively crude behavioral measures. Handbooks dealing with the input-output relationships between such variables as reach envelopes, strength requirements, sensory demands, and even perceptual criteria were produced (Van Cott and Kincade, 1972) and have contributed enormously to the design of efficient work spaces, vehicles, and total environments.

The intrinsically limiting problem with so-called behavioral approaches is that they typically treat the human as an amorphous "black box" or, at best, a series of undifferentiated processes. Input-output relationships are studied by manipulating independent variables of interest and observing effects on sensitive dependent variables of overt behavior. Certainly there is nothing basically wrong with such an approach. It forms the foundation of the scientific method itself. However, if the human is treated as a "black box", it is evident that many processes intervene between the input, as expressed in energy or stimuli impinging on the person, and the output, as expressed in behavior. In terms shown in Figure 1, the problem is to relate those input variables to the output variables over the very long series of events which intervene. As any behavioral experimenter can attest, this can be a difficult and delicate task, often producing confusing and apparently contradictory results.

This behavioral paradigm works best when the question being asked constrains the "black box" to a very few possible modes of action. In such cases, the person is really not free to behave in very complex ways (or if subjects do so, their data is eliminated from the experiment). This explains why human engineers have been most successful in scientifically studying sensory and motor systems where the input-output relationships are usually rather straightforward, even if tauntingly elusive. The paradigm begins to falter when the questions being asked deal either with complex interactions between sensory and/or motor systems, or when questions deal with cognitive behavior. Even in their simplest forms, questions dealing with information processing, problem solving, decision-making, and other cognitive functions have proven extremely difficult for the human engineer to handle. Questions of vigilance, fatigue, attention, workload, etc., while stimulating great interest in basic science, are perplexing and disturbing to the applied scientist, and are providing the basis upon which human engineers are re-evaluating the adequacy of their techniques.

Performance Measurement

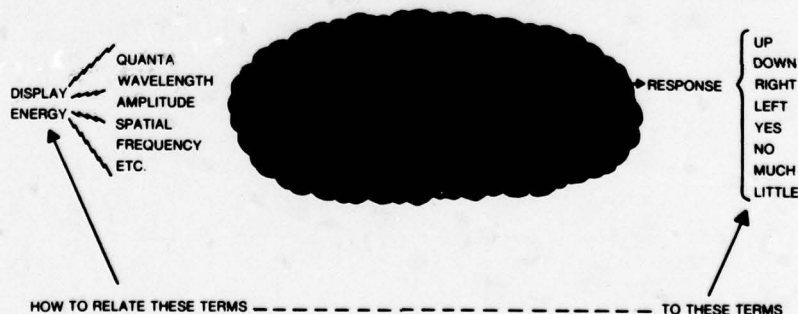


Figure 1. Traditional performance measurement paradigm using behavioral measures.

Nowhere is the impetus for such a re-evaluation of methodology stronger than in the area of aircraft design. Aerospace systems are undergoing dramatic conceptual changes which are not only supplying a quantum increase in complexity, but also are changing the very nature of the demands placed on the operator (O'Donnell, 1975). For example, the advent of digitally controlled aircraft systems, with extensive use of on-board computers, has reduced much of the manual-control requirement of the pilot. It is conceivable that an entire flight, from taxi to engine shutdown, could be programmed, and flown without a single manual control input from the operator (Krippner and Fenwick, 1975). Thus, the load placed on the pilot is shifting from one of sensing and responding in a closed-loop control manner, to one of monitoring, interpreting, problem solving, and commanding. These executive, cognitive functions are, of course, much more difficult to assess behaviorally. Most behavioral metrics assuming the "black box" approach attempt to do so indirectly, if at all.

It is becoming increasingly clear that if such questions are to be answered accurately, the "behavior" of the operator will have to be defined in a much more microscopic way. Put differently, when the questions reach a certain level of complexity, interactions within the black box itself may make the simple input-output relationship uninterpretable. For example, it is no longer sufficient to know that performance with a given dial design is "better" than with another candidate design. In such a case, it may still be possible that the operator is obtaining better performance by increased workload, or that the improved performance is utilizing a processing skill which is required by another task in the aircraft. In such a case, the "better" dial may in fact lead to a long-term unexpected overall decrement in performance, causing catastrophic failure within the total man-machine system. Such a decrement, caused by interactions so subtle that they are virtually impossible to anticipate in the laboratory, make it imperative that the behavioral approach be supplemented by techniques permitting more detailed study of mechanisms within the operator per se.

Only by knowing HOW the operator is performing a task can the various tradeoffs involved in a complex system be made. Understanding this has led scientists and engineers to open up the "black box" which is human performance, and to search for ways to interpret the processes going on between stimulus input and behavioral output. Theorists in many areas are therefore postulating stages and processes which intervene in sensory, motor, and cognitive functions. In all such theories, there is a universally recognized need for objective ways to measure such intervening processes, and correspondingly, there is also interest in any techniques which do so.

A logical place to search for such information is in the indirect manifestation of physiological processes underlying the behavior in question (Figure 2). Electrical measures of the central or peripheral nervous systems, biochemical measures of endocrine function, direct monitoring of muscular or cardiac activity, and a variety of other techniques should all indicate the manner in which the human is responding, and therefore supply needed information on the "process" of a given response. In this way, the "box" can be opened, if only slightly, and interactions which previously were uninterpretable might now be understood and studied scientifically.

Performance Measurement

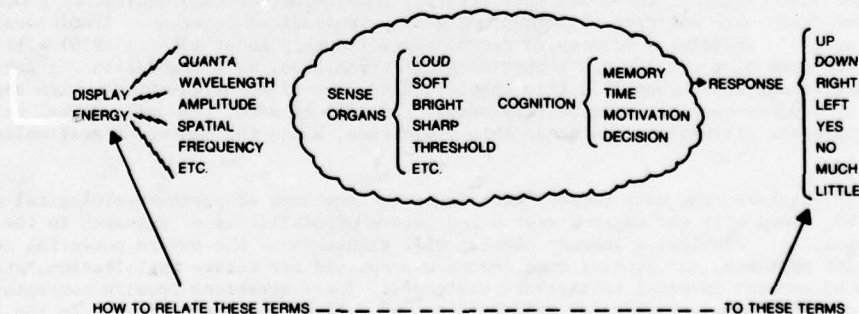


Figure 2. Performance measurement paradigm using physiological measures.

Obviously, there is nothing mysterious or even theoretically different in using physiological measures as opposed to "behavioral" measures. In fact, it is often overlooked that a physiological response to a stimulus is a true behavior produced by the subject. A cardiac acceleration, or a change in brainwaves, is just as much a behavioral response as a button press or a verbal report, and should have just as much (or as little) validity and stability as any overt response (Thompson, 1967). Therefore, it is seen that the difference between Figure 1 and Figure 2 is one of measurement specificity, not involving a qualitative change in what is measured. The change is being brought out because of the obvious need to measure behavior at a more microscopic level than has been done generally. Many researchers feel that if physiological processes can be shown to be manifestations of processing "stages" in the overall behavior, they should exhibit less variability than the final performance, and therefore should be more valuable as aids in human engineering.

As tempting as the above argument for increased stability and specificity is, however, technical problems have, in the past, prevented the scientist from capitalizing on these possibilities. With some notable exceptions, efforts to use techniques such as the electroencephalogram (EEG), the galvanic skin response (GSR) or the electromyogram (EMG) succeeded in demonstrating only that the variability of the measurements was too large, individual differences were excessive, application of the techniques to operational problems was impractical, or data obtained in laboratory conditions was uninterpretable in terms of real world situations. A healthy skepticism, or at most a guarded optimism, was justified by the numerous failures to find significant applications for these and other psychophysiological measurement techniques. It began to appear that there was an intrinsic between- and within-subject variability in many measures of physiological function, and that this variability introduced large experimental instability.

For basic investigations, where conditions could be precisely controlled or where large numbers of subjects are available in repeated measures or independent group designs which can be used in well-defined environments, such variability could be tolerated. Many successful studies of basic physiological correlates of behavior were carried out in laboratories around the world (Boiko, 1957; Masterson and Berkley, 1974; Riggs and Wooten, 1972; Thompson, 1967). However, when attempts were made to use physiological measures in operational settings, or to predict real-world behavior from such measures obtained in the laboratory, little success was achieved. Either the measurement variability was so great that large numbers of subjects were necessary in order to permit statistical evaluation, or, even when statistical significance was demonstrated, the amount of variance accounted for by the physiological measures in the total system context was so small that it was useless in a practical sense. Even in the case of successful experiments, extrapolation to real world specific situations was frequently extremely difficult. While producing valuable knowledge concerning basic mechanisms, expression of the results in limited physiological terms such as heart rates or muscle changes made no sense to the world of human engineering and operational planning.

Purpose and Plan. It is the premise of this AGARDograph that over the past five to ten years a subtle but highly significant change in this state of affairs has occurred. Technical developments, both in hardware, analysis techniques, and psychophysiological theory have resulted in the ability to design reliable and valid experiments which can reasonably be applied to real world situations. Such experiments have in fact already been used in some operational contexts, and the incidence of their use is increasing. The present work will attempt to survey and evaluate this trend, and to provide a speculative estimate of its future direction. It is our express purpose to document as many techniques as possible of proven or potential value to applied areas of human engineering in general, noting instances of their application to the human factors of aircraft design in particular. Throughout, emphasis will be given to those techniques which offer hope of revealing the processes intervening between stimulus and response, not simply those providing another correlate of observable behavior. In most cases it will be clear that if a question can be answered by using observable overt behavioral responses, it should be. There is no need or intent to duplicate or supplant existing behavioral metrics. In the end, it will be seen that the new capabilities represented by these psychophysiological techniques supplement existing metrics nicely, with little overlap.

The next chapter will present a brief overview of the better known applications of physiological measurement in the past. Many of these are appropriately considered strictly medical or biological applications, since they were not directly concerned with psychological function. Traditional data acquisition and analysis procedures used up to recent times (roughly about 1965 to 1970) will be reviewed in this chapter for each of the major psychophysiological techniques, as a foundation for subsequent discussion of more recent approaches. In this chapter, also, the kinds of questions which appropriately can be answered by psychophysiology will be considered. As will be seen, this will reveal the need to limit the scope of the present discussion to a manageable proportion, since the number of available techniques is quite large.

In subsequent chapters, the most current techniques and problems of psychophysiological measurement will be considered, along with the current status and future possibilities of research in the areas of sensation and cognition. Finally, a summary chapter will elaborate on the future potential and limitations of this measurement approach, and provide some concrete proposals for future exploitation, structured around questions of current interest to aircraft designers. These questions involve concepts such as workload, fatigue, attention, and operator reliability, and will be discussed briefly in the last chapter.

Essentially, the space devoted to each kind of measure within the major sections constitutes an evaluation of the amount of work that has been done or is presently feasible using that measure. Thus, since the EEG, EMG, heart measures, and GSR constitute the bulk of the psychophysiological techniques presently used, these will receive the greatest amount of attention. Other techniques which are less frequently utilized at the present time will be noted but not extensively covered. On the other hand, certain techniques which might be considered psychophysiological, and which are extensively used, will be covered very briefly. These include, most notably, biochemical analysis techniques. These procedures are used most often to assess long-term changes in subject state, and while they may be useful in measuring long-term effects of design principles on the human, they are not readily useable in early design stages.

BRIEF HISTORY OF PSYCHOPHYSIOLOGICAL TECHNIQUES

In a very real sense, psychophysiological observation has been carried out informally since animals began to bare their teeth to each other, or humans began to notice that other humans blushed, trembled, or perspired under certain emotions. Early cultures used lie-detector tests based on the observation that one of the sympathetic nervous system responses to stress is a reduction in the rate of salivation. If an accused individual could not chew a mouthful of dry rice or bread, it was assumed that this was a stress response due to lying. Obviously, one could question the validity of this assumption, since it is just as likely an innocent person might be scared spittleless (Hassett, 1978).

In spite of such problems, inferences concerning underlying psychological states from observation of physiological conditions continued to be made throughout the centuries. The struggling medical sciences used such observations to differentiate organic from functional illness (Mesulam and Perry, 1972). Psychological phenomena such as pain perception and phantom limb experiences were related to underlying physiological bases (Melzack, 1973). Drives such as hunger and sex were studied physiologically, and the hormonal or neural substrates of these "mental" states began to be unravelled (Cannon, 1929; Beach, 1958). Perhaps the greatest utilization of these observational techniques in the early twentieth century, however, developed from interest in studying the twin topics of emotion and arousal.

Near the end of the nineteenth century, William James and Carl Lange had proposed that emotions arise only after the occurrence of a bodily reaction to an emotion-producing stimulus (i.e., we feel afraid of the bear because our heart pounds, not vice-versa). Walter Cannon (1927) argued against this position, upholding a more traditional view. The results of the controversy generated by this disagreement are not at issue here, although neither argument is now considered valid in all respects (Grossman, 1967). The important point is that the theories, and the excitement they created, helped set the stage for many investigations into interactions between feelings or behavior and physiological responses (Hassett, 1978).

It was established during such early investigations of these and other questions that many separate indices of central, autonomic, and peripheral nervous system function bore apparently lawful relationships to emotional behavior and performance (Woodworth and Schlosberg, 1954). The most widely used of these was the electrical conductance or resistance of the skin (the Galvanic Skin Response, or GSR). The GSR varies with the diurnal cycle (Wechsler, 1925), the degree of activation required by a task (Davis, 1934; Duffy and Lacey, 1946), the perception of a sensory stimulus in almost any modality (Woodworth and Schlosberg, 1954), adaptation and habituation to a task (Davis, 1930), the apparent emotional tone of certain words (Smith, 1922), and the degree of stress associated with mental work (Sears, 1933).

Another physiological index found to differ with psychological functions was muscle tension. Early, crude mechanical techniques for measuring the level of tonus in specific muscle groups (Davis, 1942) gave way to measurement of the electrical activity of muscles (the electromyogram, or EMG). Such EMG recordings revealed good correlations between levels of muscular tension and certain thought processes (Jacobson, 1951) as well as general level of alertness (Travis and Kennedy, 1947).

The measurement of the electrical activity of the brain itself (the Electroencephalogram, or the EEG) was used infrequently in such studies, in spite of the obvious direct connection between brain and behavior. Technical problems in isolating the small EEG signal and in recording it reliably limited its usefulness to clinical and basic science applications. However, Lindsley (1951) showed that unexpected stimuli affect the raw EEG, causing a rapid desynchronization of the 8 to 12 Hz (alpha) activity, and Darrow (1946) used electroencephalographic evidence to describe the psychophysiological regulation of the brain.

Although many such productive psychophysiological techniques were developed during the first half of the twentieth century, relatively few clear-cut examples of utilization by human engineers can be found prior to 1940. It is true that ingenious instruments for measuring such things as finger tremor (Pullen, 1944) and motor performance (Seashore, 1951) were utilized. However, many of these were designed primarily to measure the physiological response itself, or to be used in a laboratory setting for specific purposes. For example, Darrow measured the sweating response by having subjects press their fingertips against a glass plate, which he then observed with a microscope (Hassett, 1978). Wendt (1930) attached a pen to the knee through levers in order to study the knee-jerk produced by stimulation of the patellar tendon. By and large, however, the human engineer basically was content to utilize well-established psychological, observational, or psychometric techniques to answer questions of interest (Fitts, 1951).

During and following World War II, however, it came to be realized that many questions of interest could not be answered by these traditional measures. In aviation psychology, it was particularly evident that the man-machine system was becoming so complex and demanding so much from the human that new techniques were needed. Applied researchers therefore turned to psychophysiology for the first time, and began utilizing the techniques described below to answer questions affecting system design.

The high level of interest in techniques such as heart rate, blood pressure, respiration, pupillary response, tremor, and eye blinks stimulated much technical progress. By the 1950s, Lindsley (1951) had developed an overall activation theory of emotion descended primarily from Cannon's approach. In this, emotion was seen as a general expression of the level of activation of the individual, expressed on a continuum from coma to extreme activity. This view was taken by many researchers to mean that any psychophysiological variable was interchangeable with another (Hassett, 1978) and produced many attempts to measure psychological phenomena with inappropriate physiological measures. Since intercorrelations between such measures is not high, this is not likely to be a productive approach.

From such a simplistic view, researchers have now moved to the position that each individual measure of physiological response is a piece in the overall behavioral puzzle. Each measure yields a specific kind of answer, and it is necessary to ask questions very carefully, and to utilize the best measure for answering them. In the following sections, each of the specific types of measure which has survived the test of time will be considered from the viewpoint of basic techniques, and specific attempts to apply the technique to real-world problems in human engineering will be described. For each measurement technique, a broad historical overview up to the early 1970s will be given in order to set the stage for subsequent discussion.

of the current state-of-the-art in these areas.

SURVEY OF SPECIFIC PSYCHOPHYSIOLOGICAL TECHNIQUES

THE ELECTROMYOGRAM (EMG)

Basic Methodology. A muscle essentially consists of a variable number of motor units which, in turn, are made up of a number of muscle fibers. Each motor unit is innervated by a single motor neuron whose cell body is located either in the spinal cord or in the brainstem. When a muscle is to contract, a large number of motoneurons must fire, contracting many motor units. However, the neurons will not normally fire simultaneously, or in synchronized volleys, since this would cause the muscle to tremor or twitch (as sometimes happens in spinal diseases). For a smooth muscle contraction, the motoneurons must fire in asynchronous volleys, producing contraction of different motor units over a period of time.

Electrodes placed on the skin surface over a muscle will measure (in a fairly complex way dependent on electrode spacing and muscle depth) the firing of motor units involved in contraction (Davis, 1959). The amplitude of the electrical activity of the EMG is linearly related to the force exerted by a muscle, at least within certain nonstrenuous limits (Goldstein, 1972). The frequency of the EMG similarly bears a direct, though fairly complex, relationship to the force exerted (O'Donnell, Rapp, Berkhout, and Adey, 1973). Typically, amplitudes may range from a few microvolts up to many hundred microvolts. Frequencies for most larger muscles seem to peak between 45 and 60 Hz, with significant power as low as 14 Hz and as high as 100 Hz.

Technically, the EMG is not a difficult signal to obtain. For precise study of single muscles, or even single motor units, needle electrodes can be inserted directly into the muscle. However, this is seldom necessary in most applied contexts. For these purposes, standard silver or silver/silver chloride electrodes can be attached to the skin directly over the muscles of interest. Inter-electrode distances, optimal placements on the muscle, and allowable resistances will differ depending on the muscle used and the purpose of study (Davis, 1959). It is necessary to amplify the signal only 5 to 10 thousand times for most large muscles, although for very small muscles, or to pick up slight changes in activity, amplification of 50 thousand or more can be used.

Analysis Techniques. Analysis of the EMG signal is not nearly as easy as obtaining the signal in the first place. The frequency of the signal, and its relatively high amplitude, produce a record which is visually complex and for many years proved essentially impossible to classify except in very gross characteristics. Jacobson (1951) progressed from measuring and counting the printed EMG waves, to rectifying the signal and then integrating the total power under the curve. This voltage was then fed into a circuit and the buildup of integrated voltage was printed out until it reached a maximum limit. At that point, the integrator and pen reset and began to accumulate again (Figure 3). This procedure became the standard research technique and has been used in many studies.

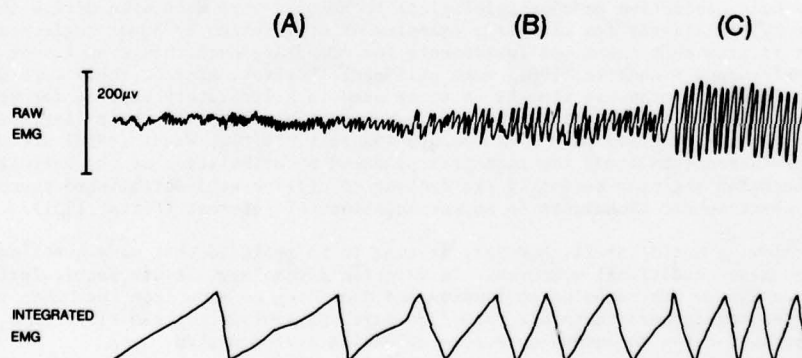


Figure 3. Electromyographic (EMG) activity at low (A), medium (B), and high (C) levels of muscular tension, and typical integrated EMG patterns.

If the rectified voltage is passed directly to an audio source, a tone whose frequency is related to the amplitude of the EMG signal can be produced. This yields a rapid, on-line evaluation of the EMG amplitude. Although no more precise than visual inspection, this audio procedure allows the subject to adjust tension and maintain a reasonably constant level due to the feedback provided by the sound. Physicians also have found this technique useful in probing muscles for pathological reactions.

Precise measurement of the EMG, however, depended on the development of techniques for accurate frequency and amplitude analysis. Early attempts at this consisted of analog filters which broke the complex EMG waveform into a finite number of "bands" (Hayes, 1960). The rectified, integrated voltage in each band could then be calculated with some degree of precision. This rather cumbersome and inflexible procedure became unnecessary when the Fast Fourier Transform (FFT) technique permitted rapid calculation of the frequency spectrum for data within the EMG range (Walter, 1963). Using this approach, precise amplitude measurements at every frequency of interest could be made, and, combined with existing computer programs, could be used to calculate autocorrelations, coherences between channels, or powers within any pre-defined bandwidth (for example, UCLA Health Sciences Center Bio-Med Computer Packages). With the advent of these techniques, analysis of the EMG has finally come of age. For the first time, it was possible to consider using this measure in sophisticated, on-line analyses, giving enough latitude to be sensitive to small variations in stimulus conditions.

Early Applications. On what basis would one expect this measure to show significant variation with meaningful psychological factors? Jacobson was the first to note that when a subject's muscles were as completely relaxed as possible, there was a state of reverie -- the mind was a blank. Conversely, with great mental work, there appeared to be an increase in generalized muscular tone (Jacobson, 1938). These observations are so reliable, and extend down to such specific levels of thought/muscular interaction, that some authors attempted to locate thought almost literally in the muscles (Max, 1937; Watson, 1919). Although this conclusion can in no way be justified from the data, the above studies do reveal an intimate connection between brain and muscle, and it is not unreasonable to look for subtle and covert psychological changes in the degree of muscle tension.

Early studies applying the concept of muscle tension to psychological factors used simple strain gauges, dynamometers, or mechanical ergographs to assess level of muscle activity. Bills (1927) attempted to have subjects produce muscle tension by compressing a dynamometer while performing mental activities such as arithmetic or reading. Higher muscle tension was associated with greater mental output. However, these results have not been uniformly replicated (Block, 1936). Although there appeared to be an "optimal" level of muscle tension for specific tasks, the amount of variability between subjects and even within the same subject over time precluded effective utilization of the EMG as an index of psychological factors such as "effort", "motivation", "stress", or "workload" in most applied contexts.

Of much greater value in a practical sense was the use of ergonomic measures as indices of the work involved in doing specific manual tasks (Davis, 1932; Freeman, 1948). In systems design, it is crucial to determine whether the human operator is able to manipulate all controls under all expected conditions, and it is a straightforward task to establish force envelopes for the specific system of interest. These envelopes can then be combined with reach envelopes, anthropometric data, and system task analyses to provide extensive design criteria for the engineer. Perhaps no single physiological measurement technique contributed as much to overall system, especially aircraft, design up to the present era. This was especially true in Europe, where data produced in many laboratories eventually found their way into standard design handbooks (Van Cott and Kinkade, 1972).

Thus, while the study of muscle activity was producing valuable data concerning purely physical design criteria, efforts to demonstrate the utility of the EMG as a measure of psychological function were faltering. In general, then, prior to 1960 the use of muscle potential measures *per se* in aviation research was restricted to a few attempts to measure effects of imposed stresses of flight on the person, assessment of some muscular concomitants of psychological stresses, and promising but limited attempts to measure alertness, effort, and other psychological phenomena difficult to measure behaviorally. It remained for technical developments in EMG analysis, and the discovery of the phenomenon of biofeedback to create a resurgence of interest in this technique during the late 1960s and into the 1970s.

THE GALVANIC SKIN RESPONSE (GSR)

Basic Methodology. There are over two million sweat glands on the human body, with great concentrations on the palms of the hands and soles of the feet. The majority of all sweat glands are called eccrine glands, and produce a sodium chloride solution which is important in thermoregulation and is not principally responsible for body odor. Eccrine glands on most of the body are controlled by the hypothalamus and respond primarily to heat stimuli. However, some eccrine glands, located principally on the palms, soles, forehead, and underarms, do not respond primarily to heat, but rather to perception of stimuli, especially stressful stimuli.

Féré was the first to report that when a weak electrical current was delivered to the arm, the resistance of the skin changed during perception of emotional stimuli. It was eventually established that these resistance changes were due to alterations in the sweating response of the eccrine glands, and the "Galvanic Skin Reflex" (GSR) was born. Tarchanoff later reported that similar changes in resting resistance of the skin could be elicited without imposing the external electrical current on the subject. These two techniques, the Féré and Tarchanoff methods, survive to the present as the dominant ways of obtaining a GSR.

Terminology in this area is quite confusing (see Edelberg, 1972). The original term "Galvanic Skin Reflex" soon came to be seen as a misnomer. The phenomenon was not a reflex in the true sense, but was rather a *response*. The term GSR therefore came to represent Galvanic Skin Response or Galvanic Skin Resistance. Some psychophysiologicals object to either of these terms on technical grounds. They point out that a more descriptive term would be "Electrodermal Response" (EDR) which would cover both the Féré and Tarchanoff techniques.

There are meaningful differences between the two methods. The Féré exogenous procedure produces a

measure which precisely should be called skin resistance, since this is what the galvanometer reading means. The Tarchinoff endogenous measure should be called skin potential, since it is a "passive" difference between two electrodes. To further confuse the issue, the resistance measured by the Fere technique has certain biological characteristics which make it difficult to interpret and unwieldy to use statistically. If, instead, its reciprocal is used (conductance, given by OHMS = 1/MHOS) (Woodworth and Schlosberg, 1954) the measure becomes statistically acceptable. Further, it has been shown (Darrow, 1964) that the skin conductance is directly related to the number of active sweat glands. Thus, in current practice, it is customary to report the measure obtained by using the Fere exogenous technique in terms of skin conductance, expressed in mhos. Since most literature still refers to the overall phenomenon as GSR instead of EDR, this term will be retained here.

The GSR is, like the EMG, not a technically difficult signal to obtain. Typically, a very weak current is delivered to the subject through one electrode, and the other electrode is used to pick up the transformed signal. Obviously, many factors such as electrode size and distance between electrodes, will determine the absolute value of the conductance measured, although relative changes will still be meaningful. Typically, the fingers or palms are used as electrode sites, although the soles of the feet, or less desirably the forehead or underarm may be used. The latency of the response is highly variable, but is slow by the standards of most electrophysiological measures (1.2 to 4 sec, with a typical palm response latency of 1.8 sec) (Edelberg, 1972). Simple commercial devices for this purpose now make obtaining the GSR extremely easy and non-technical. General discussions of techniques and technical problems are given in Edelberg (1967; 1972).

Analysis and Interpretation. Interpretation of the resulting signal is quite a different story. There are several different types of measures one can obtain, and different ways to classify each one. Since one is always looking at a change from baseline in a given subject, it is important to know whether that baseline has shifted during the experiment (it usually does) and even more importantly, whether a given conductance change from one baseline is the same as from another. Use of conductance as a standard measure has tended to reduce the magnitude of this problem because of the linear relationship to the number of active sweat glands. Thus, most investigators are content to report absolute magnitude of changes in conductance, ignoring baselines except in extreme cases.

Two general classes of GSR measures are typically distinguished. Tonic measures are those that occur over a relatively long period of time (e.g., several seconds or minutes), and do not consider magnitudes of events which occur briefly. The most basic tonic measure is simply the average resistance or conductance level over a long period. This may be sampled as fast as once per second or as slow as once every 15 minutes or more. This is enough to establish an overall level of sweat gland activity. Phasic GSR measures are those which occur rapidly (in less than a few seconds) and which may be of two types. Those which can be directly related to the occurrence of a stimulus (elicited) or those responses which occur without an obvious external eliciting stimulus (spontaneous). These spontaneous responses are sometimes counted over a long period of time (e.g., while a subject watches a disturbing movie) to provide a tonic measure of sweat gland reactivity. They are seldom measured individually, being used ordinarily to classify states or subjects. Elicited responses, on the other hand, are usually timed, measured, and analyzed to provide specific information on the stimulus-response relationship. One of the major current trends in GSR research is toward understanding the different origins for different kinds of GSR response measures, and this may eventually produce a much richer and more useful range of metrics. For the present, it is sufficient to recognize that two broad principles concerning GSR are now generally accepted: (1) sweat gland activity is an index of events in the brain; and (2) the amount of sweat gland response is lawfully related to the intensity of conscious experience (quoted from Hassett, 1978).

Early Applications. As pointed out earlier, it was recognized very early that GSR varied systematically with many conditions causing activation of the individual. These included diurnal shifts, habituation to a novel task, and emotion-laden words. It was not always possible to obtain specific agreement between a given GSR phasic response and the reported magnitude of an emotion, or the pleasantness-unpleasantness of a situation, but group averages generally produced ordinal classification with good agreement (Woodworth and Schlosberg, 1954). McGinnies (1949) used GSR measures to demonstrate that subjects shown "dirty" words slightly below the perceptual threshold actually had an emotional response to the supposedly non-seen word. Lazarus and McCleary (1951) confirmed this finding in a well-controlled study, and established that GSR was able to index emotionally toned perceptions which were not available to the individual's conscious report.

Such studies and findings might have led to widespread use of the GSR in applied settings to measure such things as emotional level during real world operations, stress, and workload. In fact, relatively few studies were done which had any lasting effect on applied fields. In advertising, there was a brief flurry of GSR activity after it was shown that a product to which housewives gave the greatest GSR response also turned out to be a best seller (Eckstrand and Gilliland, 1948). However, this failed to hold up in other tests, and the GSR eventually fell into only occasional use.

This is not to say that the GSR technique did not play a significant role in theory development during this time. Lindsly's (1951) activation theory produced considerable controversy. Discussion arose concerning the implications of such a general statement of the organisms activation on performance in many areas. Malmö (1959), using evidence from many areas of psychophysiology, synthesized much of his work in defense of a modified form of Lindsly's theory. During this time, many investigators were attempting to assess the value of the GSR as an index of overall arousal (Edelberg, 1972). In most cases, a classical inverted "U" relation was demonstrated (Burch and Geiner, 1960), with GSR rising directly with activation up to a point, then falling as activation rose higher. Other investigations found direct relationships (Stennett, 1957) or no relationship (McDonald, Johnson and Hord, 1964). These apparently contradictory findings were probably due in part to different definitions of arousal and different techniques to induce activation. They were mostly due, however, to casual and over simplistic use of the GSR measure itself, with little regard for what it actually measured.

Summary. The GSR has had considerable difficulty being accepted beyond limited basic science and theoretical applications. In retrospect, it appears that much of the confusion generated by the variability of GSR studies has been due to the failure to appreciate just how complex the measure really is. The

realization that different measures of GSR may actually be measuring different phenomena may finally bring order to this confusion (see Hassett, 1978). For instance, Kilpatrick (1972) found that subjects taking an "IQ" test showed increased tonic levels of skin conductance, without rise in the number of spontaneous phasic increases. If an emotional stress was added to the cognitive stress by telling subjects that the same test was an index of "brain damage", both tonic level and spontaneous increases in conductance became higher. Thus, a rise in the number of phasic GSR increases may be specific to emotional stress, whereas tonic level may reflect both emotional and cognitive stress.

At the present time, most investigators appear to feel that considerable basic investigation into the origin and meaning of GSR measures will be necessary before they will become generally useful in applied situations. While many are optimistic about the ultimate outcome of such investigations, the use of the GSR is, at the moment, on the decline in most operationally oriented laboratories.

MEASURES OF CARDIOVASCULAR FUNCTION

Basic Methodology. Ever since the heart was seen as the seat of the soul, investigators have looked to measure cardiovascular function to reveal hidden emotional or affective cues. Lambrosso was a pioneer in using blood pressure measurements in the interrogation of suspected criminals, and from this impressive beginning, measures of cardiovascular function have generally proven to be the most useful and most popular technique for assessing overall emotional tone. Einthoven discovered, in 1903, that if electrodes were placed on many parts of the body, a consistent potential difference could be recorded which corresponded to the beat of the heart. This electrocardiogram (EKG - a derivative of the original German term, which is still used more often than the English ECG) has been traced to the specific electrical pattern of the complex heartbeat, and forms the basis of many psychophysiological measures presently used. Its major components are shown in Figure 4.

The standard P, Q, R, S, and T points designate changes in potential direction related to specific cardiac events. The QRS complex, easily detected from almost any point on the body, separates the two major segments of the waveform, and can be used to count "pulse" beats and determine heart rate (although the two are not strictly identical). The S-T segment is the systole (highest blood pressure) and the T-P segment is the diastole (lowest pressure).

Blood pressure itself is a subject of great interest to the psychophysiologicalist. At the systolic point, the heart is pumping with enough force to overcome the resistance of the peripheral arterial circulatory system, causing the blood to be moved along. In between beats, when the heart is "resting", a force is still maintained which is sufficient to keep the pumped blood moving. This diastole is the minimum pressure maintained by the heart, and is usually represented by the second figure in a typical blood pressure reading: i.e., systolic/diastolic.

Blood pressure is classically measured by inflating an air cuff (sphygmomanometer) around a limb with enough force to occlude blood flow entirely. As the cuff is slowly deflated, the first clear sound is heard at a certain pressure reading, as the cuff pressure becomes low enough for the heart to pump past it. This

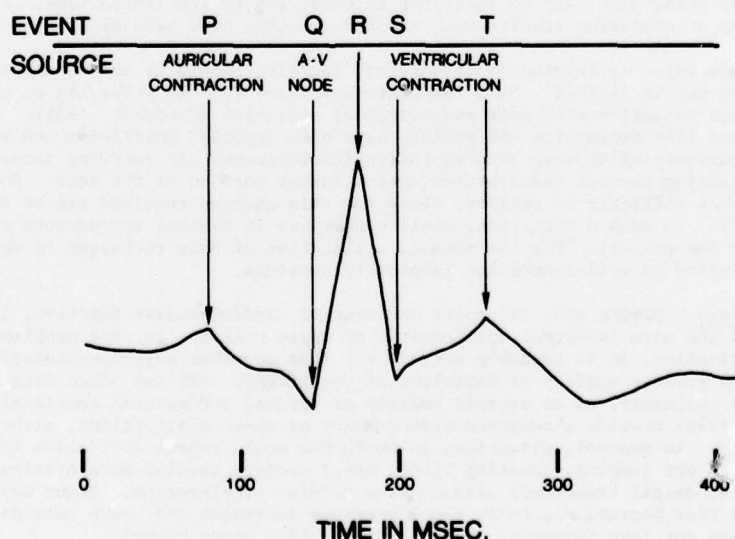


Figure 4. Principal peaks in a typical electrocardiogram, and sources of activity.

"Korotkoff" sound is taken as the systolic blood pressure. With further deflation of the cuff, the Korotkoff sounds disappear altogether as the cuff pressure becomes low enough for the diastole to pump past it. The pressure at which the sounds disappear is taken as the diastolic blood pressure. Both of these underestimate the true systemic pressure slightly, but are consistent enough for most applications. Such readings are relatively easy to obtain, and can be automated. However, they do not provide a continuous readout, and it is not a trivial matter to obtain multiple blood pressures from a given subject. The procedure is not entirely comfortable, and repeated inflation of the cuff can cause petychial bleeding under the skin and result in bruises or worse.

One attempt to obtain a more continuous blood pressure reading involves inflating the cuff above the diastolic but below the systolic level. As the heart beats, the blood volume in the arm changes, exerting pressure on the inflated cuff. This pressure can be continuously monitored. However, this measure is not really blood pressure, but blood volume, and the relationship between the two is certainly not direct. Although this system is used in many lie-detectors today, it really is not very precise and cannot be justified where accurate blood pressure readings are required.

A somewhat more precise, if not safer, technique has been proposed for biofeedback research by Tursky (see Tursky, Shapiro, and Schwartz, 1972; Tursky, 1974). A microphone on the arm is used to pick up the Korotkoff sounds, and a cuff is inflated to a point where it is just below the systolic pressure. For each beat (detected by the EKG) an electronic circuit determines whether or not a Korotkoff sound occurred. The system can then be adjusted to further inflate or deflate the cuff in order to maintain a certain percentage of beats on which the sound will occur. This technique provides a reasonably accurate and quick-response method for tracking pressure, although it does not avoid the safety problems of continued cuff inflation, and the beat-to-beat pressure readings are not entirely accurate (i.e., there may be some lag in finding the correct pressure).

A different kind of circulatory system measure is blood volume itself. This measure is frequently expressed as blood flow, and can be measured by a plethysmograph. For fairly crude measurements, or where pulse rate is simply to be counted, a finger plethysmograph can be used. This is a device which fits over the end of one finger, sealing off the finger and creating a closed volume. As the pulse of blood enters this finger, the volume inside the device changes, and this can be detected as a pressure change. Such devices continue to be used in many psychophysiological applications.

Somewhat more accurate readings of blood volume can be obtained from a photoplethysmograph. A light is passed through a thin body part (such as the finger or earlobe) and detected on the other side by a photo-sensitive source. Changes in blood volume will cause changes in the amount of transmitted light, and these can be measured and calibrated. This technique, though somewhat cumbersome, is accurate and direct. It is also possible, by including color detection in the photo-sensitive circuit, to estimate the oxygen level of the blood from moment to moment, and this provides a crude but in some cases more convenient way to measure oxygen consumption than traditional techniques. More sophisticated techniques for measuring blood flow, including ultrasonic Doppler techniques with telemetry, are discussed by Gunn, et al (1972).

The measures of cardiac and circulatory function which have been discussed so far actually assess quite different things in many different cases. Schwartz (1971) has shown that heart rate and blood pressure do not necessarily co-vary. It is important, then, to choose the appropriate circulatory measure for the purpose of the study being done, and to interpret it according to its limitations. Failure to do this has resulted in a number of confusing results when cardiac measures have been applied to real-world environments.

One final measure which is related to circulatory function should be noted, although its use in operational settings may be limited. Skin temperature appears related primarily to arterial blood volume, and shows significant variation with some psychological variables (Plutchik, 1956). Changes in skin temperature in states like depression and anxiety have been reported (Mittleman and Wolff, 1939). In addition new techniques of thermal autoregulation (biofeedback) are renewing interest in use of multiple thermistors for obtaining average temperatures over a larger portion of the body. However, this technique still remains somewhat difficult to utilize, since the skin changes involved are so small (sometimes as low as .01 degree F). In such a situation, small variations in ambient temperature or air flow can cause spurious changes in the measure. For the moment, utilization of this technique in operational contexts will probably be limited to well-controlled laboratory settings.

Early Applications. Having seen the major measures of cardiovascular function, it is now appropriate to consider some of the more important applications of these measures to past problems. Although it is a dangerous generalization, it is probably safe to say that no other psychophysiological measure has been used more often in a greater variety of experimental conditions. EKG has often been used interchangeably with GSR, its major contender, as an overall measure of arousal and general emotional tone. Gunn, Wolf, Block, and Person (1972) provide a comprehensive summary of these applications, although primarily from a clinical view. In general, situations in which one would expect activation (driving, psychodrama, stressful interviews, ski jumping, donating blood, etc.) produce cardiac acceleration. Anticipating stress (examinations, shots, dental treatment) also produces cardiac acceleration. Anger may be able to be differentiated from fear because diastolic blood pressure increases and heart rate decreases seem more common for anger than for fear (Hassett, 1978, quoting a 1953 study by Ax).

These results, in general, are not surprising and would not contribute a great deal to our observations in an applied context. Interest however, has been generated in using cardiac changes to assess more subtle differences between conditions in applied settings. This measure has been especially used in those situations where the stress levels are all high, and there is interest in distinguishing between several high levels of activation. Thus, cardiac measures have been seen as providing a metric with a significantly large "top" or upper end, since the heart rate can go from a baseline of 50 or 60 beats per minute (bpm), up to over 200 bpm. In aviation contexts, where interest is frequently in high stress situations, this could prove very important.

From the viewpoint of actually interpreting the meaning of a change in cardiac function, a most important

theoretical controversy has occurred between the Laceys on one hand, and Obrist on the other (see Lacey and Lacey, 1974; Obrist, 1976; Hassett, 1978). The Laceys have pointed out that the entire physiological response to an arousing stimulus does not always move in one direction. While doing a mental task like arithmetic, for instance, both heart rate and GSR measures might be increased. However, if the task involved attending to a stimulus occurrence, the heart rate might decrease while GSR measures increased. Lacey interpreted this type of result to indicate that "environmental rejection" (i.e., doing something internally not dependent on attending to the outside environment) yielded increases in heart rates. "Environmental intake" (attending to the environment in order to receive information) produced heart rate decreases. This is one example of what the Laceys refer to as "directional fractionation", the specific and often apparently contradictory direction of each index of autonomic and central nervous system function under differing psychological stimuli.

The practical implications for the Lacey's view of heart rate can be seen in their further speculations concerning the connection between cardiac activity and brain function. As the heart beats, pressure rises to a maximum during the systole (S-T and T). As the pressure rises, baroreceptors located in the aortic arch and carotid sinus are activated to provide a homeostatic mechanism which reduces pressure. Some evidence exists that activation of these baroreceptors can cause deactivation of the cortex of the brain. Thus, there should be a cyclic activation-deactivation of the brain with each heartbeat. If this is true, then performance which depends on brain activation should be better during the diastolic portion of the cardiac cycle, since low baroreceptor activity should allow cortical activation. The Laceys have reported several studies confirming this prediction (Lacey and Lacey, 1974) but other studies have failed to find effects of the cardiac cycle on reaction time (Thompson and Botwinick, 1970) or auditory thresholds (Delfini and Compos, 1972). Boyle, Dykman, and Ackerman (1965) attempted to find direct EEG correlates of the cardiac cycle, and were unable to do so.

Obrist completely rejects the Lacey view of heart rate changes, and refers to phasic effects as "biologically trivial". The activity of the heart is considered simply a reflection of tissue needs, and the heart will beat faster or slower depending on how active the individual is. In solving problems, the person tenses muscles. This increases the oxygen need, and the heart beats faster. The apparent "directional fractionation" of the Laceys is dependent on whether the heart is being controlled by the vagus nerve (Parasympathetic nervous system) or by the sympathetic nervous system. The former occurs when the person has little control over the environment ("passive coping"). Under these conditions, the heart and body are both coupled, so that they will co-vary. In conditions where the person is actively affecting the environment ("active coping") sympathetic control may permit heart rate changes which are NOT necessarily correlated with other body changes.

In addition to raising a wealth of critically important basic questions about cardiac function, this controversy has specific importance to any attempt to apply cardiovascular measures in operational environments. The bulk of evidence tends to support Obrist's view that the heart is responsive to tissue needs in many circumstances. Obviously, it is important to consider this in the interpretation of any study. Beyond this, there appears to be ample evidence that the brain controls the heart under most circumstances (Gunn, et al; 1972) and not vice-versa. However, these facts do not necessarily negate the importance of the Laceys' position, and certainly do not in themselves render phasic heart rate changes biologically trivial. Under at least some conditions (Obrist's "active coping" and Lacey's "environmental intake") the heart appears to be uncoupled from somatic requirements, and its principal determinant becomes the sympathetic nervous system. It would seem reasonable to expect that at least in these circumstances, cardiovascular measures (albeit perhaps more sophisticated measures than may be presently available) should provide a differential index of psychological state.

In spite of these interpretation problems, cardiac measures have been widely used in laboratory studies of arousal and stress, and in some applied studies. One of the more impressive applications of this measure involved a study of U. S. Navy Combat pilots on an attack mission from an attack carrier in Viet Nam (Roman, Older and Jones, 1967). These investigators used a portable recording pack which utilized three dry electrodes developed by NASA (Roman, 1966). The signal conditioner and recorder were carried in the aircraft mapcase and were capable of recording 120 minutes of data continuously from up to seven channels during launch, bombing, and recovery, and were averaged to produce single representations of each flight phase. Nine flights were analyzed.

The data revealed a surprising relationship between tasks and apparent stress levels. Figure 5 shows that heart rate was lowest during the bombing operation, and was significantly higher during launch and recovery. During the attack segment, which would have been expected to be the most dangerous, and therefore most stressful portion of the mission, pilots showed a decrease in cardiac activity. Further analysis of the data revealed an overall pattern of higher heart rates during the first mission of the day than during the second. The authors point out several cautions which must be taken into account in interpreting the data, including the high level of experience in these pilots and the relatively routine nature of these particular combat missions. However, they interpret the results as indicating that risk or danger are not major factors in determining heart rate in these conditions.

This study points up one of the difficulties in using crude psychophysiological measures in applied settings. In effect, since so little is known about the psychological meaning of cardiac acceleration, it is extremely difficult to make meaningful interpretations about the applications of the kind of data observed above. Did the increased heart rate really mean that the launch and landing operations were most stressful? Or do those operations simply involve a higher physical workload than attack, or is some combination of stress and workload interacting during launch and landing to produce higher heart rates? Is it possible that stress was highest during the attack phase, but that "concentration", "inhibition", "environmental intake", or some other autonomic focusing served to reduce heart rate? Does the heart respond to increased stress by accelerating up to a certain point, and then with further stress show a reduction?

At the time of this study, few solid answers to these questions could be provided. Lacking such answers, the simple observation that heart rate is maximal during launch and landing has little real, applied value. Perhaps it would have been better if such a study had been done in a laboratory setting, where specific factors could have been better controlled, and where the results would have been more clearly interpretable.

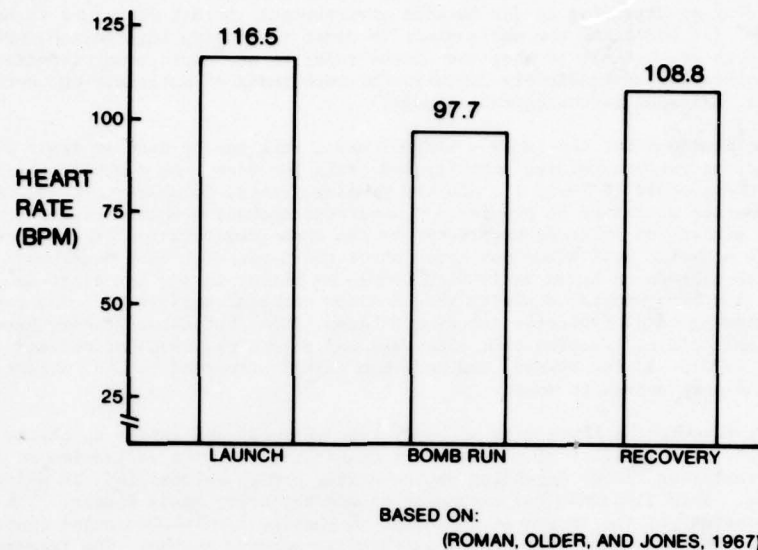


Figure 5. Heart rate results during aircraft carrier operations in Viet Nam.

In general, then, whenever basic information about the meaning and source of a psychophysiological measure is lacking, it is safest and most productive to use that measure in tightly controlled laboratory studies rather than in field situations. As more and more becomes known about the origin of the electrical signal in the body, its behavior under specific environmental conditions, and the kinds of stimuli which cause it to change, it becomes safer and easier to use that measure in real-world environments. In the former case, one gives up generality and face validity to the operational situation in return for precision of interpretation. The trade-off is certainly worth it. As impressive as heart rate changes in operational environments may be, they provide little basis for decisions if they are uninterpretable.

Changes in heart rate have been studied in a large number of other stressful situations such as parachute jumps, diving, sky diving, etc., Miller (1976) reports heart rates up to the 110 to 140 bpm range under acceleration loads from +3 to +5 G_z . At +6 to +9 G_z , or after repeated exposure to +3-5 G_z , heart rates reached 160 to 205 bpm. Other cardiac abnormalities also appeared under these loads (Cohen and Brown, 1969). Heart rates of parachutists have shown some of the highest rates ever recorded, well above the normal rate considered to reflect tachycardia. Again, except for the general indication of severe physical and perhaps emotional stress, these patterns are difficult to interpret.

Increased rates may or may not reflect the actual physiological condition of the subject, but in any case they do not provide a firm basis for a broad range of applications, principally because they do not allow interpretation of differences between individuals. The ability of heart rate to differentiate between states in the same individual may have some value. If this measure can index the training level of students learning a highly activating task (like parachute jumping, flying, diving, etc.) it could be of value. However, more research will be needed to determine whether such an index is more reliable or precise than simple verbal report or performance measures. It is interesting to note that subjects reported verbally that the attack phases of the carrier-based operations in the study by Roman *et al* (1967) were less stressful than launch or recovery. Similarly, good correlation is found between verbal report and heart rate measures in parachute operations. If verbal reports are sufficient in these cases, there really is no need for physiological measures. A number of other cardiac measures have been used in aerospace research to evaluate the physiological effects of stressors. These include such measures as the phonocardiogram (Bergman, Wolthuis, Hoffer, and Johnson, 1973; Tavel, 1968), echocardiograms (Fortuin, Hood, Sherman and Craige, 1971) and various flow techniques for measuring cardiac output and venous pressure (Denison, 1970; Thomas, 1974). However, even more than heart rate itself, these techniques are less reliably interpreted in any psychological sense, and are principally of value in determining when a particular external condition is likely to result in physical damage to the individual.

Summary. For the moment, most investigators appear to feel that, except as a crude and generalized ordinal index of high-level activation within a subject, cardiovascular measures are not likely candidates for large scale application to real world situations. Continued basic research, particularly directed to opening up new measurement techniques (e.g., measuring contractibility) may eventually reveal that the cardiovascular system yields a rich array of measures, each indexing different physical and psychological functions, but these await significant development.

MEASURES OF BRAIN FUNCTION

Basic Methodology. Since Hans Berger demonstrated, in 1929, that alterations in brain electrical activity could be recorded from the skull, the prospect of understanding psychological function directly through the central nervous system has lured researchers into using millions of miles of paper. Unfortunately, except for a few classes of clinical applications, most such efforts were wasted. Early workers never clearly appreciated the fact that the 10 billion-plus cells in the human brain were simply not going to produce a signal that could be easily reduced to a single line on paper, and early enthusiasm faded into almost total disillusion. A height of skepticism was expressed by one researcher who likened electrical recording of the brain to lowering a microphone from a helicopter into a crowded athletic stadium and trying to guess from the crowd's roar not only what game was being played, but what each player was doing.

Yet, the fact remained that such electrical activity (the electroencephalogram, the EEG) did show certain consistencies. Early demonstrations proved that the activity recorded was different when the eyes were open or closed. Brain pathology which had behavioral correlates revealed itself in altered EEG patterns, and individual differences in the EEG were temptingly as numerous as personality differences. Scientists persisted in looking for new ways to tease out the specificities from the overall recording and, to their own surprise, appear to have had a very large measure of success. The EEG and its derivative measures now appear to be the most promising technique available to the psychophysiologicalist, and these techniques have demonstrated considerable value in applied settings.

The EEG is, first of all, a true reflection of brain activity. Attempts have been made to question that significant kinds of EEG activity really are generated in the brain (Kennedy, 1959; Lippold and Novotney, 1970; Bickford, 1964; 1967). However, the data supporting these attempts have been shown to be either artifactual or based on unique and rare conditions. It is, of course, possible to record the brain's electrical activity from microelectrodes deep in particular structures. Similarly, electrodes can be located directly on the surface of the brain (sometimes called the electrocortigram). Since these techniques are obviously not readily available to the applied psychophysiologicalist, they will not be discussed here.

The typical EEG utilizes small electrodes, usually silver, gold, or a silver/silver chloride combination, which are attached to the scalp. The scalp may be rubbed or scraped slightly to remove the cornified layer of dead skin in order to improve the electrical contact. This contact is measured by the resistance between two electrodes, and is one of the critical determinants of the quality of signal which will be obtained. Resistances of 20K to 50K ohms have been commonly used in the past, and are still tolerated in clinical settings. With modern electrodes, however, there is no reason such resistances need ever be used. With minimal preparations, resistance can be reduced to 5K ohms, and in any research setting, no resistances over 2K should be tolerated. This is particularly true if the subject is not to be tested in a shielded room. The common-mode rejection of ambient interference allowed by very low resistance electrodes is one of the major reasons that EEG can now be reliably recorded in operational environments.

Amplification of the EEG signal can range from 10 or 20 thousand up to a million. The signal in the typical "raw" EEG ranges between 10 and 200 microvolts. However, it is now known that, buried within this raw signal, microevents occur which can measure as low as .5 microvolts or less. Typical frequencies range from about .5 hz up to 30 or 35 hz, with most individuals content to filter below 1 hz and above about 25 hz. Again, however, new discoveries have revealed that real activity is present in the EEG at much lower and higher frequencies. For specific purposes, it may be necessary to record DC levels on the one hand, or to pass all activity below 3000 hz. These revised amplitude and frequency standards reflect the radical changes which have occurred in EEG technology in recent years.

Placement of electrodes in EEG work has become fairly standardized. Bipolar recordings are those in which both recording electrodes are placed on presumably active electrical sites on the scalp. The resulting EEG signal therefore reflects the differential activity between the two sites. Monopolar recordings use one electrode on an "active" scalp site and the other on a site which supposedly is electrically inactive, like the earlobe or mastoid bone. In fact, almost no site is really electrically inactive, so even monopolar records reflect a difference between two activities. However, the intent of monopolar recording is to get as close as possible to seeing the "real" activity in a particular location, so that is is probably safe to consider the "inactive" electrode as simply electrically neutral, exerting a non-significant influence on the final record. At the present time, there is no general agreement with respect to whether bipolar or monopolar recording is better. Each appears to have some value in specific cases, and it is necessary to decide in each case the advantages of each technique.

Most commonly, the active electrode is placed over one or more of the four major areas of the brain; the frontal, parietal, occipital or temporal. It may be placed at the midline, or over either hemisphere. Jasper (1958) has presented a standardized measurement system for placing and designating electrodes over the entire scalp, and this terminology is in general use for virtually every EEG application.

Analysis. The procedures and types of measures obtainable from the EEG have mushroomed during the past ten years as new technology and applications appeared. The standard clinical procedure, still used, involves placing many electrodes on the scalp, having the subject recline in a comfortable position, and presenting a number of tasks and stimuli while recording both bipolar and monopolar derivations. Resting, eyes open and closed records are obtained, along with records taken while the subject hyperventilates. Other records are taken while a strobe light is flashed at various frequencies. An attempt may also be made to have the subject sleep or at least enter a drowsy state. Records are made on paper strip charts, and are visually analyzed for the occurrence of unusual frequencies, atypical waveforms, or specific firing patterns associated with known pathology. This visual analysis was essentially all that was available to the researcher until the late 1940s. It is no wonder, then, that the early promise of the EEG failed to materialize up to that time.

With the advent of FM analog tape recorders of very high quality, and of fast computers, new analysis techniques became possible. Burch, Breiner, and Correll (1955) described a technique for counting and classifying waveforms of different frequencies by looking at the zero crossing points and calculating the number of "waves" at each frequency. Autocorrelation techniques also provide an estimate of the frequency components of a complex waveform. Analyses which use higher-order derivatives can isolate the high-frequency

components of a waveform, independent of its zero-crossing. Each of these procedures, as well as others which have been described, has advantages and disadvantages, and are still used in some contexts (Shagass, 1972). However, for the most part, these have been overshadowed by techniques which became available when the Fast Fourier Transform (FFT) allowed rapid calculation of the frequency spectrum of a complex waveform.

The Fourier principle indicates that any complex waveform can be mathematically reduced to a series of simple sine waves of various frequencies and intensities. Thus, any given segment of EEG can be described by a number of sine waves and a power level at each of the sine wave frequencies. It is important to realize that the frequencies revealed by Fourier analysis may not have been visible in the raw EEG records. They are, instead, mathematically accurate components of the visible EEG. If the raw EEG consisted of a pure sine wave at, say 10 Hz, the Fourier analysis would reveal that all the power in the recording was concentrated precisely at 10 Hz. However, if the raw EEG was a complex waveform created by superimposing five separate sine waves of different frequency and amplitude, none of the five individual frequencies might be visible in the record. Fourier analysis would still reveal the five frequencies and their separate power contributions to the final wave pattern. The procedure is mathematically complex, and there are technical features which demand it be used cautiously, but it has become by far the most commonly used procedure in EEG analysis, and has opened up the EEG waveform to a level of sophisticated and detailed interpretation which was unavailable for any other psychophysiological measure (Walter, 1963).

"Evoked" Responses. Another major development which has drastically altered the utility of the EEG as a psychophysiological measure is the use of specific, temporally well-defined stimuli. In the early days of EEG analysis, attempts were made to correlate particular waveforms with a variety of poorly defined psychological states such as "attention", "motivation", "intelligence", "depression", etc. Similarly, in line with other kinds of psychophysiological investigations, such vague terms as "activation" or "arousal" were commonly used in EEG research (see Lindsley, 1951, discussed later). When new analysis techniques became available, it became clearer that the EEG had many components. Not only was it not always a general index of overall brain arousal, but even one recording from one electrode site might be measuring many separate phenomena independently.

This insight generated a reappraisal of the type of stimuli which should be considered appropriate for eliciting the EEG. Overall EEG spectra could still be computed over epochs of time ranging from seconds to hours, but the conditions of the subject during these periods had to be rigidly controlled. The studies of Zubek (1964, 1966; Zubek, Welch and Saunders, 1963) on sensory deprivation illustrate this point. The EEG was still analyzed in epochs, but the environment was systematically and precisely altered to yield correlates of the EEG changes.

To an even greater extent, it soon was realized that to obtain precise control over the stimulus condition, it would be necessary to visualize the brain's response to a single, discrete stimulus. In 1947, Dawson used a procedure in which the peripheral nerve was stimulated, and the EEG was recorded. Since the point at which electrical stimulation was begun could be precisely known, it was possible to take a photograph of the EEG waveform which began with the stimulus and continued for a very brief time. Presumably, this segment of EEG contained the response to the nerve stimulation, plus a great deal of random EEG noise. If, now, this same procedure was repeated many times, and if the photographs were superimposed, the constant signal due to the nerve stimulation should produce a very thick line, while the "noise" from the EEG should essentially be invisible due to its random nature. This is in part what happened (Dawson, 1947) and provided one of the first demonstrations of what has come to be called the evoked potential (EP) or evoked response (ER).

It was soon demonstrated that the EP could be elicited by repeated discrete stimulation in any modality (Shagass, 1972; Perry and Childers, 1969). Further, the crude photographic process soon gave way to mathematical averaging techniques which were incorporated into compact special purpose computers (e.g. the Computer of Average Transients, CAT). It became possible to calculate the EP on-line, and to perform complex amplitude, latency and frequency analyses with great precision.

As more EP work was done, it was clear that even the "pure" response to a discrete stimulus was in itself quite complex and multi-faceted (Figure 6). Many lines of evidence (summarized by Beck, 1975) indicate that the relatively early components of the EP (prior to about 250 msec) reflect the "qualitative" aspects of the stimulus, such as color, sharpness, intensity, or pattern. The later components appear sensitive to psychological state or "meaningfulness".

As early enthusiasm for the EP grew, many investigators failed to appreciate the sensitivity of the measure itself, and frequently assumed that most stimuli were equivalent in producing an EP, and that the EP to different stimuli were essentially equivalent. The strobe light, since it produced a strong EP and was precisely timed, came to be used in the majority of visual EP studies. It was only later, with the work of Regan (1972) that it was realized that the intense "on-off" stimulus of the strobe stimulated many different visual receptors, confounding the EP's morphology unnecessarily. By using a counterphase-flickering stimulus, or a simple "on" stimulus, a simpler, more precise EP could be generated.

The net result of the voluminous work done on evoked responses in the last twenty years has been to produce one of the richest and most precise psychophysiological techniques presently available to the applied researcher. While other techniques have become mired in theoretical controversy, this one has weathered the few theoretical challenges to its validity (See Bickford, 1964; 1967; Bickford, Jacobson, and Cody, 1964; Vaughan, 1969) with relative ease (Beck, 1975). New techniques have been developed in the past five years which promise to drastically enhance its practical utility as a performance measure. These will be described under appropriate headings in later sections of this AGARDograph. Current applications, and those becoming possible, will similarly be described in those sections.

Contingent Negative Variation. A final modification of the EEG may also prove to have considerable practical value. Walter, et al (1964) discovered that if the subject is given a warning signal (S_1) that another stimulus requiring action (the imperative stimulus - S_2) is about to occur, there will be a slow negative shift in the entire EEG. Thus, as the subject is allowed to "expect" an S_2 by being warned it is coming, the EEG waveform, while maintaining its normal complexity, shows a negative DC shift. This

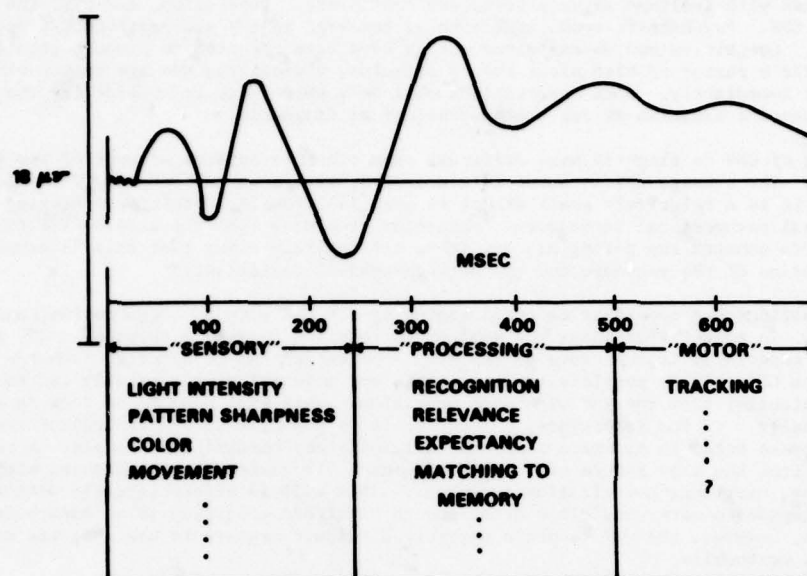


Figure 6. Idealized visual evoked response showing major components and sections.

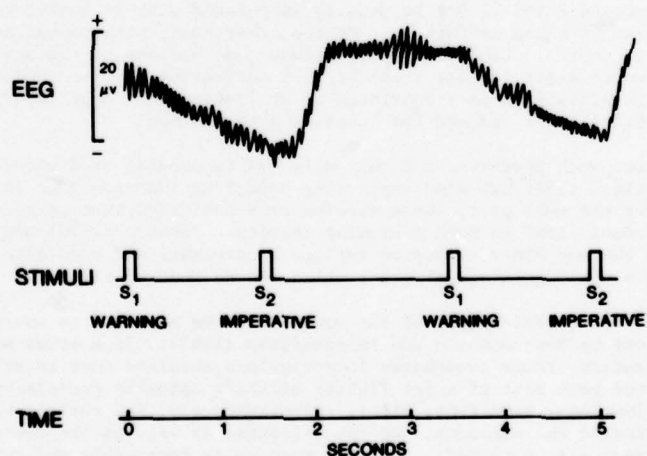


Figure 7. Schematic representation of the contingent negative variation paradigm.

phenomenon has been called the Contingent Negative Variation (CNV) and has been studied under a wide variety of circumstances. The anatomical basis of the CNV is not perfectly clear. It can be found over both hemispheres, and there is some evidence that it may be greater over the hemisphere being used for a hand movement, as well as in the dominant hemisphere during numerical and verbal manipulations (Butler and Glass, 1971). Large CNVs are also associated with quicker reaction times.

Cohen (1974) provides a good summary of the attempts to find correlates of the CNV. Basically, CNV appears to increase with feelings of well-being and confidence. Depression, anxiety, and distraction tend to produce lower CNV. Psychiatric conditions such as cerebral injury and psychopathic deviancy tend to produce low CNVs. Compulsive and obsessive neurotics have been reported to produce good CNVs, but to remain negative for a period of time after the S₂ stimulus, whereas the average person will regain positivity almost immediately. Such observations obviously have value in considering the use of this measure for personnel evaluation or for neuropsychological diagnosis.

The recording of CNV is slightly more difficult than other techniques because of its slow nature (Hillyard, 1974). The average CNV is about 20 microvolts, with a variability of 10-50 microvolts (Hassett, 1978). As such, it is a relatively small signal in most individuals, requiring averaging over 10 to 20 stimuli, or special mathematical techniques. Monopolar recording from the midline Vertex (C_z) position has been recommended to enhance the potential, but it is not entirely clear that this is necessary in view of the wide distribution of the response and the between-subject variability.

Special precautions are necessary to avoid confusing the CNV with a corneo-retinal potential generated by eye movements. In most CNV designs, the subject is required to make a response. If, in doing so, the eyes move -- especially if they move downward -- a potential can be detected from the scalp which coincides with the CNV. It is possible to detect this eye movement electronically and to subtract the corneo-retinal potential from the CNV with some precision. This will have to be done in any attempt to use the CNV operationally. In the laboratory, of course, it is possible to simply monitor eye movements, arrange the stimulus/response setup to minimize them, and eliminate any contaminated trials. A related procedural precaution stems from the slow nature of the CNV response. If resistance is measured with an ohmmeter, as is frequently done, electrode polarization can occur. This will interfere severely with recording of the CNV. Use of an impedance meter and close attention to electrode condition is necessary in CNV work. With these precautions, however, the CNV is not a terribly difficult measure to use, and its utility can make the extra effort well worthwhile.

Hassett (1978) discusses the possibility raised by McAdam (1974) that the CNV may be a special case of the "readiness potential" which precedes a voluntary act. While the theoretical discussion is not of critical importance in the present context, several observations concerning this readiness potential suggest applications which have only been partially explored. A negative shift resembling the CNV occurs about 1.5 seconds prior to the beginning of a voluntary motor action. The motor action may be a limb movement or speech. The potential is greatest over the cortical area most related to the action in question. Although based on relatively few experimental demonstrations, and requiring some design sophistication to obtain, it would seem that the possibilities offered by this phenomenon for monitoring and predicting movements in the human should be explored for possible applications.

Early Applications. Utilization of the EEG in operational contexts has been such a recent phenomenon that most such descriptions appropriately will be covered later. However, a few early attempts at utilization did occur, and can be summarized here.

The activation theory proposed by Lindsley (1951) was essentially based on the observation that activated EEG states correlated with activated behavioral states and vice versa. Alpha (the production of highly synchronized activity between 8 and 12 Hz) is usually associated with an awake, relaxed state. Lower frequencies usually appear in sleep or fatigue. On the other hand, with mental activity, the brain typically shows activity over 13 Hz (beta). Lindsley related these observations to the activity of the reticular formation, which is known to exert overall regulation of activation level. Therefore, Lindsley proposed that all emotion could be classified on a continuum of activation and, most importantly from the present viewpoint, that EEG activity level indexed the level of such emotion.

This theory stimulated much research, not only directed to proving or disproving the theory itself (Duffy, 1972; Lacey, 1958; Lindsley, 1956) but also implicitly accepting the view that an "activated" EEG meant an activated person. For the most part, these studies only confirmed that using overall EEG measures to assess cortical arousal could lead to some confusing results. Johnson (1970) has documented the lack of correlation between the EEG and other autonomic indices of arousal, and concludes with the pessimistic view that there is no single psychophysiological state which can be indexed simply.

One of the earliest, and certainly one of the most ambitious attempts to use EEG in actual real-world situations was carried out by Sem-Jacobsen and Sem-Jacobsen (1963). In a study which was years ahead of its time in conceptualization, these remarkable investigators obtained four to seven channels of EEG from nearly 200 subjects in the back seat of a jet fighter aircraft actually performing aerobatic maneuvers (e.g., G turns, rolls, Immelman, bomb runs, etc.). Simultaneously, the subject was photographed with 8mm movie cameras as he performed the maneuver, and the G-loading as well as the control surface positions were recorded. EKG records were also obtained. The data were quite remarkable and consistent (Figure 8). In 18 of 55 experienced fighter pilots, both behavioral and EEG indices revealed severe pathological responses during the stressful maneuvers. These individuals showed flattening of the EEG followed by high voltage slow and very slow activity characteristic of pathology. Behaviorally, they showed smacking movements, eye turning and rolling, loss of muscular tone, and convulsive jerking. Further, all of these individuals had "pilot error" accidents on their records. When these same subjects were tested on the centrifuge at the same or higher G-loads, the symptoms did not appear (Sem-Jacobsen, 1961). The authors suggest that an in-flight stress test of this type might increase flight safety by decreasing the number of "pilot error" accidents.

Considering the importance of these results, it is surprising that no direct replication has ever been carried out. The nearest thing to a replication was done at Brooks Air Force Base, Texas, (Berkhout,

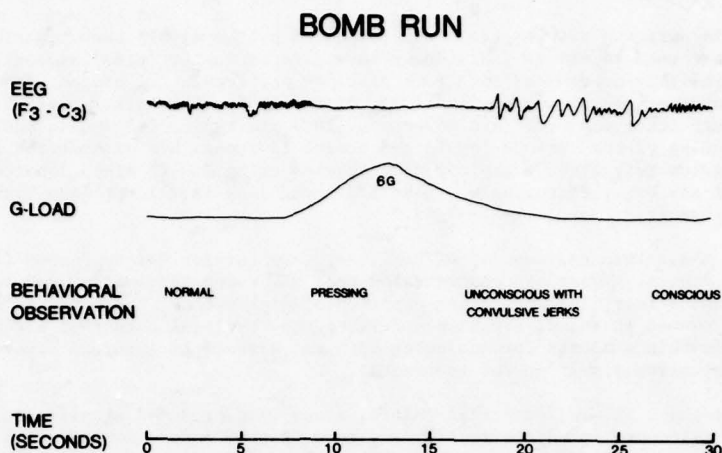


Figure 8. EEG and state of consciousness of a pilot during high-G bomb run. (Based on Sem-Jacobsen and Sem-Jacobsen, 1963).

O'Donnell and Leverett, 1973; see also Squires, Jensen, Sipple, and Gordon, 1964). Using the human centrifuge, subjects were exposed to as much as $+6 G_z$ for 45 second periods while EEG records were taken from bilateral parietal-to-occipital bipolar derivations. The subjects in this study wore a lower body G-suit and performed a straining maneuver (M-1) designed to counter the effects of the G-load. Therefore, they showed no signs of blackout, although some reported loss of the peripheral visual field. The muscle tension accompanying the M-1 maneuver caused considerable interference in the EEG signal. However, spectral analysis was able to isolate the muscle component from that generated in the brain, at least below 14 Hz, (O'Donnell, Berkhout, and Adey, 1974). The results showed increases in lower frequency theta activity (3 to 7 Hz) postrun for most of the subjects. However, the overall effects of this experience were considerably less than what one might have expected and certainly more benign than those reported in flight by Sem-Jacobsen. Overall shape of spectral profiles was not changed, and small increases seen in spectral intensity did not exceed normal levels. In all subjects, a return to normal EEG intensity levels occurred within 30 seconds after acceleration was terminated. Data from this study supplied part of the evidence which encouraged researchers to permit much higher G exposures, and the general lack of harmful cortical effects tends to confirm the validity of these original findings (Miller, 1976).

Another spectacular use of EEG was carried out by McElligut at the Space Biology Laboratory, Brain Research Institute, University of California at Los Angeles (unpublished data). Using a small, battery-powered amplifier/recorder pack which had been developed by NASA for space applications, this researcher recorded EEG continuously during parachute jumps in both experienced and inexperienced jumpers. Cardiac measures were also taken during the entire pre-flight and jump, and revealed the characteristic tachycardia noted earlier. The EEG showed generalized activation which was not observably different between experienced and inexperienced jumpers. Activated EEG patterns did not diminish until long after the jump had been completed, and in general failed to differentiate between stages of the jump. At the point of parachute opening, the jolt caused so much electromechanical interference that all records were lost for a number of seconds. One jumper experienced a fall upon landing, hit his head and was knocked unconscious. The EEG record accurately reflected this loss of consciousness and subsequent recovery. Overall, however, the EEG portions of this study supplied little in the way of new or practical information, and only validated phenomena which were available in other contexts.

It is striking that many of the attempts to apply EEG in practical environments were carried out where the subject was extremely activated. In another case of EEG recording in a high activation, real-world environment, Berkhout (1973) studied California Highway Patrolmen driving around a high speed training track at speeds in excess of 100 mph. In addition to EEG measures, heart rate, eye movements, and vehicle performance measures were taken. This study differed from others in that the data was telemetered from miniature transmitters attached to the subject's helmet to an analog tape recorder located in the back of the car. The subject therefore was not attached to his vehicle in any way and experienced little or no interference with his normal movements. The basic attempt in the study was to find differences between groups of drivers who had been involved in accidents and those who had been accident free. Few simple measures provided such discrimination. Some heart rate acceleration patterns during complex maneuvering appeared to separate the accident group from those with no accidents, as did combined analysis of performance data. Eye movements revealed some differences in the way accident drivers behaved during turns, but these patterns were

not consistent enough between subjects to permit good differentiation between groups.

The EEG in this experiment was particularly disappointing. No consistent patterns of changes were seen in the spectra of either accident or no-accident drivers, and no high correlations with other physiological measures appeared. Complex coherence (cross-correlation) measures between EEG derivations also failed to show interactions between brain sites related to the task, driving style, or ability. Overall, except for a few tentative speculations, no productive results came out of the use of EEG in this study.

These somewhat discouraging results are counterbalanced by the significant contributions made by EEG measures when they were used to assess individuals in states of low or normal arousal. The demonstration that sleep could be reliably subdivided into four distinct and meaningful stages, and that a fifth stage (REM, or rapid eye movement sleep) was qualitatively different from the other four, was primarily an EEG contribution. Although other measures (eye movements, EMG) are typically used in addition to EEG to monitor sleep, real understanding of the various levels and stages is impossible without the EEG (Dement and Kleitman, 1957). These discoveries triggered an explosion of studies using EEG in sleep laboratories around the world, and perhaps more than any other factor helped make EEG technology familiar to many researchers who otherwise would have failed to use it.

Basic studies on the origin and meaning of REM sleep were carried out by Jouvet (1967) in Paris, and Dement (1960), among others. These studies revealed that REM sleep deprivation was accompanied by significant alterations in mood and affect, and that after a period of deprivation, there was a "rebound" tendency in which the individual showed increased REM time. Further, the cyclical nature of sleep stages provided stable patterns between and within subjects for analyzing both the effects of external stressors on sleep, and the effects of sleep deprivation itself on the individual.

In the latter category, it was found that initial sleep loss produced significant decrements in arithmetic speed and accuracy, vigilance, immediate recall, and mood (Lubin, Moses, Johnson, and Naitoh, 1974). With longer-term deprivation of 220 hours (Luby, Frohman, Grisell, Lenzo, and Gottlieb, 1960) the subject showed extreme behavioral changes including paranoid thinking, visual hallucinations, and episodic rage. Physiological indices revealed an extreme stress response active by the fourth day of sleep loss. Even this adaptation began to fail by the seventh day. By the ninth day, the subject was virtually untestable. Recovery from sleep deprivation is usually fairly rapid. Lubin *et al* (1974) report that even when subjects are selectively deprived of either REM or stage 4 sleep during recovery, recuperation rate is about the same as those not disturbed during recovery.

The other major class of sleep studies uses sleep measures to index various types of stress, particularly the stress imposed by noise and toxic exposures. O'Donnell, Chikos, and Theodore (1971) exposed humans to carbon monoxide levels up to 150 ppm, for 9 hours, producing blood carboxy-hemoglobin levels as high as 12.9 percent. All night sleep recordings revealed a slight increase in the amount of deep sleep, but no change in REM time or sequencing. There were no other behavioral changes associated with this sleep alteration, and the authors concluded that no high-level cortical functions were affected by the CO exposure. Although this is a limited use of sleep measures, to define the effects of toxic stress, this study also pointed up the fact that inferences concerning subtle aspects of brain function are available to the researcher through such metrics.

Closely akin to the area of sleep research is that dealing with sensory deprivation. When Heron, at McGill University in Montreal reported that hallucinations, dissociations, and performance changes occurred with cessation of sensory input, an avalanche of research using this technique was unleashed (Heron, 1957; Schultz, 1965; Zubek, 1966). Although the early spectacular results showing bizarre perceptual changes after deprivation were tempered somewhat by later work, a strong trend emerged regarding EEG patterns. Many studies revealed an overall slowing in the EEG with continued deprivation (Zubek, 1964; Zubek, Welch, and Sanders, 1963). This slowing was reported to persist for some time after deprivation was terminated, and to correlate well with perceptual and performance changes. The slowing manifested itself both as a reduction in the amount of alpha (in favor of increased theta) and in a lowering of the average frequency within the alpha band (O'Donnell, 1970). These changes appeared quickly reversible by moderate activity if the length of deprivation was not too great, but tended to persist longer after prolonged deprivation. In addition, recurring periods of deprivation appeared to be cumulative in their EEG effects.

The space program supplied significant momentum to the study of the EEG. Culmination of the space application of this technology came with the measurement of brain signals from Astronaut F. Borman during the initial 55 hours of Gemini Flight GT-7 (Adey, Kado, and Walter, 1967). Flight data were compared with extensive baseline data from Commander Borman on a simulator and during sleep. EEG records revealed significant arousal before launch, with strong orienting reactions indicated during the first orbit. During the remaining 55 hours, there was an increase in theta power (between 4 and 7 Hz) which the authors interpreted as a physiological response to the weightless environment, similar to an orienting response. Sleep records revealed minimal sleep during the first night in space, with normal 90-minute sleep cycles returning on the second night. These results essentially agreed with Russian reports of EEG records taken in space on Cosmonauts Nikolayev, Popovich, Bykovsky, and Tereshkova. They permitted documentation of the effects of the weightless environment on the central nervous system. While the techniques used were crude by today's standards, especially in the lack of specificity of the overall environment and individual stimuli contributing to the gross EEG, these efforts stand as courageous attempts to carry the state-of-the-art in bioelectric recording to new heights. The best indications that they were successful in this is the success of new techniques in answering applied problems.

The brief summary above gives some indication of the scope of EEG applications up to the near-present time. However, unlike most other areas of psychophysiology, discussion of the EEG cannot be structured around a distinct cutoff point where interest waned and then returned. The EEG has shown a slow, steady growth in application from the post World War II period. There were no massive surges of interest in which the EEG was presented as a cure-all for psychophysiological measurement, as there were for other techniques. Such surges were usually followed by disillusionment and disuse. The EEG was technical enough that interest in it was always restricted. Its contributions were minimal in the early days. This permitted serious workers to develop it systematically, and the dividends from this approach are beginning to come in. The EEG is by

far the most powerful single psychophysiological tool presently available for applied research, and subsequent sections on sensory and cognitive measures will detail its uses.

MEASURES OF EYE FUNCTION

Introduction. The eyes are, at one and the same time, our most active receivers of information from the outside world, and the most direct and accessible source of information about the internal state of the central nervous system. Quite literally, the eye is a part of the brain, and it is unavoidable that it should be considered a primary candidate for psychophysiological measurement. In addition to its attractive accessibility, the eye does so many things, and does them with such a range of variability, that it could provide lifetimes of research effort exploring its subtleties. The eye moves -- in all directions -- innervated by six striated muscles. The 100-degree diameter circular motion permitted by these muscles provide a wide field for viewing, and for the psychophysiologicalist to monitor. Not only do the eyes move, but they can perform several kinds of movement. A single eye can move in the normal scanning way, or it can rotate about its own axis to produce a "roll". The two eyes together can converge (turn inward), diverge (turn outward), or make conjugate movements (together).

The pupil of the eye can also change size. The pupil is continually showing microscopic physiological oscillation in size called "hippus". Beyond this, the pupil changes in size to every change in ambient light, and to every change in reflected light entering the eye. The range of pupil size is from 1.5 mm to 9 mm, and the latency of the changes is small enough (as low as .2 sec) to make rapid response possible.

The eye also covers itself periodically -- it blinks. There are different types of blinks, but the ones of primary concern to the psychophysiologicalist last about .35 second and occur on the average of 7.5 times per minute, with extremely wide variation (Hassett, 1978). The pattern of blinking seems, even on superficial examination, to reflect psychological states such as anxiety or embarrassment, so the psychophysiologicalist again has an ideal measure to use in assessing the individual.

Because of all these factors, the study of the eyes is perhaps one of the oldest in psychophysiology. Observations of eye patterns have been used since the time of Confucius. Yet, it was only with the development of objective recording of eye parameters that this area was able to be scientifically studied, and common sense observations put to the test. These early, objective techniques included mechanical linkages to the eyelid to detect blinks, primitive photography to record pupil size, and contact lenses on the cornea linked to a recorder. Needless to say, they were not very successful in obtaining records that were even close to the normal, unobstructed activity of the person. This is, to some extent, a problem which still exists today even with modern recording techniques.

Measurement Techniques. More modern approaches to recording the eye's activity range from extremely sophisticated to extremely simple. Blinks provide perhaps the best example of the range of sophistication. On the one hand, complex electronic circuitry has been designed to detect blinks and measure their duration (see McGillem, 1979). In this system, a modification of one described by Bahill, Clark and Stark (1975), eyeglasses containing an infrared (I.R.) source are worn. The I. R. light is reflected off the eye, and a photoelectric cell is used to measure very small changes in the intensity of the reflected light. If a blink occurs, the reflection is, of course, blocked. Such blocks can be counted and timed. Obviously, one cannot simply expose the eye to indefinite amounts of I.R., so the overall utility of this technique is limited, although in practice this is not usually a serious problem.

At a slightly less sophisticated level, one can simply record blinks by placing electrodes appropriately around the eyes. Almost any electrode arrangement will yield a large deflection whenever there is a blink. However, the problem is that many other things yield similar deflections, even in the absence of a blink (e.g., certain eye movements, or cheek or forehead twitches). Most eye movements can be discriminated from blinks if an electrode montage recommended by Rechtschaffen and Kales (1971) is used (See Figure 9A). In this technique, electrodes are placed on the temporal side of each eye, with one placed 1 cm above and the other 1 cm below their respective canthi. A single mastoid electrode is then used as reference for each eye electrode, creating two channels of this "electrooculogram" (EOG). When properly connected, these two channels will produce deflections in the records which are in-phase whenever there is a blink (or muscle twitch) and out-of-phase whenever there is a conjugate eye movement. Again, while allowing a good approximate count of eyeblinks, this technique may confuse twitches. If used carefully, however, this EOG system, or a number of others, may accurately reflect the occurrence, duration, and even the shape of an individual blink (Stern, 1974).

At the lowest end of sophistication, most researchers dealing with eyeblinks simply use some form of photographic or direct recording. Given the general imprecision of defining an eye blink in an electrical recording, this is not always a bad alternative. At any rate, since there is no generally accepted procedure for counting blinks, one will either have to use a specially designed procedure, or obtain estimates from techniques like the EOG.

Recordings of pupil size also show a wide range of sophistication. At one extreme, photography, with or without an automated system of measuring the pupil from the photo, seems to be the most practical and time-tested technique (Hess, 1972). Such systems have been reported for use in infants, and in situations where the subject can be at some considerable distance from the camera. Lowenstein and Loewenfeld (1958) report an electronic pupilometer which had some popularity in clinical applications. Hess (1972) gives a list of commercial manufacturers of pupilometers up to that date. Automated pupilometers are on the market which measure and record pupil dilation, project the scene being viewed on the screen along with the point of regard, and put it all on FM and video tape. Modern systems for measuring changes in pupil size are extremely precise in the laboratory setting (Beatty, 1966). Using these, differences as small as one or two millimeters are usually consistent enough to be statistically significant when extraneous influences are well controlled.

Perhaps no area of eye recording has received as much attention, either procedurally or experimentally, as eye movements themselves. Again, photography provides the simplest method for determining the approximate position of the eyes. However, this is seldom sufficiently precise for research applications, and a long series of increasingly sophisticated pieces of equipment have become available which allow the researcher to

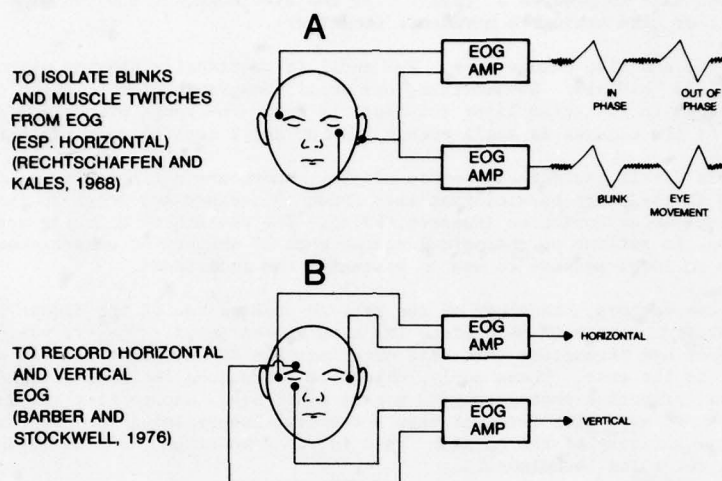


Figure 9. Two techniques for recording eye movements for specific purposes.

choose the degree of precision desired. These are well described in an excellent review by Tursky (1974b) and only a few examples of these techniques will be described here.

The Royal Aircraft Establishment developed an Eye Point of Regard Recording System for use in flight (Cox, 1973). This system used a video camera attached to a head harness in such a way that the scene being viewed by the pilot is recorded on video tape. A beam of light is projected from the harness and reflected off the subject's cornea and into a second video tube. The images from the two video tubes are then superimposed on the video tape. As the subject's eyes move, the reflected image is displaced in a directly proportional way to the movement. With careful calibration, this image can be adjusted to fall at the point the subject is viewing, with reasonable accuracy. The composite video tape picture then displays a changing visual scene as the pilot moves his head, along with a "flying dot" which always reflects the point of regard. Similar point of regard systems have been described by a number of others (see Leycock, 1974; Tursky, 1974b). These systems can provide good precision (less than 1 percent when operating at top calibration) and are rapid enough to detect eye movements which last only .1 second. However, calibration is rather difficult, and can take up to 30 minutes for a difficult subject. Users report that, once calibrated, the signal is relatively stable. However, it is hard to design a head harness or helmet which fits tightly enough to remain stable with head movements, and is still comfortable. The Royal Aircraft Establishment system appears relatively unencumbering and trouble free in this respect.

A more sophisticated system for tracking eye movements without encumbering the operator has been designed by Honeywell Radiation Center, Lexington, Massachusetts, for the U. S. Air Force and NASA. This system has

gone through several versions, mainly differing in the range of head movement allowed to the pilot before measures could no longer be obtained (Merchant, Morrisette, and Porterfield, 1974). In one version, the sensor unit is located 28 to 40 inches from the subject. This contains a silicon target vidicon operated in a standard TV camera. An infrared radiation source reflects a beam off the cornea and into the vidicon. In addition, an image of the illumination aperture is formed on the retina and refracted back out of the eye and into the camera. This produces a "bright pupil" image on the screen, along with the reflected corneal IR spot. From these two sources of information, computer programs calculate the position of the eye. In earlier versions, the subject's eye was free to move within a one cubic inch area without breaking lock. In more recent versions, addition of a two-axis moving-mirror system and a servo-controlled focusing lens permits subjects to move within a one cubic foot area. However, if head movement is very rapid, the image can be lost. In this case, the system attempts to re-locate the eye within the cubic foot space. Line-of-sight angles possible with this instrument range from ± 30 degrees in azimuth and from 10 degrees below to 30 degrees above the instrument. However, below 0 degrees, eyelashes can cause interference, and above 20 degrees azimuth, tears can distort signals from the temporal side.

Different forms of this oculometer have been used in the Air Force Aerospace Medical Research Laboratory, and the Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio., to study eye tracking of targets. These have been incorporated in laboratory studies of helmet-mounted displays and sights, and plans are under-way to incorporate the entire system into a helmet. Other types of oculometers, using slightly different principles, are also available, but have not received the degree of attention given to the Honeywell system.

Unlike the automated data analysis of the Honeywell oculometers, data reduction from other eye point of regard systems can be very laborious. Videotape records must be manually scored in most cases. This means that for some very brief interval (usually .1 second to be safe) a person has to look at a frame and determine the position of the eye. In a long session, it could take days to reduce the data from one subject. Electronic techniques for reducing such data are of course conceivable, but these turn out to be as expensive as the apparatus itself. It is still accepted as a fact of life by many researchers using point of regard systems that they will spend a much longer time in data reduction than in doing the experiment itself.

If one can tolerate less precision than the above systems, a simple electrooculogram (EOG) can prove adequate. Again, many arrangements of electrodes are possible, and they differ primarily in their ability to differentiate eye movements from blinks (Figure 9), and in their sensitivity to vertical and horizontal movement. In general, horizontal eye movements are easier to detect than vertical. If blinks are no problem, electrodes can simply be placed near the outer canthi and linked together. This is frequently done in recording nystagmus (Barber and Stockwell, 1976). If vertical eye movements are important, one electrode can be placed above or below the eye and one to the side. This provides a very rough estimate of eye position. However, it must be remembered that vertical EOG is frequently not linear, and in combination with horizontal EOG, as in this electrode derivation, can yield many unknown derivations. Calibration is therefore extremely difficult if high precision is required.

If such precision is necessary in recording eye movements, blinks, and even accommodative changes in the eye, an oculometer such as the one commercially available from Stanford Research Institute, International, Palo Alto, California, is virtually required (Cornsweet, 1970). In this complex system, four I.R. sources (of very low intensity) are reflected off the subject's eye. To determine eye position, the fourth Purkinje image reflected from the back of the lens is detected automatically and processed by a hardwired computer. With extremely careful calibration, the instrument can be accurate to within 1 minute of arc, and even under less stringent conditions, accuracy within 10 minutes of arc is common. However, as might be expected, this apparatus is extremely sensitive to movement and to the subject's position. A very stable location is required. Although the equipment will attempt to maintain "lock" once it has been calibrated, it is not always possible to regain lock once it has been broken. Therefore, it is sometimes necessary to perform some recalibration after breaks in the experiment, etc., even though a bite bar is used and the subject's head is in virtually the same position. A more difficult problem involves nonlinearity. Although, in theory, the oculometer should calculate eye position very accurately, it assumes linearity in some of its measures. This may be true for a limited segment of the population. However, subjects with spherical abnormalities, or many other peculiarities of the eye, have produced significant non-linearities in our experience. This problem has made the instrument extremely difficult to use in our laboratory. Of course, these nonlinearities can be compensated for by computer programs, but they do not constitute trivial efforts. In spite of these difficulties, however, there are many laboratory situations in which a system such as the Stanford oculometer is essential. Although these primarily lie in the basic research realm, many operational questions can best be answered by subjecting them to the precise laboratory analyses permitted with such a system.

It has been mentioned that the eye is capable of "rolling" movements, and this "ocular counterrolling" phenomenon has been studied in various operational contexts (Miller, 1961; Miller and Graybiel, 1965). No electronic way to record such counterrolling has been described, so researchers have had to rely on photography, using extreme closeups of the eye. Landmarks on the iris are then identified, and as the eye rolls around its axis, the degree of such roll is measured. Since the range of roll is less than 15° , this is a rather difficult data reduction problem, as well as consuming enormous amounts of time. In spite of such problems, the U. S. Navy group at Pensacola, Florida, have shown that the measure is reliable, and that it is a valid index of vestibular function. It has been used as one of the standard tests of vestibular sensitivity, and has been administered to the U. S. astronauts, as well as to hundreds of aviators and other research subjects.

The above discussion of systems for obtaining eye measures is, of course, incomplete. It is meant mainly to suggest the types of measurement available and the range of options available. For a more complete description of eye movement recording devices, the reader is referred to Tursky (1974b). Few techniques have been standardized, so the worker in this area must struggle to find the balance between precision and practicality which is optimal for a particular purpose.

Applications. Of all the more common eye recording techniques, perhaps the one that has received the least attention in actual human engineering contexts is blinking. Although a very large number of studies involving reading behavior have used blink measures, these were seldom related to other specific real-world tasks. Similarly, many investigations using eyeblink measures as dependent variables studied generalized

psychological states such as 'information processing need' (Poulton and Gregory, 1952; Baumstimler and Parrot, 1971), 'visual fatigue' (Luckiesh and Moss, 1942; Carmichael and Dearborn, 1947); 'fear', and 'increased activation'. However, controversy about such studies was always high, and Hall and Cusack (1972) provide an extensive critical review of the eyeblink literature to that date. Emphasizing the disturbing individual differences found in so many studies, and pointing out numerous other sources of contamination in such studies, these authors conclude that not one study to that time could be considered adequate. They do believe, however, that blinking rate increases on both ends of the activity continuum (the familiar inverted U relationship) with minimal blink rates at normal processing levels of attention.

There is no intrinsic reason, even considering the above criticisms, that the study of eyeblinks could not prove to be a useful adjunct to other psychophysiological techniques. New developments in this area are beginning to emerge. Stern (1978) has described techniques which measure the duration and timing rather than the frequency of eyeblinks. Based on observations that longer blinks appear to occur in some states where the central nervous system might be assumed to be degraded (Kopriva, Horvath, and Stern, 1971) it is proposed that these long duration blinks may index "drop-outs" in performance. Stern has suggested that blink durations of 130 milliseconds or perhaps even less may actually indicate behavioral "blackout" periods and may predict performance capability. In view of this kind of renewed interest, it is likely that the study of eyeblinks will assume a much greater role in investigations of such phenomena as fatigue and attention.

If eyeblinks have received the least attention from the applied researcher, the study of pupil size has received the most well publicized attention. This is due primarily to the efforts of Eckhard Hess (1975). The scientific study of changes in pupil diameter, although many years old, really dates from 1960, when Hess and Polt (1960) published the observation that male and female pupil dilation differed when pictures of members of each sex were viewed. In each case, pupil dilation corresponded to increased interest. Hess also reported pupillary contraction in situations which could be considered unpleasant or aversive (Hess, 1972; Bergum and Lehr, 1966; Barlow, 1969). A large series of studies were stimulated by this work, and these succeeded in demonstrating that sexual preference, and preference for certain commercial products could be identified with some accuracy by pupillometry (Hess, 1968; Hess and Polt, 1966). Although these studies were already coming under attack by the early 1970s, Hess declared in 1972 that pleasant stimuli or positive affect had a dilating influence on the pupils, whereas negative or aversive stimuli had a constricting effect. He related the positive-pleasant dimension to sympathetic firing, generating dilation, whereas the negative stimulus caused a parasympathetic type of response. Hess also did not eliminate the possibility that pupil size was directly affected by the central nervous system (Hess, 1972).

The range of uses to which the pupilometer has been put is enormous. Researchers have used it to study eye disorders, political and racial attitudes, drug effects, teachers' reactions to students, and effects of pictures of children on child molesters (Rice, 1974). However, pupillometry, after enjoying a long period of financial and scientific success, has been attacked vigorously over the past ten years. Reviews of the evidence of Goldwater (1972) and Janisse (1973) were extremely critical of Hess' contention that pupil constriction or dilation was at all related to likes or dislikes. At most, it was conceded that the pupil may respond with dilation when the scene being viewed is interesting. A variety of methodological problems may account for the apparent constriction of the pupil, including a "rebound" from previous stimuli, the brightness of the supposedly unpleasant scene, and the movement pattern of the subject (Janisse and Peavler, 1974).

One other aspect of pupillometry has fared better than the "like/dislike" interpretation, and forms the basis for the potential application of this technique to aircraft design problems. Hess (1965) reported that the pupil dilates during the time a person is solving a problem such as mental arithmetic, reaching a maximum just before the solution. As soon as the answer is given, the pupil usually returns to normal size. Further, the degree of dilation appeared related to the difficulty of the problem or effort involved in solving it. Kahneman and Beatty (1966, 1967; Beatty and Kahneman, 1966; Kahneman, Beatty, and Pollack, 1967) have validated this relationship, and extended it to indicate that there is a linear relationship between pupil dilation and the amount of material stored in short term memory. In their experiments, they used a digit recall task, and loaded the subject up to the limit of short-term memory. The same measure indexed the difficulty of a tone discrimination task. Distraction, or loading by a secondary memory task, also appeared to affect pupillary dilation, reducing its magnitude below that in the undistracted state. These types of results have been confirmed by Paivio and Simpson (1966) in Ontario, using abstract and concrete words to be visualized. The relatively more difficult task of visualizing abstract words led to larger pupil dilation (Simpson and Paivio, 1966; Paivio, 1966).

Practical applications of these later results have been reported. Janisse and Peavler (1974) report a study using pupillometry to assess the relative difficulty of a telephone operator's task. Two methods of looking up numbers were instituted, and wider pupil dilation was found with the method behaviorally described as being more difficult and causing more fatigue. Rice (1974) reports that an airline has used pupillometry to assess the stress response of job applicants for a pilot's position. Obviously, these applications are a long way from widespread use of this methodology in human engineering, but they do indicate the potential of the technique.

On the other end of the utilization continuum from eye blinks stands the phenomenon of eye movement. If, by eye movement recording, one includes all efforts to monitor the position or activity of the eyeball itself, this certainly has to qualify as the most frequently used eye measure. From eye point of regard studies, to measures of nystagmus, to all-night sleep recordings, researchers have been interested for centuries in telling whether, when, and where the eyes are moving. The study of eye movements in reading has, by itself, produced an entire literature which will not be covered here (see Tinker, 1958). Similarly, eye movements indicating hemispheric activation (Ornstein, 1973) and those used to study basic visual processes (such as stabilized images) will not be discussed due to their tenuous connection to immediate applications. Instead, the present section will concentrate on eye point of regard studies in applied settings, and on the study of nystagmus and related vestibular phenomena.

Eye point of regard measurement apparatus has already been discussed (p.19). The range and frequency of attempts to apply such techniques is at least as large as the number of techniques. Leycock (1974) has developed a good bibliography of these early applications, and only a few representative ones will be described here. These were chosen because they present a reasonable range of studies, and because they point up both the advantages and limitations of this technique. Perhaps the major purpose of most investigators using eye movement recorders in applied settings is to determine the instruments, controls and information sources used by the operator in performing a task, as well as the sequence and timing of such viewing patterns. Thus, cockpit displays, instrument layouts, visual scenes, etc. are given to the operator, and scan patterns are studied.

Philco Corporation used an eye point of regard apparatus to determine the scan pattern of subjects viewing a series of dials, instruments and tapes for required information. Various configurations of the instruments were presented to the subject, and the speed of scanning was forced by requiring actions in briefer and briefer intervals without reducing information input requirements. There was a strong hint that as learning progresses, the human begins to take in information peripherally rather than centrally. As the time limit became shorter, eye fixations began to be made not on the controls themselves, but at a point between them. It was as if the person was obtaining information from two places simultaneously in the interest of time (Goldbeck and Charlet, 1974; 1975).

The use of eye movement recording in flight simulators has become virtually routine with many aircraft manufacturers. Similarly, many attempts to measure eye movements in-flight have been carried out. For example, Cox (1974, 1975) used an eye point of regard system to measure pilots' scan patterns, both inside and outside the aircraft, during an ILS approach in Devon, Comet and VC-10 aircraft. Data were reduced on a frame by frame basis, and histograms of instrument utilization were calculated. Results revealed considerable differences between pilots in instrument utilization, even under similar conditions and missions. Figure 10 shows that pilots A and B utilized the HSI most, but that the second most utilized instrument for Pilot A was the ASI, while for pilot B it was the AI. Such differences apparently reflected real differences between the two which could be related both to personality differences and to landing style.

A more developmental, laboratory use of eye movement recording is represented by the work of Spicuzza, Pinkus, Klug and O'Donnell (1974). A computer graphics terminal was used to generate a simulated set of flight displays from an aircraft approximating the dynamics of a lightweight fighter. Subjects then "flew" this aircraft at different levels of mission difficulty, imposing different workloads within the same overall mission requirements. Eye movements were videotaped and analyzed in brief intervals. Interest in this study was in developing a system analyze the pattern of eye movements over a long period of time, and in seeing whether such patterns related to the workload of the operator. For the pattern analysis, a technique developed by Nirenberg, Haber, and Moise (1973), was adapted to these kinds of data. Conditional probabilities were

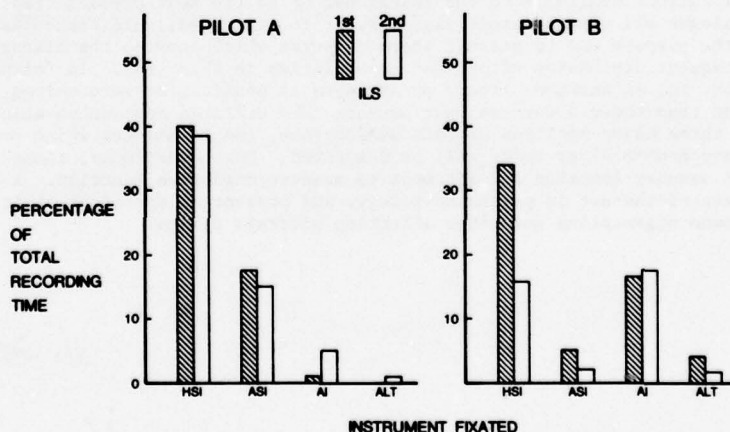


Figure 10. Eye fixation time on four flight instruments for two pilots during two successive ILS approaches. (Based on Cox, 1975.)

calculated for each of the six displays in the array. Each time the eye moved to a particular display dial, the probability that the subject would next move to each of the other displays was determined. For instance, how likely was it that the subject would move from the altimeter to the vertical velocity indicator? Next, the probability of a given eye movement after two preceding eye movements was calculated. Given that the pilot had viewed the altimeter and then the vertical velocity indicator, what was the probability he would then view the G-meter? These calculations were made for all combinations of up to five successive displays. Obviously, this creates an enormous amount of data. However, since only certain combinations of display sequencing are usually of interest, it is possible to reduce the data quickly to those patterns that are meaningful. Using this technique, it was demonstrated that the pattern of eye movements does in fact change with increased workload, and that this change tends toward elimination of certain non-essential information. Such an analysis system, while time consuming, provides a sophisticated, comprehensive way not only to study time spent on one display, but the overall interaction between displays.

In general, then, it appears that eye point of regard recording provides one of the more reliable and stable measures for applied human engineering. It is likely to be used as a control for many other measures, in order to assure that the subject is in fact doing the task assigned, even when it is not in itself the primary focus of the study. For this reason, it is critically important that a safe, accurate, non-encumbering technique of reasonable accuracy be standardized. Many of the current eye "trackers" available (such as the one produced by the Honeywell Corp) are close to fulfilling these criteria, and their further development should be encouraged.

A second major reason for using eye movement recording deals with the functioning of the vestibular apparatus (Barber and Stockwell, 1976). The intimate anatomical links between the inner ear and the eye allow precise information about many inner ear functions to be monitored by watching or recording the eye. From simple rotational nystagmus, to the ocular counterrolling measures already discussed, the eye is a sensitive indicator of the functioning of both the otoliths and canals. Obviously, from a basic medical physiological point of view, this is extremely important because it permits assessment of a mechanism inaccessible in any other non-invasive way. Less obviously, these mechanisms are important for the human engineer, especially one concerned with aerospace vehicles. At one extreme, much information about an aircraft is still gained from "the seat of the pants" which, to a large extent, is really located in the ear. Even in systems where instruments are the major source of information, it is still necessary to design in such a way as to minimize vestibular-visual conflict. At the other extreme, space flight provides a new vestibular environment about which we know very little. The basic mechanism of the vestibular apparatus has been exhaustively studied. Using measures based on the degree of nystagmus (Miller, Graybiel, and Kellogg, 1966), ocular counterrolling (Miller and Graybiel, 1965), and the oculogyric illusion (Roman, Warren, and Graybiel, 1963), investigators defined the sensitivity of the vestibular system in normal gravity and zero-G conditions. A variety of devices, including spin chairs, the centrifuge, and even a slow rotating room were used to generate virtually every kind of accelerative input, and to test interactions between factors likely to be encountered in flight or in space. In view of recent renewed interest in the problem of motion sickness in space, such studies will continue to have high priority, and eye movement measures will continue to be required.

SUMMARY

This section has presented a general summary of the major psychophysiological techniques, methods, and representative applications available to the researcher up to the near-present time. It intentionally did not attempt to catalogue all physiological measures, or to mimic available textbooks in physiological psychology. Instead, the purpose was to present those measures which provide the historical perspective and foundation for subsequent discussion of present capabilities in this area. In following this limited purpose, many techniques, and an enormous litany of attempts at application were omitted, not because they were less important than those presented, but because they utilized approaches which were already included. In the next three major sections of this AGARDograph, the techniques which are assuming greatest importance in applied psychophysiology today will be described. For convenience, these are broken down into attempts to assess sensory function and attempts to measure cognitive function. A final section will discuss the general state-of-the-art in psychophysiology, and present an appraisal of its future possibilities in the assessment of human engineering questions affecting aircraft design.

EVALUATION OF SENSORY FUNCTION

INTRODUCTION

In a very real sense, an aircraft or other vehicle represents an extension of the sensory and motor capabilities of the human. The system reports objects, positions, or environmental conditions in the same way that human senses report such factors. Receiving inputs from the person, the mechanical system responds with movements or alterations in capacity much as the human motor system responds to the dictates of the brain and nerves. It is well-established in engineering design that the status of such sensing and responding capabilities in the mechanical system must be monitored often so that changes in sensitivity, accuracy, or reliability will be corrected prior to catastrophic failure. It is equally important, though less generally recognized, that the sensory and motor capabilities of the operator should also be evaluated as continuously as possible.

In practice, this proves to be a much more difficult task with respect to the human operator than it is for the mechanical system. In the vehicle, it is possible to perform periodic maintenance and routine inspections. In the human such gross periodic checks take the form of medical examinations, vision and hearing tests, etc. These alert the operator or supervisor to long-term degradations in sensory and motor capabilities of the person. However, in mechanical systems, it is also possible to provide on-line analysis mechanisms which monitor the system continuously. At the very least, these activate a warning signal when failure is imminent. This is typically done through the use of probe points. A sensor is permanently attached to the system in question, and these probes provide a continuous read-out on the status of the system. Engine instruments, temperature sensors, and landing gear indicators are common examples of such test points. With respect to the human, no readily accessible test points can be identified, and it is an important function of human engineering to search for and develop alternative ways to continuously monitor the sensory and motor capability of the operator. Psychophysiological measures are assuming greater importance in this respect (Donchin, 1978). The types of techniques and measures presently used to accomplish these goals will be surveyed in this section.

It has been estimated that 80% of the sensory input utilized by the operator in an aircraft is provided visually. For this reason, emphasis in the present chapter will be on visual input. Audition constitutes the next most important channel of sensory input to the operator, and therefore will similarly be emphasized. Finally, vestibular input, as noted in the last section, continues to be of interest to the human engineer, and will be discussed briefly. Throughout, a somewhat arbitrary and rather narrow definition of sensory function will be maintained, to involve only those techniques which evaluate the present state of the pure sensing apparatus. Therefore, differential thresholds, where the subject is required to decide whether two things are different, and signal detection tasks, where the subject must identify the target detected, are treated in the following section on cognitive functions. It is recognized that this is a somewhat unorthodox approach. However, within the context of the human engineer's evaluative function, it is more productive to treat the pure sensing and the decision making capabilities separately. In many cases, of course, the techniques used to probe the human system will be the same for both sensory and cognitive function. It will therefore be possible to describe many techniques in the present section and then simply to reference them in the later section.

VISUAL INPUT

It has been said that the most difficult part of a visual display evaluation is the last few inches, between the operator's eyes and brain. Relatively well-developed techniques exist for evaluating the engineering characteristics of visual displays and visual scenes. These enable the design engineer to calculate modulation transfer functions through various optical devices, and to specify the characteristics of a displayed scene in terms of spatial frequency, contrast ratios, grey shade characteristics, and a variety of other objective measurements (Cornsweet, 1970). Up to recent times, psychologists have not been able to achieve this kind of precision. Major problems arise when energy begins to be transmitted through the eye. Formulations of attenuation, scatter, absorption, etc., by the ocular media have been developed (Zegers, 1959). However, these are probabilistic and, to some extent, idealistic estimates, and are of limited value in the on-line evaluation situation. With respect to physical events occurring in the visual system behind the retina, basic theory provides even fewer clues with respect to human engineering applications, though significant advances are being made.

The present section describes the techniques which have been developed to evaluate the last few inches of the visual display problem. In most cases, these techniques evaluate the quality of the final product of an interaction between environmental characteristics (the sharpness of the viewed scene, contrast, color, spatial frequency composition, etc.) and the integrity of the visual system. To be used accurately, it is necessary to remember that these probes reflect both factors. Experimental designs must carefully control one factor if inferences are to be made about the other. In psychophysics, this is usually done by employing well-trained, idealized subjects operating at maximum efficiency. This effectively controls the variability which might be introduced by the physiological and psychological side of the interaction. These designs are excellent for revealing basic effects and functions. Too often, however, these basic results are then extrapolated to real-world situations with no modification. For example, for years it was argued that because carbon monoxide in small doses caused changes in the psychophysical threshold of vision, complex behaviors such as driving and flying would be affected. Yet, when such complex behaviors were tested, no effects were found (O'Donnell, Chikos, and Theodore, 1971).

The basic point simply reiterates the obvious but not always appreciated fact that, while it is important to measure basic sensory capacity, and even more important to understand how sensory systems work, it is not always a trivial matter to extrapolate from subtle psychophysical measures to real world performance. The tools described in this chapter permit rather precise measurement of changes. It is quite another matter to decide that these changes are meaningful in a system context. If they are to be used most productively, the impact, in real, meaningful terms, of a sensory change or condition must be demonstrated in an operational sense. The contribution of psychophysiology to human engineering is in providing more precise and detailed information. This contribution could be useless or even counter-productive unless it is realized that an alteration in sensory capacity, or a new measure of sensory skill, is only of academic interest to the human

engineer until it is demonstrated to produce a change in some real-world performance.

Visual Sensitivity Measurement

One of the elementary questions concerning the visual system deals with minimum detectable quantities, usually quantities of light (Brown, 1973). The most common, traditionally accepted techniques for testing these threshold limits are psychophysical. They utilize a dark or light adapted eye, presenting very low intensity light, usually as a small pinpoint in a dark field. Complete light adaptation can be obtained by exposing the subject to a bright light for some period of time over one minute in order to assure uniform bleaching of all the photosensitive pigment. The light is then extinguished and, in total darkness, the subject either adjusts a small light until it is seen, or the intensity of a flashing light is increased until the subject reports seeing it. The light intensity is then either increased above threshold and slowly decreased until the subject cannot see it, or it is reduced below threshold and increased until the subject sees it. This procedure is carried on for an extended period of time (typically from 10 to 30 minutes) and the intensities at which the subject saw the light, when plotted against time, reveal the changing absolute threshold as the visual pigment is replaced after bleaching. In this way, photopic and scotopic sensitivity curves can be plotted. If the test stimulus covers both rods and cones, the familiar "rod-cone break" is seen between 5 and 10 minutes as cone vision becomes maximally sensitive, while the rods continue to increase in sensitivity (Woodworth and Schlosberg, 1954).

In addition to the above techniques for determining the recovery of sensitivity after light adaptation, the completely dark adapted subject can also be tested with targets of different size and wavelength to determine absolute sensitivity. Such studies reveal remarkable sensitivity for both rod and cone systems (Hecht, Schlaer, and Pirenne, 1942; Zegers, 1959) indicating that a visual experience is possible with stimulation by as few as two quanta of light per cone in parafoveal vision.

Several types of adaptometers exist for determining absolute visual threshold in the ways described above. Some of these utilize a split field technique (Dixon, 1958; McLaughlin, 1954) while others use a solid test field (Blackwell, Pritchard, and Ohmart, 1954; Schaefer, 1949; Wald, 1945; Weidemann, 1952). Several devices have been used to obtain a single estimate of the subject's sensitivity after maximum dark adaptation. These utilize a Landolt Ring (Pinson and Chapanis, 1945; Rowland and Mandelbaum, 1944), Aircraft Forms (Rowland and Mandelbaum, 1944) or Graded Series of Figures (Della Casa and Birkhauser, 1946). In general, these techniques establish the theoretical limit of detectability for these kinds of targets. Although they have had limited use in on-line applied contexts, they have been used extensively to develop handbook specifications of minimum detectability criteria for such things as external lights, paint schemes for aircraft, and survival markers.

Modern psychophysics has tended to define both the problems and measurement of visual sensitivity somewhat differently than traditional techniques. Advances in the basic visual sciences made it clear that there is not one single value or adaptation curve which is adequate to represent the absolute sensitivity of the visual system. Rather, there may be specific detection mechanisms tuned to particular orientations of the visual stimulus, or to narrow bands of spatial frequencies (Blakemore and Campbell, 1969; Blakemore, Nachmias, and Sutton, 1970; Carter and Henning, 1971; Hubel and Wiesel, 1959). Westheimer (1972) discusses the reasons for utilizing spatial frequency analysis to describe visual stimuli, and the implications of its use (see also Cornsweet, 1970). Although the theoretical advantages and disadvantages become quite complex (see Sekuler, 1974) the impact on techniques for determining threshold sensitivity in applied contexts is certain to be considerable. If the visual system shows differential sensitivity to various kinds of stimuli such as spatial frequency patterns, as is obviously the case (Campbell and Greene, 1965; Cornsweet, 1970; Davidson, 1968) it is no longer sufficient simply to measure the visual sensitivity to a pinpoint or whole field flash. Techniques must be established to determine the sensitivity across the entire range of spatial frequencies. Such "contrast sensitivity" procedures are being utilized at the present time in several laboratories and have implications in the psychophysiological assessment of visual sensitivity. These will be discussed in more detail later.

The Electroretinogram. Electrophysiological methods of obtaining the absolute sensitivity functions of the human visual system have been confined predominately to the electroretinogram and the visually evoked response. The electroretinogram, being the representation of the gross electrical response of the retina to stimulation, has been used to plot dark adaptation curves in several laboratory situations. To obtain the traditional electroretinogram (ERG) it has been necessary to attach electrodes to a contact lens which is placed on the cornea. Another electrode is placed somewhere on the head where it can be expected to be influenced by the rear of the eyeball (Johnson, 1949). The signal recorded when a light pulse is seen in the dark adapted eye consists of a number of recognizable patterns which have been related, with varying degrees of success, to subjective phenomena (Bartley, 1951).

Basically, two distinct parts of the ERG can be identified in the above conditions. One part, reflecting receptor activity, is a short diphasic potential in which the cornea is initially negative (the a-wave). The second part is scotopic, and is a prolonged (.1 to .5 sec) monophasic potential in which the cornea is positive (b-wave). This wave is thought to reflect bipolar cell activity. The size of this potential varies with the degree of adaptation, and is directly related to visual sensitivity. This component can be used to approximate the dark adaptation curve, although there are a number of difficult methodological problems associated with this interpretation. It is thought that the ERG is primarily scotopic, and reflects rod activity (Armington, 1964; 1966; 1968; Armington, Corwin, and Marsetta, 1971). In any case, the use of a corneal electrode, which is very encumbering and uncomfortable to the subject, obviously limits the use of this technique.

A technique for obtaining the ERG from an electrode placed on the surface of the skin above the eye has been described (Tepas and Armington, 1962; Armington, 1974) and recent advances in standardizing the procedure have been made (Giltrow-Tyler, Crews, and Drasdo, 1978). This technique appears to sacrifice only minimum reliability and consistency in measurement, although it is not a trivial task to control for muscle potentials and other extraneous signals. Some averaging of the signal is typically necessary in order to reveal the underlying retinal response.

In spite of these limitations, the ERG as recorded from either corneal or skin electrodes could have considerable utility in operational environments. Since it is a rather specific and sensitive index of retinal function, it would be expected to be an early indication of any stressful condition which would affect transmission in the retina. Sub-acute anoxia typically diminishes the amplitude of the b-wave, as does any condition resulting in reduction of the blood supply to the eye (Ward, 1968; Wulffing, 1964). As such, it has been recommended that the ERG might be of value in observing the O_2 saturation of the eye after acceleration stress in the centrifuge (Miller, 1976) although care must be exercised in recording the signal. ERG measures taken during actual centrifugation, however, have failed to reveal any amplitude changes, even up to the point of blackout (Lewis and Duane, 1956). The ERG would be particularly useful if there was some question whether a given visual degradation was due strictly to retinal effects, or to insult further down in the nervous system. By combining this measure with cortical evoked response techniques described in the next section, it is possible to differentially locate the source of such degradation. Normal ERG with disrupted evoked responses definitively locates the insult at or beyond the ganglion cell level.

The Cortical Evoked Response. The second electrophysiological technique available for measuring visual sensitivity is the cortical evoked response. Some early investigators utilized pinpoint flashes of light to produce the response. Using the psychophysical method of constant stimuli, a range of intensity conditions was presented, and the lowest intensity giving a consistent evoked response was considered the threshold (Irwin, 1974). There is a considerable number of difficulties with this technique. It is probably true that some index of the visibility of the stimulus is obtained in this way. However, it is likely that this index is obtained from the later components of the evoked response (the P3 or P300) rather than earlier components which would truly reflect the sensory reception of the visual system. Thus, it would be difficult to isolate the sensory and perceptual components of a threshold.

On the other hand, a different type of evoked response technique discussed by Regan (1972; 1977a) adds considerable precision to such determinations. In the classical evoked response, the stimulus is presented at a rather slow rate, usually slower than one per second. The evoked response to each stimulus is then stored and averaged with all responses from preceding stimuli. In effect, this reveals the "transient" response of the brain to the stimulus, analogous to the transient produced in an electronic system when a pulse is introduced. If, instead of pulsing the system very slowly, the stimulus is presented rapidly, a "steady-state" is achieved in the brain (Figure 11). In effect, a microportion of the brain's activity becomes entrained with the temporal frequency of the stimulus. This microportion can be isolated from the other brain activity, and can be displayed as a sine-wave at the same frequency as the stimulus. In this way, a direct input-output relationship can be established. Since a sine-wave is obtained at the output it can be related directly to the pulsed or sine-wave modulated light at the input. The steady state response can vary only in amplitude or phase angle (delay) with respect to the input. This entraining of the brain's response, and the simplification of the waveform, eliminates much of the variability (and information) of the transient response, but produces a much more stable measure. The temporal frequencies which have traditionally been used to generate the steady state evoked response range between 4 and 30 Hz, with most stimulation being between 8 and 20 Hz.

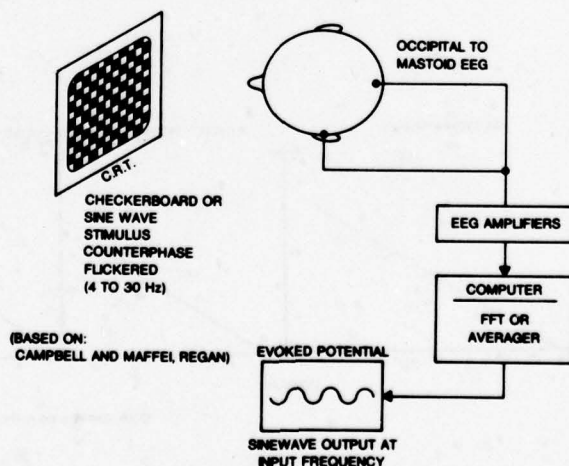


Figure 11. Schematic procedure for obtaining the steady-state evoked response.

Technically the steady state evoked response is somewhat more complicated to obtain than the transient evoked response. Not only is EEG amplification required, but some technique for filtering out all of the evoked response activity except at the stimulating frequency must also be employed. Two methods are used to accomplish this. In the first, analog or digital filters eliminate all activity outside of a narrow bandpass. In practice, this is a reasonably easy and precise technique. Responses to stimuli are still ensemble averaged, as with the transient evoked response. However, since most of the extraneous activity has been filtered out, the signal is built up rapidly. Thus, the experimenter can view the waveform at the stimulating frequency as it develops.

The steady-state response can be further quantified by performing spectral analysis on the already averaged waveform. This will, of course, reveal a sharp peak at the stimulating frequency. Under certain conditions, there will be additional peaks since harmonics may be present. The sharpness of the major peak will indicate the degree to which the brain is following the stimulating frequency precisely. This technique has been attempted with some success in evaluating optical properties of aircraft windscreens (Gomer and Bish, in press). Tentative indications were found that a "broader" spectral distribution about the major peak was found with optically unsatisfactory windscreens (Gomer, unpublished data).

A second major technique for measuring the steady-state evoked response requires a Fourier analysis to isolate the input frequency in the EEG output. In determining the Fast Fourier Transform (FFT), two terms are derived which can be used to determine both amplitude and phase angle of the Fourier component at the input frequency. It has been demonstrated by Regan (1973; 1975a; 1975b) that these terms accurately reflect the power in the ongoing EEG at the input frequency, and that the phase lag can be used to determine "apparent delay" between the input and the appearance of its representation in the evoked response. This apparent delay represents the transmission time of the optic pathway for that stimulus. Further, 10 seconds of data are usually sufficient to calculate these values, and under optimal conditions less time may be required. This technique virtually provides the researcher with an on-line procedure for utilizing steady-state evoked response.

Using an early version of the steady-state response, Campbell and Maffei (1970) determined the amplitudes produced when sine-wave gratings of three spatial frequencies were flickered at eight times per second. The contrast ratio between the dark and white areas of the sine wave grating was systematically varied for each spatial frequency. As contrast ratio increased (up to a limit) steady state amplitudes increased. These amplitudes were logarithmically related to the contrast ratio for each spatial frequency tested. (Figure 12). In addition, when the plots were extrapolated to a theoretical point where no evoked response would have been obtained, the contrast ratio agreed quite well with the subjects' psychophysically determined behavioral threshold, at least for the spatial frequencies tested.

The implications of this study are considerable with respect to methodology in psychophysics. If the above technique reveals the true absolute threshold of the visual system by a psychophysiological technique not requiring subjective report, it should contribute a great deal to reducing much of the variability typically found in psychophysical studies. It would also permit determination of psychophysical thresholds in untrained, non-ideal subjects, permitting a more valid determination of real sensitivity in the unselected populations which are frequently of more interest to the human engineer than the idealized subject.

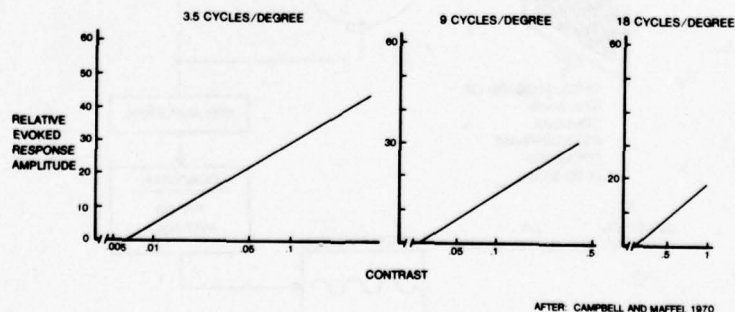


Figure 12. Plot of steady-state amplitude versus contrast ratio for three spatial frequencies.

It should be noted also in Figure 12 that the slope of the function relating steady state amplitude to contrast ratio is virtually the same for all three spatial frequencies tested. In this case, the slope of the function represents a change in physiological response (amplitude) with external visual stimulation (sine wave contrast ratio). In effect, the slope value represents unit physiological change per unit sensory change, and this appears to be constant for a range of stimulus values above absolute threshold. This slope, therefore, could be taken to indicate a "suprathreshold" sensitivity of the visual system to contrast ratio changes. This would be analogous to what psychophysicists measure with the small "just noticeable difference" (j.n.d.) metric which attempts to obtain the differential threshold of the subject. However, in behavioral psychophysics, it has not been possible to measure the absolute threshold and the differential threshold with a single metric. Absolute threshold is usually determined by presentation of a single stimulus many times at varying intensities, while differential thresholds virtually require comparison between two stimuli. If the steady state evoked response provides a single metric for both types of threshold, it will contribute immeasurably to standardizing the way that visual displays can be calibrated in aircraft systems. Such standardization would be even more desirable because of the increasing utilization of sine-wave and spatial frequency metrics by optical and display engineers. The ability to specify display psychophysical thresholds in the same metric that the engineer uses would represent a considerable advance over present, fractionated techniques, and could finally provide the psychologist with the same precision demanded from the system engineer.

Unfortunately, the application of these metrics to more complex imagery is not quite as simple as one might hope. In the study cited above by Campbell and Maffei, it was found that if the stimulating pattern consists of the addition of 2 spatial frequencies, the slopes of functions such as those presented in Figure 12 increase dramatically. There would rapidly be a limit to how steep this slope could become before it would lose interpretability from the viewpoint of the system designer. Since most complex imagery will consist of many spatial frequency components, the determination of steady-state evoked responses directly from the realistic visual scene will not readily produce a useful metric, at least as far as threshold sensitivity of the visual system is concerned. In spite of this difficulty, the attraction of a single metric to measure both threshold and suprathreshold functions, even if limited to one or two levels of spatial frequency complexity, is extremely powerful, and possible ways to utilize the evoked response in this way are presently being tested. One of these is described below under the heading of the modulation transfer function area.

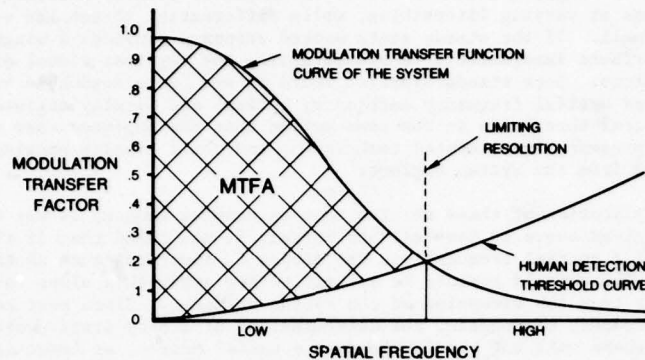
Specific characteristics of the cortical evoked response must be considered in the design of applied experiments. For instance, using both transient and steady-state techniques, general principles concerning the interaction between retinal location and evoked response amplitude have been determined. With increasing distances from the fovea, a pattern stimulus must be increased in size in order to produce the same amplitude evoked response, at least to 12 degrees (Harter, 1970). Stimulation of the upper retina produces the largest evoked responses, with the central retina producing the next largest, and the lower retina producing the smallest responses (Eason, White, and Bartlett, 1970). With respect to the horizontal and vertical meridians, larger responses are found closer to the vertical meridian, with a polarity change as stimulation changes from the upper to the lower field (Brown, 1973; Halliday and Michael, 1970). Overall evoked response variability has been studied by Callaway and Halliday (1973). These and a host of other parameters must be carefully controlled if the evoked response measures are to have any interpretability (Perry and Childers, 1969; Regan, 1972).

It is clear that the cortical evoked response is not a technique to be used in cavalier fashion. When questions involve such things as basic visual processes, and particularly where basic sensory physiological mechanisms are to be inferred from the results, the evoked response provides a good but extremely delicate instrument of measurement. It is, after all, the final representation of a number of intervening processes which can be influenced by many factors. Extremely tight control must be maintained over these factors if one is to make inferences concerning the meaning of a change in the response relative to the manipulation of some environmental factor. On the other hand, in many cases of applied psychophysiology, not all factors need always be specified in such detail. In these instances, one is interested only in knowing whether a difference exists between two complete designs, two systems, or two operator states. A change in the evoked response may only indicate the existence of an otherwise undetectable difference between conditions, and need not be over-interpreted in terms of mechanisms. It will still be of considerable value to the human engineer if it can simply be related to real world events. Like most measures, use of the evoked response for assessing visual sensitivity requires care and experience, but it can add significantly to the measurement capabilities of the investigator.

Modulation Transfer Function Area. One proposal for permitting complex aircraft display problems to be analyzed in manageable form utilizes the concept of the Modulation Transfer Function Area (MTFA) first proposed by Charman and Olin (1965) and further applied by Snyder (1976) and Keesee (1976) (see also Beamon and Snyder, 1975; Snyder, Keesee, Beamon, and Aschenbach, 1974). Although there are other techniques for describing display parameters, this one will be described here as illustrative.

The concept makes use of two sets of information. One curve is derived by measuring the response of a given optical or electrooptical system. Such a curve may be derived by presenting a standardized set of visual targets on the system in question. The target's spatial frequency composition is systematically changed, and the system response is measured. In the example shown in Figure 13 (based on Snyder *et al*, 1974) the upper curve shows a fall-off in system response as spatial frequency increases. In other words, this system produces less displayed modulation for smaller targets. This curve essentially represents the system modulation transfer function.

The second set of information is determined from the human observer, and represents the eye's ability to resolve the same target over a range of spatial frequencies. To obtain this, a contrast sensitivity curve is determined for the target in question (Cornsweet, 1970). This plot can be interpreted in terms of a "detection threshold curve" by converting it into the modulation terms used for the system MTF curve. This curve is then superimposed on the first curve, as shown in Figure 13. The intersection of the threshold curve and the system curve determines the limiting resolution of the system with respect to the human observer. Further the squared area between the curves (MTFASQ) is an indication of the overall quality of the displayed image. This measure has shown high correlation with observer performance.



AFTER: KEESEE, 1976

Figure 13. Schematic representation of the modulation transfer function area concept.

One of the problems with the MTFA concept has always been the inability to determine suprathreshold values in any usable way. As presented in Figure 13, the lower curve represents the absolute threshold of the human visual system. In practice, it is more important that such things as differential thresholds or changes in sensitivity of the visual system be determined. The steady state visual evoked response technique described above may provide one means for specifying the suprathreshold limits of the visual system. For instance, the slopes of the evoked response/contrast sensitivity functions, as discussed by Campbell and Maffei (1970) could be used at each spatial frequency, instead of the contrast ratio itself to determine the visual MTF curve (see Figure 12). Other derivatives of such a measure might also be used. The search for such a metric is currently being carried out in the Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, and if a valid usable suprathreshold metric is found, it will have a major impact on the methodology for specifying minimal criteria for visual displays in aircraft and other systems.

Visual Acuity

The absolute resolving power of the eye is its acuity. Traditionally, acuity has been considered to be the minimum visual angle that the eye is able to resolve, and has been defined for high contrasts in terms of either dark or white lines subtending minimum visual angles. Hecht (1947) had subjects report the presence or absence of a fine wire silhouetted against a bright sky. Using a criterion of 75 per cent correct response, he found adequate discrimination when the diameter of the wire subtended an angle of only .43 second and its length subtended only about 1 degree. More standardized procedures are used clinically and in many research applications. The Landolt Ring was adopted in 1909 as the standard test object by the International Ophthalmological Congress in Naples (Sloan, 1951). To some extent, this technique has been replaced by the familiar Snellen letters, first proposed in 1862. These present a series of figures, usually letters of the alphabet, which are graded in size. The subject views these letters from a fixed distance (frequently 20 feet) and determines the smallest size letter which can be read. The size of the letter read at this distance is then converted to a ratio by placing the actual viewing distance over the average distance at which a "normal" subject can resolve that letter. For instance, if at 20 feet the individual can only resolve letters typically read at 100 feet, visual acuity would be described as 20/100.

Several modifications of the Snellen letters have been carried out. Sloan (1959) proposed a system consisting only of capital letters. The acuity measurements given by these letters has been shown to be related to those obtained by the use of the Landolt Ring. Therefore, the two tests could be used interchangeably if repeated measurements on a given subject were required. A number of mechanical devices for testing visual acuity (among other visual functions) have been developed. These include the familiar Keystone Telebinocular and US Armed Forces Vision Tester. The Ortho-Rater, manufactured by Lafayette Instrument Company, USA, uses increasingly smaller squares to determine an acuity rating.

Several integrated systems for testing monocular and binocular functioning have been developed (see Decker, Williams, *et al.*, 1975). An extremely elaborate device is the Vision Analyzer manufactured by the Minneapolis Honeywell Corporation, USA. In this test, the subject sits at a box console approximately 4 x 4 x 4 feet. A separate console programs and delivers a preset series of stimuli, to which the subject responds. Landolt rings are used to measure acuity, and they are presented in the darkened field several times throughout the test session. An entire test, evaluating muscle imbalance, binocular fusion, phoria, color sensitivity, and other standard tests of visual function can take as long as 45 minutes, although portions of the test may be given separately. Scoring is automatic and is given in terms of stanines. Therefore, this test represents only a crude screening for visual function, and is essentially an automated version of the traditional tests noted above. Its major advantages lie in the consistency of test administration and interpretation, and in its automated administration, which makes it suitable for screening large populations. In that respect, it probably provides much more precision and standardization than previous tests.

A further consideration involves determination of dynamic visual acuity, since in many cases of operational interest one would like to measure the subject's ability to see small, rapidly moving targets. It is well known that acuity falls off rapidly when the target velocity exceeds the capability of the ocular muscles to produce smooth pursuit movements (Brown, 1972; Reading, 1972). No standard technique for measuring such acuity has been developed which is readily adaptable to operational environments. However, it is possible to present moving or drifting sine-wave gratings at various spatial frequencies on a CRT. The typical contrast sensitivity function can then be determined for these moving targets, yielding an estimate of dynamic visual sensitivity and, indirectly, acuity.

Psychophysiological measurement of acuity was limited for many years to the various optometric techniques available only to the optometrist or ophthalmologist. These included such standard devices as the ophthalmoscope and retinoscope, which could give the trained clinician an accurate estimate of the refractive error of a given eye. These continue to be the standard techniques available to the clinician. However, they have been of little value to the researcher looking for rapid and accurate ways to test individual subjects, particularly, on-line.

Although the above techniques have provided useful clinical approaches, and continue to be valuable even in limited research applications, they no longer can be considered precise for certain types of applications. Snellen acuity of 20/20 represents a resolving power of one minute of arc, or 30 cycles per degree (Marg, Freeman, Peltzman, and Goldstein, 1976). Since this is over twice the optimal resolving power of the visual system, it represents, at best, a convenient statistical convention concerning visual acuity. Something more precise than that is certainly required for the vast majority of experimental work. Further, the determination of visual acuity by most techniques is primarily a psychophysical procedure, and depends on a highly subjective judgement by an individual, as well as on numerous stimulus determinations. As such, its application in operational environments is limited. Therefore, it has been used predominately in developing the specifications found in handbooks and in other laboratory procedures.

It has been extremely difficult to develop measures of acuity which are useful in an operational setting. Yet, there is considerable need to do so. Visibility of targets from an aircraft cockpit or observation post, whether they be ground targets or other aircraft, continues to be of prime consideration in the design of systems. The introduction of radar and other automated sensing systems has not diminished this requirement. In fact, technological developments such as increased speed and maneuverability of aircraft have introduced a whole new series of problems affecting visual acuity. For instance, the increased thickness of aircraft windscreens in order to protect them from bird strikes had introduced serious questions concerning specifications of visibility through the windscreens. The need to specify the optical characteristics of the windscreens in terms of the meaningful visual acuity of the operator has proven to be an extremely difficult task. In view of requirements such as these, it is entirely appropriate that new concepts and techniques for measuring visual acuity be investigated.

Acuity and Spatial Frequency Contrast Sensitivity. As in the case of visual sensitivity, the question of visual acuity may be enriched significantly by consideration of spatial frequency analysis (Sekuler, 1974). The modulation transfer function discussed in the last section represents not only the sensitivity function for the human visual system, but also the ability of the visual system to transfer information at various spatial frequencies from stimulus input to output. In many ways, this is more valuable information than simply establishing an arbitrary acuity standard. Thus, an important concomitant of acuity can be viewed as the modulation transfer function of the individual across a wide range of spatial frequencies. Standard MTF curves can be calculated, and have been well reported (Campbell and Greene, 1965; Cornsweet, 1970; Davidson, 1968; Ohzu and Enoch, 1972). It is possible, therefore, to compare the contrast sensitivity curve of any individual to these standards. While this does not give a measure of visual acuity in the traditional sense, it may provide more information about the individual's real ability to resolve visual imagery than traditional measures.

Campbell and others (Mecocci and Spinelli, 1976; Van Ness and Bouman, 1967) have proposed that there may, in fact, be distinct channels in the human visual system which are differentially tuned or sensitive to stimuli of various spatial frequencies. Further, it has been argued (Ginsburg, 1978) that these channels can be independently decremented in some individuals, and that such decrements may not always be detectable by traditional acuity measures. Such an individual could test out perfectly, for instance, on Snellen acuity, and still have significant visual deficit. In fact, it has been shown by Ginsburg (*in press*) that the Snellen letters actually demand sensitivity in only a small portion of the entire range of spatial frequencies in order to be detectable. All of these factors argue for a redefinition of the techniques for measuring visual acuity in operational environments. While no standardized techniques have been yet developed, research along these lines is being carried out in many areas, and may eventually result in considerable alteration of the concept of visual acuity. It will be necessary that any psychophysiological techniques purporting to measure acuity be capable of adaptation along these altered lines.

Transient Evoked Response Measures of Acuity. As might be expected from the previous discussion on visual sensitivity, the cortical evoked response has been extensively utilized to assess visual acuity. Harter and White (1970) were among the first to note that there were systematic changes in the transient

visual evoked response to a patterned stimulus with progressive defocussing. These authors, using a checkerboard pattern flashed at the rate of one per second, found that the transient visual evoked response contained peaks which were maximal when the image was in focus. These peaks consisted of a negative-going peak at approximately 100 milliseconds, and a positive peak at approximately 200 milliseconds. As the image of the checkerboard was defocussed through use of lenses ranging from +6 to -6 diopters, the negative peak at 100 milliseconds gradually became positive, and the positive peak at 200 milliseconds disappeared. The author suggested that the spherical correction for an unknown eye could be determined by systematically inserting corrective lenses until an optimal evoked response was obtained. In subsequent studies (Harter, 1970; Harter, 1971; Harter and Suitt, 1970; Eason and Dudley, 1971; Eason, White, and Bartlett, 1970) some of the parameters affecting these responses were studied. For instance, it was found that the optimum target size for producing an evoked response is about 9 minutes of arc, and a checkerboard pattern produces larger evoked responses than a striped pattern. Overall, the conditions under which one could reasonably perform an evaluation of spherical refractive error of an individual have been well specified.

However, there are significant problems in utilizing this procedure. First, since the stimulus is presented at a rather slow rate (1 per second or slower) it may take a minute or more to obtain a single evoked response. Since, in the course of a refractive determination, many such responses would have to be taken, the demands for attention and cooperation from the subject are considerable. Some investigators (Marg, Freeman, Peltzman and Goldstein, 1976) insure that the stimulus will be triggered only when the subject is looking at the pattern. The investigator manually triggers the flash when it is clear that the subject is paying attention. In this way, it has been possible to test infants as young as three weeks of age.

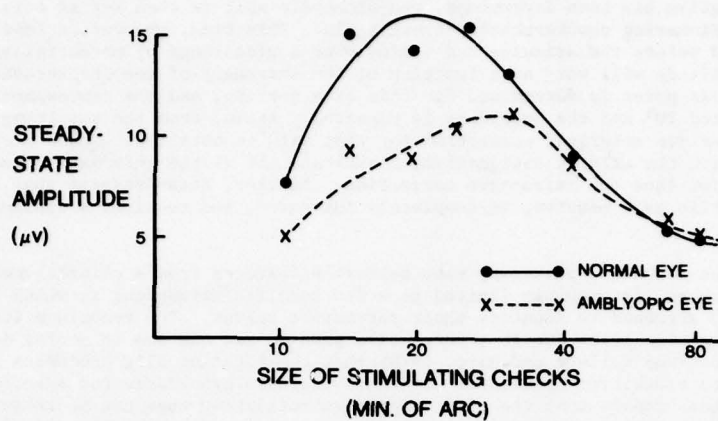
A more serious problem with this approach stems from the intrasubject variability in the transient visual evoked response (Callaway and Halliday, 1973). While the general morphology of waveforms will be relatively the same between subjects, it is not always a trivial task to identify the negative and positive peaks corresponding to those presented by Harter and White. For these reasons, the use of the transient visual evoked response to assess visual acuity in operational settings probably is extremely limited. Although it provides valid and objective index of acuity, correlating with more tediously determined behavioral thresholds, its use will probably be limited to clinically related areas. In such areas, the added precision given by this technique can have considerable impact. For instance, in the study cited above by Marg *et al.*, it has been established that visual acuity in the human infant matures to adult levels by 4 to 5 months of age. Determination of childhood visual acuity by psychophysical methods had led to the conclusion, still widely taught in clinical medicine, that acuity does not mature until 4 to 6 years of age. The increase in sensitivity produced by the use of the evoked response clearly has significant implications for clinical practice.

Steady-State Evoked Response Measures of Acuity. Increased applicability of these procedures can be obtained if the steady state evoked response (discussed on p.27) is used instead of the transient evoked response. In a typical case, a checkerboard pattern is counterphase flickered at a given frequency (preferably below 12 hertz) and the amplitude of the steady-state evoked response from an occipital/mastoid EEG derivation is determined. Regan (1977a) has shown that the amplitude of the response is directly related to the size of the stimulating checks. For normal adults, the highest amplitudes will be obtained with checks between 10 and 20 minutes of arc. This is true whether the evoked response is taken to pattern reversal, pattern appearance, or flashed pattern (Regan, 1977b). Further, the amplitude falls off quite regularly, as shown in the upper curve of Figure 14 (adapted from Regan, 1977c).

These observations lead to an accurate technique for determining visual acuity in situations where subjective response is difficult. For instance, in the lower curve of Figure 14, the evoked response versus check-size curve produced by an amblyopic eye is presented. In such conditions of reduced acuity, small checks do not produce as high an amplitude as they do in normals, although large checks are unaffected. The overall shape of the curve therefore reveals an acuity problem. Regan has suggested that this technique could be used to monitor the progress of occlusion therapy in an amblyopic child, or to determine the degree of acuity disruption in an adult. The procedure permits an estimate of the subject's acuity without requiring the difficult judgments normally needed for such determinations. However, it is still a relatively tedious procedure, requiring the subject to attend for some considerable period of time to the stimulus. Regan (1977a) reported a modification of this procedure which makes it more applicable to children and which could have considerable impact on the development of operationally useful techniques for measuring acuity. A TV-generated cartoon showing animated characters is presented to the child. Superimposed on the cartoon, a checkerboard pattern is counterphase-flickered. Since the cartoon does not have a consistent temporal frequency pattern, it does not significantly affect the steady-state response. Further, the contrast ratio of the checks can be quite low. Regan reports excellent results with this technique, and it can also be used with transient flashes of the checkerboard pattern to produce a transient evoked response. Amplitudes are interpreted in the same way as without the superimposed image. The obvious application of this approach stems from the ability to generate an estimate of the subject's acuity without intruding on an ongoing visual task. If children can watch cartoons, pilots can watch CRT displays while the evoked response is generated. Thus, this measure appears to be an almost totally non-obtrusive technique for assessing acuity in operational settings.

It is this non-obtrusiveness, along with precision, which makes the evoked response an attractive technique for human engineering applications. It could be used, either in steady-state or transient form, in many experimental and field applications. Although it would still take a considerable period of time to obtain a complete estimate of acuity, it could be done non-obtrusively, and without significantly impacting the subject's primary performance.

If it is important to obtain a refraction rapidly, a further modification of the above techniques has been developed by Regan (1973). This permits a complete refraction, including astigmatic and spherical determination, to be carried out within 5 minutes. This procedure is illustrated in Figure 15 with hypothetical data based on Regan's (1973) description. The subject is seated approximately 15 feet before a checkerboard pattern which is counterphase flickered at a given rate. Check size is maintained between 10 and 20 minutes of arc, the optimal size for steady state evoked response amplitude, with the



AFTER: REGAN, 1977c

Figure 14. Amplitude of steady-state evoked response as a function of check size for normal and amblyopic eye.

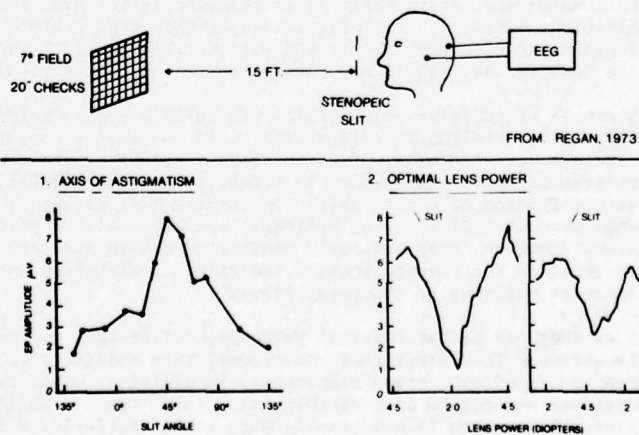


Figure 15. Rapid refraction using steady-state evoked response.

entire field subtending 7 degrees. The subject views the checkerboard pattern through a stenopeic slit which can be rotated through 180°. Using the fast Fourier Transform technique, the amplitude of the steady-state evoked response elicited by stimulation is calculated every few seconds. This amplitude is then plotted as a function of the average position of the slit. In individuals with astigmatic error, the angle of slit corresponding to their axis of astigmatism will produce the maximum amplitude of evoked response, with amplitudes falling off dramatically on each side, corresponding to the degree of astigmatic aberration. Once the axis of astigmatism has been determined, the stenopeic slit is then set at this axis angle, and the subject views the flickering checkerboard pattern again. This time, however, a lens of continuously changing power is placed before the stimulus and varied over a wide range of potential spherical corrections. Again, steady state amplitude will vary as a function of the sharpness of the checkerboard through the slit. Once the maximum amplitude point is determined for this axis setting, and the corresponding lens correction noted, the slit is rotated 90° and the procedure is repeated. Again, from the resulting amplitude versus diopter plot, the appropriate spherical correction for that axis is obtained. These 2 cylindrical corrections, combined with the axis of astigmatism, constitute all of the information necessary to determine the prescription lens for refractive correction. Further, Regan reports that this procedure can be carried out in as little as 5 minutes, is completely automated, and requires a minimum of cooperation from the subject.

Obviously, the above technique possesses many desirable features from a clinical point of view. Its operational utility, however, is probably limited to a few specific situations in which large numbers of subjects must be rapidly screened to identify their refractive status. The technique itself cannot distinguish between astigmatic error due to a refractive problem and one due to neural defect. Further, it has been reported (Bostrum, Keller, and Marg, 1978) that the rotating slit procedure has not proven to be as accurate as more exhaustive behavioral or evoked response procedures for specifying refractive error. These investigators report that the best corrective refraction that can be reliably achieved using this procedure is within about 1.0 diopter of the proper correction, due to variability. This is not as good as can be obtained with other methods, and is not clinically useful. With flashed checkerboard evoked responses, on the other hand, accuracy to .25 diopter has been reported. Even though this may be a valid criticism of the Regan procedure, the advantages of speed and objectivity make it a good candidate for screening in large groups, particularly in those clinical cases where one is frequently interested in determining whether or not there is a major deficit. Further, it may be that, for specific operationally meaningful research questions, the ability to perform such a rapid screening would make it easier to control for the refractive error of a subject in the investigation.

Color Vision

Adequate color vision is obviously related to a number of real-world tasks. In the extreme, normal color perception is necessary for differentiating traffic lights and other signals, for determining significance of flags, signs, and vehicles which are color coded, for identification of equipment handles and dials which differ only in color, and for a number of other tasks. (Sloan, 1946). In most cases, the types of differentiation required from the human are crude enough that it is only necessary to establish whether the individual is significantly deficient in color reception relative to the average individual.

Many tests are available to perform such crude determinations. Most rely on the use of cards or plates containing various colors. A digit or shape is outlined in one color. If the individual has normal color perception, the shape is able to be identified against the multi-colored background (e.g. the Ishihara Pseudo-Isochromatic plates, the Rabkin Polychromatic plates, or the American Optical Company Pseudo-Isochromatic Plates). Other tests use colored objects to be identified or classified by the subjects. These objects are carefully calibrated and retain their color characteristics over time. The individual is asked to make fine discriminations between the colors. Tests of this type include the Inter-Society Color Council (ISCC) single judgment test (Hardy, 1943), the Farnsworth-Munsell 100-Hue test (Farnsworth, 1943), the Peckham Color Vision Test (Sloan, 1946), and various tests using dyed yarn as the test objects.

More precise mechanical methods of studying color vision use anomaloscopes, colorimeters, and lanterns. These devices all present colored lights to the individual which, in one way or another, must be classified or matched. Colorimeters use mechanical mixtures of pure colors to present a given hue to the subject. Anomaloscopes present split-field views in which, typically, one half is constant and the other half is made up of a combination of colors which the subject must adjust to match the first half. These mechanical systems provide precision of stimulus delivery, and perhaps more precise quantification of small changes of color vision. However, they obviously require more time and care in conducting the experiment. Examples of these types of instruments include the Pickford-Nicholson anomaloscope (Holmberg, 1963) and the Four Color Replacement Colorimeter (Bongard, 1957).

These techniques, useful as they may be for clinical purposes, suffer the same problems as acuity measures with respect to applications. They are rather cumbersome, take considerable amounts of time to administer, and require a large number of subjective responses. In addition, small degrees of aberration in color reception by the individual may not be discoverable with these tests. While it is true that, from an aircraft design point of view, the concern is not usually with subtle differences in color vision, new display technology for both in-flight and ground crew use is increasingly utilizing color as an input channel. Air traffic controllers are being asked to discriminate aircraft symbols on the basis of color, and the use of color coding of information in the cockpit is being widely discussed. Therefore, it is no longer sufficient to simply determine whether an individual is "color blind". Increasingly, questions of subtle abilities to differentiate colors and, more importantly, the limits of color contributions to information processing are being discussed. These will require increasingly sophisticated techniques of study.

Several basic psychophysical approaches have been used to study these problems. These have confirmed the incredible complexity of the color receptors in the visual system, and have led to a correspondingly large amount of research dealing with basic color mechanisms (Walraven, 1972; Jacobs, 1976). Among the many significant developments in this area is a growing concern with the spatial and temporal characteristics of color vision. The contrast sensitivity of the component color mechanisms is beginning to be studied

with increasing interest. Results indicate that blue mechanisms show fairly low contrast sensitivity, with the peak centered at 1 to 2 cycles per degree, while the green mechanism appears to have a higher frequency peak than red, with both being more sensitive than blue. Such results are emphasizing that color vision must be studied in combination with spatial and temporal frequency considerations (Kelly, 1974).

Transient Visual Evoked Response to Color. Electrophysiological techniques for studying color vision, while not fully developed, are conceptually consistent with the recent psychophysical approaches. White and Eason (1966) were among the earliest investigators to attempt measurement of spectral sensitivity using the transient visual evoked response. Armington (1966) used the amplitude of the transient response as a criterion, and derived a sensitivity curve in general agreement with the CIE curve. Latency measures produced good agreement with CIE curves for both photopic and scotopic spectral sensitivity (Wooten, 1972). There is some ambiguity in the responses obtained under these conditions, however, and it was clear that improved techniques would be required if this measure was to become stable enough for laboratory and operational use.

White, Kataoka, and Martin (1977) among others (Krauskopf, 1973) have used the technique of chromatic adaptation to isolate a more "pure" color response. These investigators found that if an adapting field of a particular wavelength was used (the familiar Stiles technique), stimulation by a flash of another wavelength produced characteristic patterns in the evoked response. Using stimuli centered at about 450, 540, and above 680, and red, orange, yellow, blue, and blue-green background adapting stimuli, three sets of components were found. These are suggested to represent the three basic color processes. The "red" response has positive peaks at about 100 and 190 milliseconds. The "green" has peaks at about 120 and 200 milliseconds, and the "blue" has peaks at about 150 and 240 milliseconds. Interactions become very complex in this procedure. However, if it can be shown to reliably isolate the independent color mechanisms, it could be of significant value in laboratory settings. Its functional utility in more operational settings is harder to conceptualize.

Kinney and McKay (1974), however, discuss a color test based on another technique developed by White which may have more applicability. The original technique was designed to isolate a pattern evoked response from the response to an unpatterned whole-field presented at the same luminance. The evoked response obtained from one condition was subtracted from that obtained in the other condition. In the case of the pattern and whole-field comparison, if there were components added by the pattern, the two evoked responses would be identical, and the subtraction process would create a straight line. With proper controls, any remainder after the subtraction process constitutes the system's response to pattern. If, now, the pattern is composed of hue differences, where luminance levels of the hues were chosen to lie on the confusion line of color deficient individuals, the response by such an individual to presentation of the pattern would be the same as the response to a monochromatic whole-field. The subtraction process would then produce a straight line. In the normal individual, the pattern would be seen by the subject because of hue differences alone, and a subtraction process would produce a recognizable and repeatable waveform.

Kinney and McKay (1974; Kinney, McKay, Mensch and Luria, 1972) used several patterns composed of checks subtending 30 minutes of arc, with luminance combinations prepared specifically for protanopes, deuteranopes, and tritanopes. Results, as expected, revealed that normal subjects gave pattern responses even when the pattern was composed of hue differences only. The amplitude of the evoked responses was reduced, and latency increased, as hue contrast was reduced. On the other hand, color defective individuals showed a response only to luminance, and no pattern response to the hues in which they were deficient. The authors suggest that these results confirm that this technique can be used to detect color without verbal response from the subject. They recognize that, particularly with protanopic subjects, it would be possible to confuse a pattern response elicited by luminance differences with one elicited by hue, and they suggest that the luminance of one of the hues should be varied over a wide range. For color normals, such an adjustment would not eliminate the pattern response to hue, whereas for deficient individuals the response would disappear when luminance equality was achieved.

Steady-State Evoked Response to Color. If an unpatterned field of a given hue is flickered rapidly (between 45 and 60 Hz) the Fourier spectrum will reveal a "resonant" peak at the stimulating frequency. This is, of course, the steady state evoked response as described previously. Regan has noted that such "high frequency" flicker correlates with luminance, and that the red, green, and blue channels pool their responses linearly (Regan, 1970). If the unpatterned field is flickered a bit more slowly (between 13 and 25 Hz) it is found that the spectrum shows the expected peak at the stimulating frequency, and a second harmonic peak at twice the stimulating frequency. Again, this higher frequency harmonic (if it falls between 35-60 Hz) is quite sensitive to stimulus wavelength. Its amplitude can be used to measure the spectral sensitivity curve of the eye (Regan, 1975). The primary frequency, however, is not sensitive to photometric luminance. Thus, it appears that the steady-state evoked response to medium-high frequency chromatic stimulation contains two separate, distinct elements, perhaps representing different visual information travelling along parallel channels very early in the visual system.

If the steady-state response is generated by a colored patterned stimulus instead of an unpatterned field, different mechanisms appear to be involved. In this case, the pattern is defined by hue differences only. Thus, in a checkerboard pattern, the edges of each check would be defined simply by adjoining color areas (i.e., there are no lines or other designations of an edge). The intensity level of one set of checks is then systematically varied, producing a wide range of intensity ratios. That is, the contrast between adjacent checks is altered from high, to zero, to negative. Under these conditions, clearly defined pattern evoked responses are found (Regan and Sperling, 1971). Further, these have been shown to be due to the hue difference alone, and not to such potentially contaminating factors as chromatic aberration (Regan, 1971; 1973).

Such observations have led to development of a sensitive objective test for defective color vision (Figure 16). In the normal individual, as intensity of one set of checks is varied, the evoked response will be obtained over the entire range of possible intensity ratios. On the other hand, the color deficient individual will produce a steady-state evoked response only when the contrast information allows perception of checks. When contrast is effectively zero, the color deficient individual has no clue as to the location of edges, and consequently does not see any flicker at all. In such a case, there will

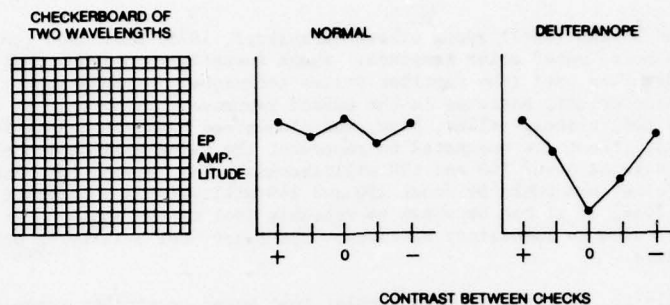


Figure 16. Objective test for color vision using steady-state evoked response.

not be an evoked response, or a very minimal one. The position of the minimum point gives the relative sensitivity to each of the colors used in the checkerboard.

If, as Regan suggests, this technique is used in a continuous mode, the determination of color sensitivity would be very rapid. Under this system, the fast Fourier transform would allow the determination of a data point in a matter of seconds. A whole range of contrast ratios could be run in one or two minutes, for a given pair of colors, and an entire survey of color deficiency could be run in 5 or 10 minutes. Presumably, all other things being equal, this would be an extremely sensitive index of color deficiency, and would enable the investigator to specify not only whether the individual was deficient, but the degree of deficiency and the "spread" in terms of range of contrast in which the subject gave a reduced evoked response.

For further discussion of spectral sensitivity determination using steady-state evoked response, see the section following on Critical Flicker Fusion (CFF).

Determination of Color Responses with Evoked Response Feedback. A creative and seminal demonstration of the power of the steady state evoked response has been demonstrated by Regan (1975). Combining extremely tight stimulus control with the speed given by Fourier analysis techniques, the entire procedure can be used in a "feedback" mode. The basic concept involves establishing an amplitude of steady-state evoked response for a given set of well-defined stimulus conditions. This amplitude is then monitored continuously. Alterations in specific environmental factors may produce a change in the amplitude of the steady state evoked response. It is entirely feasible to sense this alteration in the evoked response on a short term basis, and to make an adjustment in the display environment which would tend to restore the original amplitude of the steady-state evoked response.

In Regan's demonstration of this procedure, a 2 x 2 degree pattern of bright and dark checks of the same wavelength (676 or 544 nanometers) was counterphased flickered to generate the steady-state evoked response. Superimposed on this pattern was a 6 degree patch of desensitizing light whose intensity could be controlled by a neutral density wedge. The steady-state evoked response was calculated every few seconds. The amplitude of the response was used to drive the neutral density wedge controlling desensitizing light intensity. In this way, a given steady-state amplitude could be maintained. If the amplitude decreased, the desensitizing light was decreased to bring evoked response amplitude back up to previous levels.

This procedure also permits different stimuli to produce the same amplitude evoked response. In Regan's demonstration, a baseline amplitude was established to a red (676 nm) checkerboard and a yellow (590 nm) desensitizing light at particular intensities. After this baseline was stable, the wavelength of the desensitizing light was precipitously changed to 437 nm (blue). In response to this change, the wedge increased the desensitizing light intensity by 1.7 log units. Thus, a blue adapting light had to be 1.7 log

units more intense than a yellow light to maintain the same neurological effect to a red checkerboard. The entire spectral sensitivity curve (using a red checkerboard as a probe) was determined in this way, and it was found that sensitivity peaked at about 580-610 nm under these conditions, as opposed to the CIE curve peak at approximately 555 nm.

The full range of possibilities of this technique has apparently not yet been explored. It is clear from the precision necessary to produce the response, and the specificity of the steady-state technique, that it will not be a trivial task to apply this feedback mode in operational settings. Yet, it opens up so many possible applications that it would seem imperative that they be explored. With this technique, it is obviously possible to hope that an "optimal" level of neurological activation can be maintained in a subject. At very least, it should be possible to obtain neurological equivalence between two different kinds of stimuli. Unfortunately, it has not been demonstrated that this neurological equivalence is the same as subjective equality. Nor is it clear that a presumed "optimal" level of response will be correlated with the subject's optimal level of performance capability. It is known, for example, that objectively determined optimal levels of display intensity do not always prove the most comfortable or subjectively pleasing levels for the individual. For this reason, intensity, sharpness, or color levels of a display which yield optimum performance for each individual may have to be determined before anything can be done with the evoked response.

Further questions regarding this methodology should be explored to determine its applicability to applied problems. One deals with the stability of an "optimum" point over time. If a given level of color response in the visual system does not always produce the same maximum amplitude of evoked response, the technique and other intra-individual variations must also be considered. If one were interested in using this feedback technique to monitor or maintain performance over a long period, it would be necessary to consider habituation, adaptation, fatigue, etc., as factors which might alter the optimum criterion level of evoked response amplitude. A final problem associated with this technique deals with the fact that steady-state evoked response amplitude is determined by many sensory characteristics, such as acuity, color, intensity, temporal frequency, etc. A change in steady state amplitude therefore could not necessarily be related to changes in one stimulus parameter unless the environment has been carefully controlled. Although this will obviously be possible in many laboratories and even in some operational settings, it will not always be the case. Further research should be directed to determining whether there are unique characteristics to the changes which occur in the steady-state amplitude from each of these various sources. In the meantime, and in the absence of extremely controlled environments, changes in steady-state amplitude would not yield a great deal of information with respect to a complex real-world environment.

In spite of these difficulties, the feedback technique based on the steady-state evoked response provides an exciting prospect for evaluating sensory function in a rapid, precise, objective way. It provides a totally non-invasive and non-subjective way to obtain a sensory "point of equality". Like other EP techniques, it measures not only refractive error and other physiological factors, but also the environmental factors which may be influencing visual acuity, color reception, contrast, etc. As such, it should find wide application in the laboratory efforts to define "optimal" or at least stable sets of conditions for the design of display and other information presenting techniques. This one contribution should stimulate a great deal of research directed to applications. In no sense is it ready to be used over a broad range of applied questions at the present time, but if only a fraction of the possibilities it raises should become feasible, it could significantly impact aircraft human factors methodology.

Critical Flicker Fusion (CFF)

An intermittently flashing light stimulus produces the sensation of flicker if the frequency of flash is low enough. As the frequency of the flash is increased, the point is reached at which the individual will cease to perceive the light as flashing, and will begin to see it as a steadily burning light. The frequency of flicker or flash which is required in order to see the light as steadily burning is called the Critical Flicker Frequency (CFF) or sometimes the Flicker Fusion Frequency (FFF). Study of this phenomenon has a long history (Landis, 1953; Pieron, 1965; Sokel and Riggs, 1971). The CFF value for any individual will vary depending upon a number of subjective and objective factors. These include the intensity of the light, the area of the retina being stimulated, the position of the retina being stimulated, the duty cycle of the light and dark ratio, the wavelength of the light, and a number of other factors (Landis, 1954). Subjective factors such as fatigue may also influence CFF (Webar, Jermini and Grandjean, 1975). However, if the objective or subjective factors are well controlled, the CFF value is a very stable measure. Reported variations within a subject range from .6 to 2.9 percent. Generally, the eye is more sensitive to flicker 10° to 30° in the periphery than it is in the fovea. The more intense and larger the source, the lower the flicker threshold. Subject-to-subject variability is quite large. However, there appears to be only a very slight learning curve in the task, and most subjects produce stable thresholds after a few trials.

Flicker fusion thresholds have been obtained in a large number of stress situations, and the reader is referred to the reviews noted above for the complete list. However, representative examples will be given to indicate the range of stressors under which this measure has been taken. CFF, as would be expected, has frequently shown changes in conditions of anoxia. Scow (1950) found decrements at 18,000 feet for one hour. O'Donnell, Chikos, and Theodore (1971) studied CFF in humans under carbon monoxide exposure. Generally, decrements are not found until the CO-induced oxygen deprivation reaches a point equivalent to an altitude well above ten thousand feet. Fatigue, while apparently capable of disrupting CFF, does not do so readily. Tyler (1947) reported no CFF change in subjects who remained awake from 30 to 60 hours. Similarly, subjects doing prolonged visual work (3.5 hours of reading) did not do consistently worse in CFF (Ryan, Bitterman, and Cottrell, 1953). A large number of studies investigating the effects of drugs and other chemical agents have used CFF as a measure (Misiak, Zenhausen, and Salafia, 1966; Misiak and Rzy, 1968). Keighley, Clark, and Drury (1951) used CFF to evaluate the effects of positive acceleration on vision and found no significant change between 2.5 and 3.2 +Gz. When acceleration was increased to +4.8 Gz, a statistically significant change in CFF was found.

With respect to psychological stress, the US Dept of the Army (1963) used CFF to test subjects undergoing their first parachute jump. These individuals showed significant decrements. Subjects exposed

to 90 db of noise while taking the CFF test showed stress effects. Pulsed auditory noise was unable to cause flicker when it was absent, but when flicker was already present, it became more pronounced with auditory input (Knox, 1953; See also Miller, 1969). Ambient temperature may also affect the sensitivity to an intermittent stimulus (Lockhart, 1971).

The fact that adaptation to a flickering stimulus influences the threshold for the discrimination of flicker (Brown, 1973) raises the possibility that there may be specific receptor channels for the reception of flicker in different frequency regions. Although this is not a necessary conclusion from the studies (Smith, 1970; 1971) it would have significant and basic operational meaning if such channels could be established. Regan (1977a) has provided evidence for the existence of bands of sensitivity to flickering light. If a light is flickered at several frequencies between 3 and 12 hz, a curve similar to the first, highest amplitude curve in Figure 17 is found. This shows peak sensitivity near 10 hz. If the stimulating frequency is further increased, another "tuning" function emerges between 13 and 25 hz, and still another between 40 and 60 hz. These three ranges, at least, produce steady-state evoked responses which appear to come from different parts of the cortex, have different color properties, and have different relationships to intensity. Further, these responses to an unpatterned flicker show different tuning characteristics than the response to stimulation by a checkerboard of high spatial frequency.

It is fascinating to note that these steady-state responses are being recorded at flicker frequencies well above the subjective CFF point. In fact, it has been shown that these responses do not correlate with perceived flicker at all (Regan and Beverley, 1973; Spekreijse, 1966). Increasing the flicker rate of a stimulus can abolish the subjective perception of flicker, but actually enhances the evoked response.

These findings have considerable theoretical significance and, under the proper circumstances, could have massive implications for applied areas of research. Basically, they imply that separate populations of cortical cells, and separate perceptual functions, can be measured by stimuli which are essentially imperceptible to the human. As Regan notes "Whole experiments can easily be carried out with stimuli so weak that the subject never sees them" (Regan, 1975). This brings the whole area of flicker into the realm of a totally non-obtrusive, non-invasive measure which could be used in an operational environment in many ways. The evoked response could be generated from the flickering of aircraft instruments, surround lights, or even ambient lighting. To explore these possibilities, a number of pilot projects are currently underway to define the conditions under which these high frequency responses occur, their operational and theoretical significance, and their limitations (Moise, 1978; Wilson and O'Donnell, 1978).

At least one parameter of high-temporal frequency evoked responses has been well-defined. It was noted earlier that stimulation with an unpatterned light at relatively high frequencies (e.g., 24 hz) produces an evoked response which shows a peak at the stimulating frequency (24 hz) and another at the

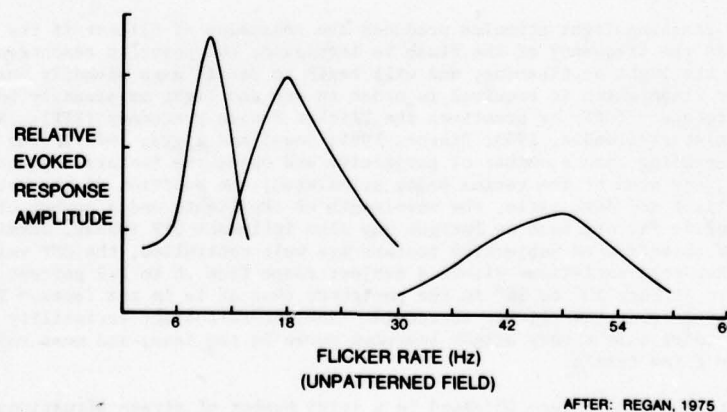


Figure 17. Peak sensitivities for steady-state evoked response amplitude as a function of flicker rate.

first harmonic (48 Hz). If this procedure is carried out by using a colored flash alternated with a white flash, the procedure approximates the psychophysical technique of heterochromatic flicker photometry. If the brightness of the colored flash is gradually altered, subjective flicker will eventually disappear. Regan (1975; 1977a) has shown that the high frequency harmonic of the steady-state evoked response produced in this way measures the spectral sensitivity of the eye, while the primary frequency does not correlate with photometric luminance. This is true even though the two are part of the same waveform. The medium frequency primary peak does not show a spectral efficiency curve, is sensitive to chromatic adaptation, and may ultimately be shown to reflect the activities of opponent color mechanisms. Again, this provides a tantalizing example of the extreme specificity of measurement possible with the evoked response, and raises many questions which can easily be subjected to controlled investigation.

Obviously, the reliability and validity of these observations must be established in other laboratories. Assuming this will be done, it is necessary then to further explore the meaning of various components, harmonics, and morphologies of the responses in experimental settings. The reliability of techniques for obtaining these micro-volt signals in operational environments must be established. Finally, the measures must be related to meaningful, real-world events. There will be problems and obstacles at each of these points. However, again, if even a small part of the possibilities offered by these measurement approaches prove valid and feasible, the implications for basic theory in vision would be considerable. The implications for measurement in applied fields would be spectacular.

AUDITORY INPUT

In terms of basic research interests, auditory input has been studied almost as extensively as visual input. However, from an applied point of view, the auditory modality has not received nearly as much attention. This may be due to the fact that it is somewhat more common to design auditory inputs in such a way as to avoid many of the problems of absolute and differential thresholds which are encountered naturally in the field of vision. Paradoxically, since language has received so much attention from a clinical point of view, there has been no lack of research interest in absolute and differential auditory thresholds. Methods of testing these thresholds over a wide range of frequencies, appropriate to many kinds of clinical problems are well established. Surveys of auditory transmission, (See Davis, 1968; Harris, 1972) continue to elaborate the physiological and psychophysical mechanisms underlying auditory sensation (see Licklider, 1951; Licklider and Miller, 1951).

In spite of the widespread clinical study of auditory thresholds, determination of real thresholds with any degree of precision is a very delicate procedure requiring apparatus and environmental controls not normally accessible to most applied researchers. The problems basically arise from the fact that achieving absolute quiet for determination of auditory thresholds is extremely difficult. Obvious solutions, such as the use of earphones or other localized sound attenuators, are not completely satisfactory. Additionally, frequency, the duration, and shape of the tones presented to the subject interact in peculiar ways. Therefore, there is not one absolute threshold. Depending on the complex of stimulus factors used, there may be many thresholds. Auditory thresholds are also extremely variable. Licklider (1951) reports studies showing that the day to day variability in a subject may be 5 times the average variability on a given day. In addition, one subject can show as much as 5 decibels (db) variation from one-half minute to the next. The particular psychophysical technique used (for example free versus forced-choice) can introduce differences in the obtained threshold. Brown (1965) recommends an augmented "receiver operating characteristics" (ROC) technique which, he believes, provides a more comprehensive measure of overall performance.

The above difficulties in obtaining a true psychophysical threshold for auditory sensitivity have resulted in this measure being used relatively infrequently in performance and human engineering studies, except where there were questions of direct insult to the auditory system. In such cases, elaborate electronic and environmental equipment must be available to the researcher, and this puts the testing of auditory function into the realm of the specialist. Recognizing these problems, however, it is sometimes adequate to obtain relatively crude estimates of the absolute sensitivity of the individual in order to ensure that the person will hear certain warning signals or communications. In terms of medical evaluation of pilots and other crew members, this can be particularly important. More important, new aircraft systems are imposing greater and greater auditory loads on the individual. Electronic warning devices, auditorily coded, are being used in cockpits with increasing frequency. These are superimposed on other communications channels. Due to these factors, the overall auditory load on the pilot is rapidly reaching the saturation point. Consequently, questions involving the advisability of adding more auditory signals are causing greater concern among system designers. To answer such questions, more sophistication in auditory testing is clearly desirable. In most cases, these questions will more properly be considered in later sections of this AGARDograph, under the heading of cognitive function. However, in the present section, some of the very basic psychophysiological techniques for assessing the absolute threshold of auditory sensitivity will be discussed. Other questions of auditory function, such as pitch discrimination, differential thresholds, etc., will be incorporated into later sections.

Measures of Absolute Auditory Threshold

Psychophysical Techniques. Recognizing the limitations discussed above, several methods have been developed to measure the approximate absolute threshold of auditory functions over a range of frequencies. The most common technique, still used clinically, was proposed by Bekesy (1947). In this technique, the subject controls the intensity of a continuous tone. As long as a button is depressed, the intensity increases. When the button is released, intensity decreases. The subject is instructed to hold the button down until the tone is heard, and then to release it until the tone can no longer be heard. A direct-writing recorder shows the intensity of the threshold tone over different frequencies, and the resulting curve represents the subject "tracking" the threshold. Five types of Bekesy patterns have been identified (Rintelmann and Harford, 1967) which attempt to localize lesions to the middle ear, cochlea, or eighth nerve, and to detect simulated hearing loss.

The other major psychophysical technique for determining absolute threshold uses discrete, pulsed tones, either presented repetitively at gradually increasing or decreasing intensity levels, or presented individually. In using pulsed tones, there are some problems introduced by the unwanted transients produced as the tones go on or off, and these must be controlled. A variation of the pulsed tone technique has been described by Reger and Voots (1957). In this, discrete tones are presented in coded form (either one, two, or three pulses) which may appear in one of three time positions. In a forced-choice response situation, the subject reports on the nature of the stimulus by pressing pre-designated response keys. This response is automatically compared to the stimulus, and future stimulus presentations are adapted to the subject's past success or failure. Testing time is reported to be about 20 minutes.

Psychophysiological Techniques. Psychophysiological techniques which have been developed to look at auditory thresholds arose predominately out of the clinical need to test infants, children, the retarded and senile. Anyone who cannot (or will not) give an adequate response to indicate when tones are heard presents a considerable problem to the clinician and the clinical researcher. Behavioral techniques, involving conditioning and other indirect procedures were developed and are used widely (Northern and Downs, 1974). However, these are time consuming procedures which are not always successful. Consequently, considerable interest developed in evolving new procedures which would require even less cooperation from the subject.

In response to this need, several objective hearing tests have been developed for clinical use. Impedance audiometry involves sealing off the external auditory canal with a probe. A tone is then introduced into the canal and a pick-up microphone is used to quantify the sound pressure level of acoustic energy reflected back into the auditory canal by the tympanic membrane. Three specific tests are usually used to determine the integrity of the middle ear and the compliance of the tympanic membrane. Although these tests provide a great deal of information, and are reported to be very reliable, they require specific training in their use, and should be limited to the clinician with expertise in their administration and interpretation. As such, they are not easily used in most applied situations.

A similar situation exists for the procedure termed electrocochleography. This attempts to record electrical activity probably generated in the middle ear itself to a tone or click presentation. It has proven to be an extremely small response which is best recorded from an electrode surgically placed in the round window niche. However, through time-locked averaging, the small signal can be enhanced and picked up from an electrode in the canal, or even on the pinna. Again, this is not a procedure which the non-specialist researcher can easily handle, and is not generally available for any applied purposes.

A number of other procedures are based on the observation that an infant (or adult) will "attend" to an auditory stimulus with a variety of autonomic responses. In the case of the infant, these include widening of the eyes, stiffening, changes in respiration, heart rate, and activity level. Such observations led to the development of a series of techniques, some simply observational, for utilizing electrophysiological measurement to determine when an infant or other individual could hear a stimulus.

One of the earliest of these techniques used the Galvanic Skin Response (GSR) to measure perception of the auditory stimulus. In audiology, this technique may be called electrodermal response audiometry (EDA, EDR) galvanic audiometry (GA) or psychogalvanic response audiometry (PGSR). The procedure, first used in the late 1940s, involves first conditioning the child to a tone and an electrical shock. The electrical shock elicits a change in skin resistance (GSR) as the unconditioned response. If hearing is present, the tone soon becomes associated with the shock, and the tone alone elicits a GSR. If the conditioning is done with high intensity tones, then threshold values can be obtained from tones of much lower intensity. Early enthusiasm for the EDA technique among audiologists was short-lived. It soon became clear that the conditioning procedure, being aversive, was frequently traumatic to the child, and most clinicians are now reluctant to use this technique. Further, infants show frequent phasic GSR changes, and it was difficult to separate true responses from these random occurrences. Finally, one cannot ignore the fact that a small current is being introduced to the patient if the Feré technique is used. No matter how trivial an amount this is, it must be considered when dealing with infants. Nevertheless, this measure was the first electrophysiological technique which attempted to measure hearing but did not attempt to use the ear itself to produce the signal. As such, it stimulated considerable interest in the field.

Shortly after this period, it was noted that if a human hears a moderately loud tone, a transient alteration in heart rhythm will be produced (Zeaman and Wegner, 1956). The direction which the change takes is variable between subjects, and is very much influenced by the cycle of respiration, tending to accelerate during inspiration and decelerate during exhalation. However, the occurrence of a change in one direction or another was relatively easy to detect, and could be reliably scored. A large number of clinical trials were carried out on humans (usually infants or babies) and it was established that the response could be recorded in all but possibly some retarded children, that the acceleration response was higher in four-day infants than in younger ones, and that there appeared to be habituation (Northern and Downs, 1974). Further, response results agreed closely with other tests of auditory function, and appeared to be a clinically feasible procedure.

Difficulties in using this procedure stem mainly from controlling the state of the subject. The response appears related to alertness, and there is some disagreement as to the desirability of sedating the child for this test. Without sedation, many of the subjects most in need of physiological testing (hyperactive, autistic, etc.) cannot be tested. With infants, it is frequently difficult to tell exactly which state of sleep or waking they are in. Yet, it is important to know this if results are to be interpreted. Because of these difficulties, audiometry based on heart rate responses has not become a common technique. It appears to be well respected, and is often recommended in difficult cases. However, it is reserved as a specialized procedure. Similarly, in applied settings it would appear to have possible but limited value.

The above techniques all depend upon autonomic responsivity. Therefore, to some extent they are limited specifically to those circumstances where perception of a tone will cause an "alerting" or other surprise response to the individual. In the case of an adult subject, particularly in any applied

situation, it is unlikely that a simple threshold tone will provide an easily measured autonomic response. Thus, it is unlikely that the above techniques would have great value in operational settings, at least for threshold determinations. On the other hand, the use of central nervous system indicators of perception would not suffer from such a disadvantage. For this reason, evoked response audiometry has developed into a reasonably large and well-utilized area (Davis, 1976).

The observation that the EEG showed changes when a sound was heard was made in 1939 by Davis. However, until the averaging computer came into use in the 1960s, results of attempts to apply this observation were disappointing. With more sophistication in procedure and electronics, it became possible to describe idealized forms of the auditory response, and several different types of response emerged (Figure 18).

The first response studied, and still the one most familiar to physicians and audiologists, is the late auditory response, sometimes called the slow response. This is produced in response to a stimulus with relatively brief rise time, and consists of a small positive peak at 50 msec, a negative peak at about 90 msec, another positive peak at about 180 msec, and the largest peak going negative at about 250 msec. These peaks are generated in the cortex, probably in the primary projection area and immediately surrounding secondary projection areas. Thus, it assesses the integrity of the entire auditory system. It is best recorded at the Vertex, although it can be recorded over the primary projection areas if hemispheric information is desired. Skinner and Antinoro (1969) found that signal parameters have a significant effect on the amplitude of the auditory evoked response latent components. As signal rise time decreased, the peak amplitude in the evoked response increased. Increasing the stimulus frequency from 250 to 8,000 Hz produced a consistent decrease in the peak to peak amplitude of the evoked response.

This response is extremely reliable if the proper testing conditions are maintained. Many investigators have used the technique to objectively assess auditory acuity (Davis, Hirsh, Shelnut, and Bowers, 1967; Rapin and Schimmel, 1977; and Suzuki, 1969). Generally, the success rate is very high, with more false negatives reported than false positives. However, there are still considerable problems with the technique. Most importantly, the amplitude and morphology of the evoked response is significantly affected by the existing state of alertness of the individual. Osterhammel, Davis, *et al* (1973), obtained transient evoked responses while simultaneously recording the level of sleep. There was a systematic relationship between the changes in evoked response and the sleep stage of the subject. Thus, like the autonomic procedures discussed above, this technique suffers from dependence on the state of alertness in the individual, and this limits its use to very controlled situations. Davis (1977) has proposed that the auditory evoked response recorded during sleep may, in fact, measure a different physiological phenomenon than it does in the waking state. Whatever the final determination on this stimulating hypothesis, it is clear that the slow or late components of the response are rather difficult to assess. For this reason, they too will probably find limited operational use.

A faster response has been described by Davis (1977). This "middle" response shows peaks between 12 and 50 msec, with a positive peak at about 35 msec being most robust. It is now known that this response is frequently contaminated by a myogenic component (the sonomotor response of Bickford (1967) shown in Figure 18). For this reason, it is a difficult response to obtain, and has not come into general use.

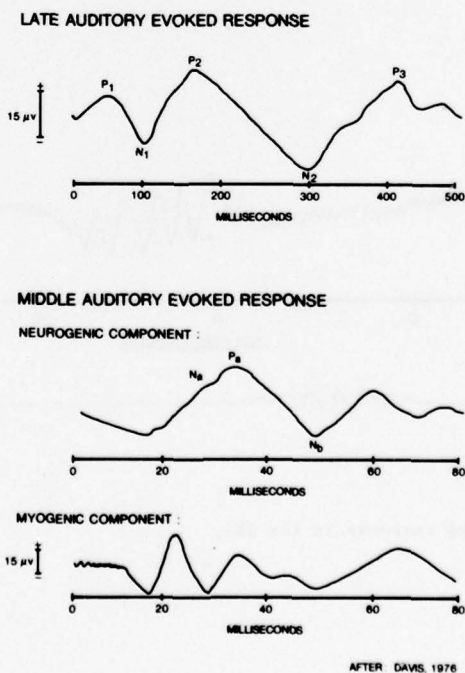


Figure 18. Late and middle auditory evoked responses.

Still, Davis believes these peaks represent the earliest cortical responses to auditory volleys, and are generated in auditory projection areas as well as in the medial geniculate.

Both of these "transient" evoked responses suffer from the difficulty that they are state dependent, since they are generated in or near the cortex. Therefore, any condition affecting the cortex will alter the response. While this may be of some value in determining whether or not there is a cortical effect from a given operational stressor, it may not be the most direct way to determine such an effect. In addition, the interaction between cortical state and the transient auditory evoked response makes determination of auditory threshold rather tenuous using these techniques.

With respect to auditory transmission itself, two relatively new techniques provide a more objective and reliable procedure. The first (Moushegian, 1977; Moushegian, Rupert, and Stillman, 1973) is based on the fact that the auditory system will "follow" the frequency of certain tones. When recorded from scalp electrodes, pure tone stimulation below 1.0 kHz produces a sine-wave like response at the stimulating frequency (Figure 19). This response has a latency of about 6 to 9 msec to short duration tones, and shows remarkable consistency between and within subjects. Although it appears in the record in rudimentary form at about 20 db above threshold, it is not clearly defined until about 40 db above threshold. The origin of this response is not well established, but may be in the medulla and midbrain.

To obtain this response, the subject is presented either with brief bursts of pure tone stimulation, or with a continuous pure tone at a given intensity. Averaging of the signal is carried out for an extremely brief period of time (50 msec for example) locked to the phase of the stimulus. This allows the very small brain signal which is at the same frequency as the input signal to be isolated from the noise. Thus, as illustrated in Figure 19, a "steady state" type of response is obtained in which the output of the brain is obtained at the same frequency as the auditory input. Many lines of evidence indicate that this frequency following response (FFR), sometimes called the frequency following potential (FFP), is a measure of the integrity of certain brainstem structures, and the peripheral auditory apparatus sensitive to tones below 1000 Hz (Galambos and Hecox, 1977). Therefore, this response can easily be used to test the gross hearing of a subject below this frequency.

This response however, is not without difficulties. Although it is not cortically dependent, and therefore is relatively independent of the state of the individual, it is a very small response, frequently measuring in the range below one microvolt. It therefore must be obtained with extreme caution. The signal is so small, in fact, that mechanical artifacts and electromechanical pickups from the stimulus generating equipment might mimic the response, leading to a false interpretation. It has been suggested (Moushegian, 1977) that in order to overcome this possibility, the pulsed tone technique be used rather than continuous tones. Since there is a lag of 6 to 9 milliseconds between the tone presentation and the appearance of the FFR response, this technique allows the investigator to be certain that the response being seen is from the brain rather than artifactual.

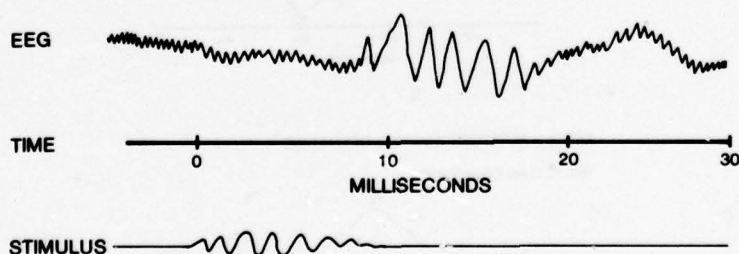


Figure 19. Frequency following response in the EEG.

Campbell, Atkinson, Francis, and Green (1977) have used a variant of this response to detect auditory thresholds. Adapting their procedure for obtaining visual thresholds, these authors presented stimulus tones at the rate of 6 to 32 per second. When fairly large numbers of responses were averaged, reliable power at the first or second harmonic of the repetition rate could be measured. When the second harmonic amplitude was plotted as a function of the stimulus intensity in decibels, a linear relationship was found at lower intensities. Figure 20 illustrates this result. Extrapolation of the curve to the theoretical zero evoked response amplitude produced good agreement with behavioral thresholds. The authors warn that enough data points must be taken to assure the validity of extrapolation, since an asymptote does occur. However, they believe a 10 minute test would allow screening of a patient's threshold to within ± 5 dB. Overall, it would appear that the FFR technique could provide a valuable procedure for measuring auditory threshold below 1000 Hz in applications of interest to the human engineer. This would be true particularly in cases where no cortical decrements would be expected, such as in determining the effects of noise stress and other peripheral auditory insults on the individual.

Obviously, extreme caution and relatively sophisticated design of experiments must be coupled with the technological sophistication necessary to obtain such a response. It will not be possible to utilize such a technique in a large number of field environments. However, for controlled laboratory experiments, the procedure has much to recommend it in terms of reliability and specificity of measurement. A similar and, in some ways, more reliable and well specified procedure has been described for the auditory system. Jewett and others (Jewett, Roman, and Williston, 1970; Jewett and Williston, 1971; Davis, 1976) have described a click-evoked response recorded from the vertex of the human scalp in the first ten milliseconds after stimulation. Unfiltered click stimuli are delivered to the ear rapidly, from 5 to 60 times per second. The brain response to these clicks is averaged for 10 milliseconds, and a large number (usually over 1,000) stimuli are delivered. The normal response obtained from an adult (Galambos and Hecox, 1977) consists of 5 to 7 distinct peaks with well defined latencies. These are illustrated in Figure 21. Jewett and others have determined that these peaks originate in specific peripheral and mid-brain structures, as identified in the figure. Starr and Achor (1975) specified latencies for each of these peaks. Typical nominal latencies at 65 dB for an adult are 5.5 msec for the V peak (inferior colliculus), 3.8 msec for the III peak (olivary region) and 1.6 msec for the I peak (eighth nerve). The variability of these measures between individuals is ± 2 msec. With decreasing intensity of the click, the latencies of all peaks increase and amplitudes decrease. For example, the V peak latency is 8.1 msec at 5 dB for a normal hearing adult. Finally, at or near threshold, the peaks cannot be isolated. It is possible, therefore, to determine the peripheral and midbrain auditory threshold, and this is usually within ± 10 dB of the behaviorally determined threshold (Davis, 1976). Further, from the shape of the intensity vs latency curve, it is possible to estimate whether a hearing loss is conductive or sensorineural (Galambos and Hecox, 1977). This "brain stem response" (BSR) is probably maximally sensitive only to frequencies above 2,000 Hz (Davis, 1976). Thus, it is theoretically possible that the individual will show no BSR response, or a degraded response, and still show normal threshold hearing values for the speech range. For this reason,

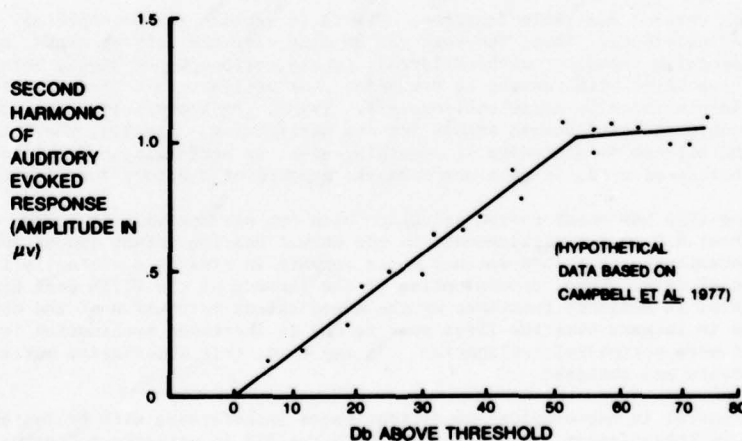


Figure 20. Use of auditory evoked response to rapidly presented tones in assessing auditory thresholds.

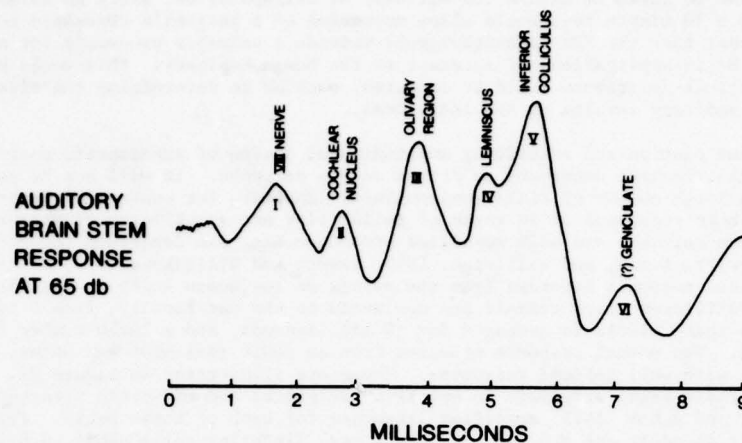


Figure 21. The auditory brain stem evoked response (BSR) to click stimuli, showing probable source of major peaks (positive up).

a complete auditory test should include both the BSR and the FFR, thus covering all frequency ranges. In most applied situations, however, the BSR alone is probably sufficient.

The BSR possesses several desirable features. First, it appears to be completely independent of the cortical state of the individual. Thus, the test can be done with the subject awake, asleep, or even sedated. This is especially important with children. Second, although the signal being recorded is extremely small, it is so well specified with respect to frequency that artifact rejection techniques can be used to isolate it, even in electrically noisy environments. Third, the stability of the signal (note the variability between subjects noted above) argues for its utilization. Finally, the ability to obtain this response when the subject is attending to something else, or performing a completely detached task introduces the possibility of using it as a nonobtrusive measure of auditory function.

The BSR technique also has other characteristics which are attractive. At birth, the latency of the V peak ranges from about 8.0 to 8.6 milliseconds in the normal hearing infant (Hecox and Galambos, 1974). These authors have established that this latency moves forward in time in a virtually linear way for the first 12 to 18 months of life. Thus, determination of the latency of the fifth peak during the first year can reveal abnormalities in auditory functions or the neurological maturation of the child. It is uncertain whether this decrease in latency over the first year is due to increased myelination in the auditory pathway, or whether there is a more peripheral explanation. In any case, this observation makes the BSR a valuable diagnostic aid in infants and children.

The BSR is also useful in neurological conditions where interference with brainstem transmission would be expected. Delays in transmission have been found with the BSR in patients suffering from acoustic neuromas, astrocytomas, diffuse infiltration of the brain stem, and other focal lesions (Starr and Achor, 1975; Starr, 1977). With such lesions, one sees intact early peaks, with later peaks disrupted or absent. While this is obviously not the major interest in most applied human engineering contexts, it is significant that the BSR is sensitive enough to provide information with such precision and reliability. The human engineer has seldom had such a precise technique available, and it will be interesting to see if this procedure will be able to be utilized in meaningful applied studies, or even in the field environment.

One such application has already been suggested. It has been found that the latency of the V peak of the BSR increases dramatically with ingestion of alcohol. Further, the authors report that this latency increase shows a higher correlation with subjective intoxication levels than the correlation between intoxication and blood alcohol. Although this is a quite preliminary report, the implications of such a result would be significant. It would raise the prospect of a non-obtrusive, rapid, reliable test for particular kinds of drug use and for other psychological decrements in the individual. Further, it is possible that such a reliable measure of neurological integrity could assist in the routine medical evaluation of pilots and other critical personnel, as well as isolating day-to-day alterations in neurological stability.

It cannot be overemphasized that much more research is needed before such applications can be considered valid. In general, however, the point should not be minimized that here is a physiological measure, showing psychological concomitants, which is extremely stable and is precisely localized neuro-anatomically. Whether the BSR itself ever proves to have significant applications to human engineering problems, it represents a new class of measurement technique which the psychophysicologist will use in applied situation. These are specifically defined, reliable, show small within- and between- subject variability, and can be obtained with a minimum of intrusion into the natural operational environment. They represent, in many ways, the first true microscope the human engineer has ever had.

COGNITIVE FUNCTION

The assessment of cognitive function is, in some ways, assuming greater importance in system design than assessment of sensory function. As pointed out in the first section of this AGARDograph, it was possible, until recently, to treat the human operator essentially as an error nulling subsystem. The pilot could be adequately described by a series of differential equations accounting for closed-loop variability in this subsystem. However, the advent of digital avionics, with fly-by-wire capabilities, can change all of this. It is possible (though perhaps not likely in the near future) that the pilot will be able to fly an entire mission, from engine start to shutdown, without producing a single error nulling input (Krippner and Fenwick, 1975). This would happen through preprogrammed digital flight commands. The operator in this system would become a monitor, processing system information, assessing the desirability of updates relative to preprogrammed maneuvers, and deciding when and how alterations in the preprogram maneuver should be introduced (O'Donnell, 1975).

The net effect of these changes is to place much more emphasis on the importance of open-loop evaluation and decision processes by the operator. This, in turn, results in significant increments in memory load and information queuing requirements. In approaching system design, therefore, techniques must be developed to answer questions about such difficult-to-measure concepts as decision making, attention, vigilance, mental fatigue, motivation, and workload. This problem is magnified when one considers the requirements for multi-operator systems (Command and Control Aircraft, communication systems, Remotely Piloted Vehicles, etc.). In these cases, it is necessary not only to assess the effect of the given system requirement on a principal operator, but to consider the effects of degraded cognitive performance in any member of the team on all other operators in the system. Questions of memory load and information transfer become incredibly complex. Existing techniques for assessing cognitive function in such multi-man systems are correspondingly complex and cumbersome. Computer simulations are seen as the only manageable way to handle the massive number of variables and interactions possible in these multi-person systems. However, even computer techniques require high quality input data which must be determined empirically. Such data are difficult to acquire at best, and most existing behavioral data sources have proven inadequate in measuring cognitive variables.

One of the major difficulties in obtaining good estimates of cognitive function in applied environments stems from the problem of taking such measurements without interfering in the very process of interest. This methodological paradox is well known in the physical sciences. It is true in the assessment of sensory function. However, to an even greater extent, cognitive performance is probably based on a fluctuating, highly adaptable set of human capabilities. Most attempts to measure such performance have used synthetic tasks, frequently superimposed on a primary task. In many cases, these are quite ingenious. For instance, Sternberg (1969; 1975) has adapted a choice reaction time paradigm to the recall of information in short- or long-term memory. By controlling the amount of information to be stored, it is possible to plot reaction time as a function of memory load, and to obtain independent estimates of sensory input, motor output, and central processing times. This task has been used successfully by itself to measure cognitive effects of various changes in stimulus parameters (Briggs, *et al*, 1972). It has also been used as a secondary task to assess workload changes in a simulated flying task (Spicuzza, *et al*, 1974). However, in each of these types of application, there is still the problem of task interference and extrapolation to real-world environments. No matter how used, these tasks require that the experimenter set up a somewhat artificial situation. Either the subject is doing a synthetic task (the reaction time task alone) or is required to divide attention between a primary and a secondary task. Although these designs have yielded a great deal of valuable information, the problems associated with such behavioral testing are well recognized, and in the final analysis probably impose an unacceptably low limit on the amount that can be learned about cognitive function (Chiles, 1978; O'Donnell, 1975).

What is required if one is to obtain a non-obtrusive measure of cognitive function are tests which would monitor the subject's system without requiring any additional attention, workload, or modifications of normal behavior. Such an approach is, of course, possible within the system itself. The behavioral output of the primary task can sometimes be used as an index of cognitive function. However, in most applications, either the behavioral output does not lend itself to such an analysis, or the task is capable of solution by a variety of approaches, and the adaptable human will select different approaches at different times. Primary task measures, then, do not constitute an adequate answer to the need for measures of cognitive function.

It is proposed here that psychophysiological measurement techniques are beginning to address this question with some success. While no complete answer has been yet forthcoming, useful techniques have already been developed, and these indicate that further developments are likely. The following section will present the more recent attempts to assess major areas involving cognitive functions. These areas have been somewhat arbitrarily chosen, based on one of many possible classification schemes. Essentially, large categories of performance which, to varying degrees, depend on or affect cognitive integrity constitute the major headings. Under these, individual cognitive tasks and individual psychophysiological techniques which have been used to evaluate them will be discussed.

THOUGHT PROCESSES

The first category of cognitive activity is deliberately named as broadly as possible to encompass all of the vague activities subsumed under the term "thought" itself. Obviously, it is far beyond the scope of this work (and the author's capability) to enter into discussions about the physiological or psychological meanings of thought processes. Instead, it is directly of interest to identify certain behaviors which are generally agreed to reflect cognitive activity as their principal determinant. Even with this admittedly operational orientation, it is clear that categories of cognitive processes will not be found to fit everyone's biases. "Problem solving" is a deceptively easy term to operationally define. Yet, not everyone will agree to exclude choice reaction time from the category, since the choice essentially solves a problem.

In the following sections, specific types of behavior have been chosen and are discussed as if they represented unitary cognitive processes. It is recognized that this may not be the case. However, from

an operational human-engineering viewpoint, it is certainly acceptable to discuss them in this way. The important goal here is to define the ways these behaviors might be measured psychophysically in operational contexts. The subtleties of taxonomic classification can be considered in another place.

STIMULUS MEANING AND RELEVANCE

If communications between people and between machines and people are to become more important in future aircraft systems, then the study of meaning and relevance of stimuli will become critical to further progress. The vast increase in the volume of information presented to the operator, along with increased requirements for speed, present staggering problems to the system designer. Information must not only be presented in its entirety, but it must be sequenced and displayed or delivered only when needed. As volume of input increases, a given bit of data must compete with thousands of others for the operator's attention. We can no longer afford the luxury of dedicated displays and dials, but must multiplex the critical information when, and only when, it can be used. This requires determining when the operator's "channel" is open to receive input, when a given input has been processed, and when its meaning is appreciated.

Basic science has not produced anything near the elegant theories of cognitive function which would be necessary to achieve these goals. Models of information processing (Norman, 1969) have begun to point out some of the complexities, and behavioral techniques are just beginning to be able to assess such behaviors. Psychophysiology is not in a much better position. The measurement techniques discussed below reveal fine sensitivity to variations in the relevance, expectancy, and meaning of stimuli to the individual. However, none have been developed to the point where they can be used in operational settings to control information presentation to the subject. They represent solid and promising beginnings, but a great deal of development is still required.

Galvanic Skin Response. As noted earlier, the Galvanic Skin Response (GSR) has been used for many years to index the emotional (autonomic) changes induced in the subject by stimuli. Such attempts have met with mixed success. Yet, under proper circumstances, this measure can produce stable results with meaningful applications. Bernstein, Taylor, and Weinstein (1975) obtained phasic GSRs to tones in which significance was manipulated by requiring different classes of motor response. They found consistent changes in the GSR which were correlated perfectly with stimulus significance. The presence of such changes with verbally-induced significance suggested central mediation of the response. Further, they were able to dissociate a motor "execute" GSR from those associated with stimulus significance. Results such as these are encouraging. However, it must be remembered that the GSR is, in general, a slow response. It is difficult to see how it might be used on-line to assess meaning. For laboratory studies, on the other hand, confirmation and extension of the above results could provide a useful measure of stimulus significance.

Pupillometry. Changes in the use of pupil size to index cognitive affect have been discussed earlier. Generally, this procedure appears to have fallen into disuse at the moment for measuring meaning and emotional tone. Current interest is limited to assessment of workload, to be discussed later.

Electroencephalographic (EEG) Measures. The EEG in raw form has been used in several experiments dealing with evaluation of short-term memory. Gale, Jones, and Smallbone (1974) found a strong association between increased EEG arousal and poor performance on an immediate recall task. However, Surwillo (1971) had found no statistically significant differences in frequency of EEG during acquisition of lists of different numbers of digits. In an extension of their original study, Gale, Jones, and Smallbone (in press) again found that increased arousal was associated with high errors. In addition, the EEG discriminated between serial position of items to be recalled, subjects, good and poor trials within subjects, and trials over time. If confirmed, these results will be most intriguing.

The transient evoked potential (EP) has, again, proven to be the most durable measure of cognitive relevance. In the early 1960s, it was recognized that the EPs to relevant stimuli are larger than those to non-relevant stimuli (Chapman, 1965). Many investigators, most notably Donchin and his co-workers, then set about establishing the particular parameters of the response which measured such relevance (Donchin and Cohen, 1967; 1969; Donchin and Sutton, 1970). It became clear that the major characteristic of the cortical evoked response sensitive to cognitive function is the large positive-going peak which occurs between 200 and 500 milliseconds post-stimulus. This peak is absent if decision or attention is not required from the subject, and when it occurs it seems to be capable of indexing a wide variety of stimulus meaning and relevance. Beck (1975) reviewed the literature dealing with this positive component (called the P3 or P300) and concluded that it is enhanced when and only when cognitive information is being actively processed by the subject.

As an example of the way in which P3 changes with stimulus relevance, Figure 22 presents data from a study by Gomer, Spicuzza, and O'Donnell (1976). These authors utilized the behavioral paradigm proposed by Sternberg (1969). In this design, a previously determined set of alphabet letters is memorized by the subject ("positive" set) and all other letters are considered the "negative" set. Individual letters, both positive and negative, are then flashed to the subject, and the task is to indicate as rapidly as possible whether the flashed letter (the probe item) belongs to the positive or negative set. Sternberg has demonstrated that reaction time for this task increases in a predominantly linear way as the number of letters in the positive set is increased. In addition to traditional reaction time measures, the authors also obtained visual evoked responses to the onset of the letters. Analysis of the peak latencies and amplitudes then were carried out to determine whether any processing differences would appear in the evoked response as a function of the "relevance" of the letter as defined by its membership in the positive set.

Results of the reaction time measures were essentially in agreement with previous Sternberg data. In addition, the amplitude of the P3 peak showed a clear enhancement for the relevant (positive set) letters as opposed to the non-relevant ones. Further, this difference was progressive over the levels of increasing memory workload (an observation which will be discussed in more detail later, in the section on workload). This result for the P3 confirms many previous observations concerning the sensitivity of this peak to cognitive meaning or relevance, and raises the possibility that the measure could be used to index reception of a message by the subject. Obviously, it would have significant operational impact if it could be determined, within 500 msec after an event, whether the subject has processed the information or not.

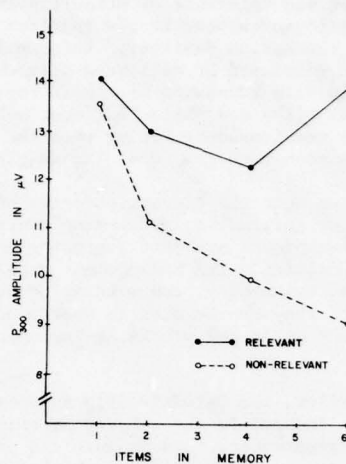


Figure 22. Amplitude of P300 as a function of stimulus relevance and memory load.

Other developments have brought this possibility somewhat closer to reality. It has been determined that the P3 of the evoked response can be reliably isolated on a single trial, without averaging repetitive stimuli, by using stepwise discriminant analysis (Donchin, 1969; Donchin and Herning, 1975; Squires and Donchin, 1976). McGillem and Aunon (unpublished data) from Purdue University are using maximum likelihood discriminators to reveal even smaller events in the single-trial evoked response. Further development of these techniques may permit on-line, operational utilization of the EP to assess stimulus relevance.

In a series of recent studies, Squires *et al.*, have further elaborated the nature of the P3 component. It was first noted (Squires, Wickens, Squires, and Donchin, 1976) that several later peaks in the cortical evoked response were sensitive to the sequence of stimuli preceding the stimulus yielding the EP. Thus, if binary events are presented in a random sequence, there will be occasions when the "rarer" of the two events occurs two, three, or even four times in a row. These chance occurrences will, of course, be very unexpected. The subject behaves as if the series is constrained by sequential rules. High expectancy stimuli (a "frequent" event after a "rare" event) yields a low amplitude P3, whereas a low expectancy stimulus (two or three "rare" events in a row) yields a high amplitude P3. These amplitudes are influenced by stimuli as far back as five preceding the eliciting stimulus. The authors propose that expectancy, dependent on a decaying memory for events within the prior sequence, as well as on other factors, determines the P3 amplitude. Later, this effect was shown to hold for visual as well as auditory stimuli (Squires, Petuchowski, Wickens, and Donchin, 1977) and for cross-modal stimuli (Squires, Duncan-Johnson, Squires, and Donchin, 1977).

Ford, Roth, and Koppell (1976) were unable to find sequential probability effects on P3, and suggest that this peak may be more determined by temporal rather than sequential uncertainty of events. However, they also recognize that the range of probabilities in their experiment may have been too narrow to demonstrate sequential effects.

This phenomenon therefore, is relatively robust although it has not been observed under all conditions. It appears to be tapping very sensitive aspects of cognitive function such as the sophisticated processing of stimuli based on prior registration and interpretation, the global probability level, and the reliability of the present situation. It is likely that such a complex and advanced level of cognitive processing would be sensitive to a number of factors in the individual, and in fact, this has proven to be true under at least one circumstance (see Wickens, *et al.*, 1976; 1977, in discussion of workload later in this AGARDograph).

It will be appreciated as further discussion of cognitive function reveals more and more interpretations of the P3, that this peak is sensitive to many variables. In fact, the P3 shows sensitivity to so many cognitive functions that it has sometimes been criticized as being too general. If this is true, then there would be little hope of utilizing the measure in any meaningful way. One could never be certain which variable was accounting for changes in latency or amplitude. However it will also be clear that in previous studies, the factors to which P3 is sensitive were able to be independently manipulated. Some skill in

design is necessary in order to make sure, for instance, that a given change in P3 is truly representing relevance rather than stimulus expectancy. In laboratory settings, such control should be within the capabilities of any competent researcher. As more is known, and more variables effecting P3 are isolated, the experimental design problem becomes more complex. However, it has in no sense reached the stage where the design problem appears unmanageable in most circumstances.

Another possibility has been raised concerning the P3 and the number of factors which affect it. It has been suggested that the large positive component which is usually considered to be P3 may in fact, consist of several smaller, higher frequency components (O'Donnell, 1977). It must be remembered that the P3 is isolated by sampling the EEG from the instant that the external event occurs (a flash, a tone, etc.). While this may be entirely appropriate for sensory components of the evoked response, it may only be an approximately correct trigger point for cognitive events. The beginning of information processing may occur somewhat variably after the presentation of the physical stimulus. A slight variation of this kind (even as slight as a few milliseconds) would tend to produce a "smeared" peak, such as that seen in the P3 component. It is possible that if we could trigger on the initiation of the cognitive processing sequence itself, we would discover that the P3 contains a number of separate components, each of which varies with a smaller number of cognitive factors. O'Donnell (1977) suggests that, as a start, eye movements may be used to trigger the evoked response, and efforts to test this hypothesis are currently underway (Moise, 1978). Given the ability to analyze the evoked response on a single stimulus presentation, it may also be possible to trigger a given sample off a peak within the evoked response itself. Therefore, the second positive peak (or any other identifiable point) could be used as a trigger for a new evoked response.

Further indications that the P3 may be made up of a number of smaller peaks comes from latency-corrected-average techniques being developed at Purdue University and elsewhere (Aunon and McGillen, 1979). When a latency correcting procedure based on overall variability of peaks is applied, it is clear that the P3 shows a number of distinct "minipeaks", each suggestive of a discrete event occurring within the overall time period. Suggestive evidence is also seen in the work of Chapman (1973) who used principal component analysis to isolate separate patterns within the evoked response. These revealed at least 15 orthogonal components in the EP waveform, many more than are usually visualized. The net result of these indications, of course, does not prove that P3 is generated by multiple cognitive sources. Even further, it does not prove that these multiple sources independently assess separate cognitive functions. However, the hypothesis is a reasonable one, and the pay-off should such independent assessment be demonstrated, would be large enough to justify significant research efforts.

PROCESSING AND DECISION

One step beyond determination of simple relevance of a stimulus, we must consider the processing and, ultimately, the decision components of cognitive function. In this area, behavioral measures have been especially unable to come up with on-line, reliable techniques usable in real-world environments. Game theory, computer modelling, and other complex techniques to define and assess the quality of processing and decisional processes have had some success. They are not suitable, however, for day-to-day evaluation of subjects in operational settings. In fact, no single test strategy to monitor or predict the moment-to-moment capability of an individual in processing information has emerged.

Psychophysiological techniques attempting to do this have involved two main areas of investigation: evoked responses, and measures of interhemispheric function. These have not yet produced a generally usable measure, but have had some success, not only in indexing the existence of decisional processes, but to some extent in measuring the type and quality of processing underlying them. Another measure, pupillometry (Simpson and Hale, 1969), has been investigated during decision making tasks. However, subsequent work has established that this measure is more appropriately considered an index of workload, and will be discussed under that heading.

Evoked Responses and Decision. It was observed in the early 1960s that the later components of the evoked response were enhanced if the subject was required to make a difficult decision, and were not enhanced for an easy, routine decision (Davis, 1964). Subsequent work confirmed this observation, and indicated that the enhancement was independent of stimulus modality. The P3 amplitude appeared to indicate not only stimulus relevance, but the difficulty of decisional processes in those cases where relevance was controlled. When the same physical stimulus was designed to produce different behavioral decisions, significantly different evoked potentials were obtained (Begleiter, 1975). Further, Donchin (1975) has established that the P3 enhancement during decision processes is totally independent of other electrical events which probably indicate "expectancy" (see section on cognitive readiness, under the discussion of attention, later in this AGARDograph).

Interhemispheric Measures of Cognitive Processing. A more recent focus of interest in assessing cognitive function emphasizes the contribution of the two hemispheres of the brain. Measures of interhemispheric function stem from early observations that the two sides of the brain appear to be differentially activated during processing of different types of information (Cohn, 1971; Galen, Ornstein, Kocel, and Merrin, 1971). These and other studies indicated that processing of verbal, sequential, logical material was accompanied by greater EEG activity from the left side of the brain in most people. Musical, spatial, wholistic material produced increased activity on the right side. Techniques used to estimate the amount and power of EEG activity ranged from simply calculating the average power in all frequencies from 1 to 35 Hz (and taking the ratio of right hemisphere over left) (Galin and Ornstein, 1972), to elegant evoked response measures (Caperell and Shucard, 1977).

It appeared possible, from early results, to evaluate the involvement of each hemisphere simply by taking EEG power ratios. This effect was even more powerfully demonstrated if only the alpha band (8-12 Hz) was measured (Doyle, Ornstein, and Galin, 1974). The asymmetry was shown to be genuinely related to cognitive task rather than personality variables or cognitive "style" (Morgan, MacDonald, and Hilgard, 1974). More recent studies have attempted to control for possible sources of artifact in the earlier results, particularly with respect to the behavioral tasks employed and the measures of brain power (Donchin, Kutas, and McCarthy, 1976).

One recent technique uses an irrelevant auditory tone to generate an evoked response while subjects are engaged in processing auditory information (Shucard, Shucard, and Thomas, 1977). In general agreement with previous studies, it was found that the left hemisphere evoked response was higher in amplitude during verbal tasks than that in the right hemisphere. For musical tasks, the right hemisphere showed higher responses. Paradoxically, the same laboratory (Caperrell and Shucard, 1977) reports that if the evoked response is generated by a visual stimulus, while the subject is engaged in mental tasks, the amplitude was attenuated in the hemisphere most involved in the cognitive process. Thus, it would appear that if the "probe" stimulus is in the same modality as the ongoing activity's input, evoked responses are enhanced in the more active hemisphere. If the cognitive activity is independent of modality, the probe stimulus will generate a reduced evoked response on the more active side. The reasons for such specificity are not well understood.

In a remarkable series of studies, Rebert (1977a; 1977b; 1977c) first established that reaction time to words was fastest when left hemisphere arousal (as measured by an alpha index) was greatest. Subsequently, Rebert used the subject's own EEG to trigger a reaction time stimulus. Depending on whether the right or left side showed enhanced activity, either word or pattern stimuli were presented. In most subjects, there was a differential effect of hemispheric activation on reaction time, although the underlying causation appeared to be somewhat more complex than would have been predicted by a simple alpha activation model. Generally, however, the studies showed that overt performance does depend on the functional asymmetry of the brain. In the most impressive study in this series, Rebert obtained bilateral EEGs from subjects playing the spatial TV game PONG (tennis). Compared to rest periods, alpha was suppressed in the right hemisphere (i.e., it was activated) during the game. In temporal and parietal areas, this effect increased linearly during the course of a rally. Further, the increase was reversed during the one second prior to an error. This effect did not show up in the central leads. Other results indicated that asymmetry is due to perceptual factors, and in some way indexes the subject's ability to respond to both verbal and spatial material.

Measures of Problem Solving. Closely related to the question of decision making is that of problem solving. From an operational viewpoint, problem solving is a much more difficult area to study than simple decision processes. The decision, often enough, has or can be made to have an "endpoint", the point at which a decision is made. Problem solving, however, is by its nature a continuous, often long-term process. It is therefore much more difficult to identify discrete points or steps to measure or to use in evaluating the behavior on-line or in triggering an evoked response. For this reason, not a great deal of progress has been made in this area.

One line of investigation (Newton, 1977) appears promising. Typically, the beta range of the EEG frequencies (between 13 and 30 Hz) has been considered to be indicative of "activation" such as might be present in problem solving. Researchers seldom record above this level. Newton found that there is significant EEG power at 40 Hz which shows clear increases during problem solving periods. This activity can be dissociated from both high beta (21-30 Hz) and from muscle activity. Patterns of shift in the 40 Hz activity over different electrode sites and for different kinds of activity were quite complex, but indicate lawful relationships at these high temporal frequencies over various brain areas for specific kinds of problem solving behavior.

Measures of Confidence Level. The confidence that an individual has in a decision or problem solution is also of considerable interest to the evaluation of operator behavior. Donchin (1968) first observed that a positive peak at about 250 msec appeared in the visual evoked response when a subject was certain about a judgement (in a threshold detection task), even if the response was not correct. Others have confirmed this basic phenomenon. Rasmussen (1972) extended these results to a five-point confidence scale, and showed that two long-latency, low frequency components with peaks at about 225 and 450 msec were related strongly to detection and confidence levels. In general, however, the possibility that these peaks would reliably serve as an index of confidence level in a subject has proven to be difficult to confirm. Although the basic phenomenon may be real, the peak at 225-250 msec is also affected by many other things. Outside of the signal detection task, it tends to become contaminated by other peaks reflecting additional aspects of the stimulus, processing or response (e.g., Chapman, 1973). Thus, while the above studies reveal intriguing possibilities, a great deal of laboratory research is necessary to delineate how these may be exploited.

Much of the research attempting to do this utilizes a signal-detection paradigm, since it is easy to obtain confidence measures in such a case. Hillyard (1969) suggested that there is a complex relationship between expectation and resolution of uncertainty which determines the amplitude of P3 and also of the Contingent Negative Variation (CNV). Based on results obtained in experiments where subjects were attempting to detect threshold signals, Hillyard postulated that the CNV indexes the environmental scanning for a specific, expected stimulus, and the P3 indexes the detection of the stimulus. Later, (Hillyard, Squires, and Bauer, 1971) it was observed that the P3 to detected signals was much higher than that to undetected signals, falsely reported signals, or correctly reported non-signals. The authors concluded that the P3 amplitude therefore reflected the certainty of the subject that a "hit" (or real signal) had occurred, and by implication, that it did not reflect the same certainty that a signal did not occur. This conclusion was challenged by Cael, Nash, and Singer (1974) who found P3 enhancement whenever the subject resolved uncertainty, whether it was to a hit, a miss, or a correctly reported non-signal. Subsequent work has tended to confirm this later view.

Signal Detection Performance. Threshold detection of visual and auditory stimuli were discussed in the last section. However, the detection of a signal can often be more than a simple threshold determination. If the signal is buried in noise, if the subject has some a priori evidence concerning its occurrence, or if "detection" involves discriminating one level of stimulus from another, the task no longer involves absolute thresholds, but now contains differential threshold sensitivity and perhaps an increased cognitive processing load. Many investigators have attempted to develop psychophysiological measures and theories regarding the optimal conditions for efficient performance (see the Lacey-Obriest controversy discussed earlier). Some of these have dealt with reaction speed to suprathreshold stimuli, and others have dealt with absolute and differential thresholds. A sample of these later ones will be presented here, and those dealing with reaction time will be discussed in the next section.

Much research was stimulated by Lacey's hypothesis that cardiac decelerations would have a facilitative effect on attentional processes. In an early study, Kalafat (1971) gave subjects an auditory detection task which was timed to appear during different phases of the cardiac deceleration which is known to occur after a warning stimulus. The probability of signal occurrence was also communicated by the warning stimulus. It was found that when the subject expected a difficult discrimination, there was a greater number of cardiac decelerations. However, no relationship was found between cardiac deceleration and sensitivity or criterion level. Similar results have been reported for other signal detection tasks (Delfini and Campos, 1972). On the other hand, Simons and Lang (1976) report that in an auditory pitch discrimination experiment the stimulus which produced fewest errors in judgement evoked the largest cardiac rate response (see also Lang, Gatchel, and Simons, 1975). From many such studies, frequently reaching apparently contradictory conclusions, a general feeling appears to be emerging among applied researchers that cardiac deceleration does indeed index the perceived difficulty of a task or the degree of external scanning being carried out. However, the micro-events of the cardiac cycle, if they influence detection sensitivity at all, do so in such a minimal and rapid way as to have little interest for applied research.

Reaction Time Performance. The speed with which a subject actually responds to a stimulus, whether it involves choice between two or more alternatives or simply rapid response to the occurrence of a stimulus, is the final product of a long series of processes. These have been well documented, from Donders, who first attempted to break the response into component parts, to the present day (see various recent sources for an extensive annotated bibliography). The exhaustive study of reaction time (RT) in its own right, as well as the number of studies using it as a dependent variable attest to its importance in the real world. In general, researchers are interested in determining the factors which contribute to or cause good or poor reaction times, and in finding methods of enhancing RT performance.

Psychophysiology has not failed to join in these attempts. Along with behavioral investigators, psychophysicists have spent a great deal of time searching for aspects of the subject's physiology which correlated with good RT behavior, and have attempted to use these correlations to predict performance. Respiration, cardiac measures, and EEG have been used most often, and examples of these approaches are presented below.

The Respiratory Cycle and Reaction Time. One recent study (Beh and Nix-James, 1974) has investigated the relationship between the phase of the normal respiration cycle and simple RT. An auditory signal was presented and subjects were required to respond as rapidly as possible. The respiratory cycle was broken into three segments, and response times to stimuli occurring during the three phases were calculated. It was found that mean RT for signals presented during the inhalation phase was significantly less than for signals presented during either the exhalation or pause phases. Although differences, in absolute terms, were so small that it would be unlikely to affect an operationally meaningful behavior, this result could have significant theoretical implications, since it reinforces the concept that the human shows constant micro-fluctuations in reception and processing. However, as will be seen below, such results must be tested very cautiously, since they are not supported by all of the evidence from other physiological systems.

Cardiac Activity and Reaction Time. As noted in an earlier section of this AGARDograph, several authors had suggested that variations in blood pressure which occur with each heart beat may be related to RT. Thompson and Botwinick (1970) investigated this hypothesis in a series of four studies. Stimuli were presented at 0, 200, 400 and 600 msec following the R wave, and during the ascending slope of the R, T, and P waves of the cardiac cycle (see Figure 4). No relationship was found in any of the studies between RT and any of the cardiac phases. Thus, any simple connection between the two can be ruled out (see also Bostock and Jarvis, 1970).

The question of whether heart rate is related to reaction time proved to be somewhat more confusing. Surwillo (1971), among others, showed that background heart rate level was not related to RT in any meaningful way. This result was true both between and within subjects. However, a number of other studies have shown convincingly that heart rate variability may indeed be related to reaction time (Porges, 1970; 1972; 1973). In a typical design (Porges, 1971) subjects were given an RT task with either a fixed (16 sec) or variable (16 to 28 sec) foreperiod. Groups were divided into three resting heart-rate variability levels (low, medium, high) based on the variance of 25 beats during a rest period. With the fixed preparatory interval, no correlations between heart rate variability and RT appeared. However, with the variable foreperiod, a correlation of $-.711$ (significant at $.001$) was found between resting heart rate variability and reaction time. Those showing more variable heart rates had faster RTs. In essence, this means that with temporal stimulus uncertainty, some factor increasing the beat-to-beat variability of heart rate shows a remarkably high correlation with fast responses. From a practical viewpoint, one could predict RT, in this type of situation at least, from heart rate response patterns. It would certainly appear desirable to repeat and extend these findings. Temporal uncertainty is a characteristic of many operational systems and tasks, and it is not always possible to obtain actual reaction times from operators on such systems. The ability to screen potential operators and to predict their reaction time would have considerable importance.

The EEG and Reaction Time. In the early 1960s, it was reported that when subjects were asked to decide between two alternatives, their decision time was high related to the average period of their brain waves (Surwillo, 1964). Subjects with "slow" brain waves required longer to make decisions. Further, when EEG frequency was held constant, the relationship between age and decision time disappeared. Thus, it was postulated that EEG slowing was a factor (if not the only factor) behind the "age associated drop in information capacity of the central nervous system". This interpretation of the EEG frequency/RT relationship was supported by a study which used sequential analysis of the EEG frequency in a vigilance-reaction time task (Morrell, 1966). Reaction time to a photic stimulus could be predicted by the EEG frequency.

However, these results have not always been obtained. Boddy (1971) performed a series of experiments using the mean alpha period and the overall mean EEG period as a correlate of RT. Under conditions of high incentive, non-significant correlations were found in two experiments. Further, the mean EEG period during the one-second interval just prior to the RT stimulus produced non-significant correlation with RT. In other cases (Thompson and Botwinick, 1968) the relation between age and EEG slowing was not supported.

These failures to replicate the Surwillo experiments can be explained on several grounds, but the fact remains that under these conditions, the raw EEG does not appear to predict reaction time.

In spite of this limitation, Surwillo (1975) has hypothesized that the speed of information processing is a function of a "cortical gating signal" and a recovery period of events activated by the gating signal. The gating signal is assumed to be measured by the half wave period of the EEG. Thus, it would be predicted that there should be a correlation between RT and the particular frequency of the EEG at the moment of stimulus presentation. In a group experiment, some evidence for this type of correlation was found. However, much more specific studies will have to be carried out if such a hypothesis can be considered supported. Spatial distribution of the EEG is not uniform, and it is difficult to believe that a crude measure such as the raw EEG would reveal "the" gating signal reliably from one electrode derivation. Nevertheless, one should not ignore the consistency of results which suggest that even the raw EEG shows significant correlations with RT under some stimulus conditions.

The Evoked Response and Reaction Time. As early as 1965, Donchin and Lindsley (1965; 1966) revealed that the transient visual evoked response showed lawful relationships to reaction time. Shorter latency of major peaks was related to faster reaction time. In addition, faster RTs were associated with larger amplitude evoked responses, and knowledge of results shortened reaction times and increased the amplitude of the evoked response. Morris (1971) confirmed such relationships for the most part, and determined that the evoked response measures were more strongly related to RT than raw EEG measures of arousal. In another study, Bostock and Jarvis (1970) obtained auditory reaction times and evoked responses to stimuli phased to the subject's cardiac cycle. A very strong relationship was found between both the amplitude and latency of a negative peak at about 250 msec and the speed of RT. Thus, it was suggested that this peak (resembling the familiar N2 component) might serve as a moment-to-moment index of the level of arousal of the subject.

Not all investigators found the P3 latency and RT to be correlated. Karlin, *et al* (1970, 1971) disagreed not only with the correlation, but with the prevailing interpretation of the P3 itself. These investigators felt that P3 was only indirectly related to cognitive aspects of a stimulus through the mediation of momentary arousal factors. In 1977, however, with a powerful demonstration of experimental design and computer sophistication, the Donchin laboratory cleared up many of the ambiguities and questions concerning the relationship between P3 and RT (Kutas, McCarthy, and Donchin, 1977). Subjects in their experiment were presented with series of words, one at a time, every 2 seconds. In each series, there was one class which was rare (20 percent) and one that was frequent. Subjects either counted the infrequent words, or pressed a button when they appeared. Under one set of conditions, subjects were told to maximize speed, while under another they were told to maximize accuracy.

It was hypothesized that the inconsistent results found in correlating P3 with RT may be due to the fact that P3 can be multiply determined. If P3 latency represents stimulus evaluation time, then whenever subjects exercise considerable care in stimulus evaluation (as they would under an accuracy criterion) then P3 and RT should correlate. Under a speed criterion, where the subject concentrates on response selection, the correlation should be lower. Figure 23 illustrates the results of this study

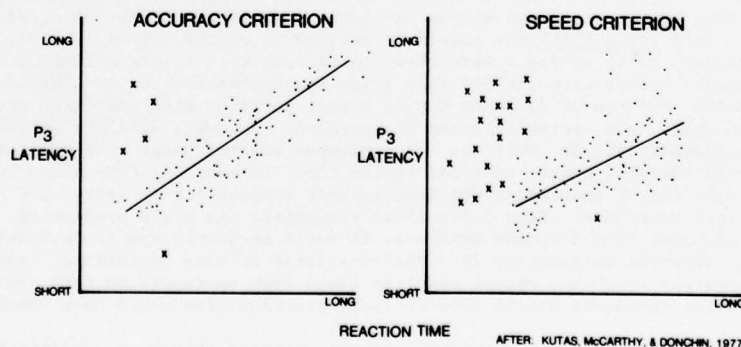


Figure 23. Reaction time and P3 latency under accuracy and speed criteria.

schematically, and they support the above reasoning. Under the criterion of high accuracy, the correlation between RT and P3 latency was .61, whereas with a speed criterion, the correlation was only .26. The authors believe that these data support the hypothesis that an RT stimulus initiates at least two processes; a response selection and execution process which can be measured by the RT itself, and a stimulus evaluation process measured by the P3 latency.

An even more remarkable feature emerged from the above data. The observed error rate under the accuracy condition was 3 per cent. Under the speed condition it was 9 per cent. These errors are schematically represented by the large "X"s in the figure. The interesting thing is that on error trials, contrary to the situation on accurate trials, the P3 latency exceeded the reaction time. The authors note that it is "as if the subject continued to process the information provided by the stimulus even through the overt response had been generated". This is a most intriguing possibility. If, in fact, the probability of error is very high on a trial in which the response selection and execution precedes the stimulus evaluation, then it should be possible to identify errors even if we cannot evaluate the response itself. In other words, preventing or disallowing responses where the P3 has not yet occurred should significantly reduce errors. Unfortunately, this is not yet feasible. Even with currently available techniques for single trial evoked response recognition, the P3 cannot be identified fast enough to allow immediate feedback (this would require virtually instantaneous analysis). Further, beyond this one impressive demonstration, the limitations, range of applications, and parametric constraints of the concept have not been studied. Still, it stands as a tempting model of the kind of cognitive analysis which may become possible with the evoked response.

Contingent Negative Variation and Reaction Time. At various times, it has been reported that correlations existed between the CNV and reaction time (Hillyard, 1973). However, these relationships were never simple and direct. They frequently held for only some subjects and not others, and sometimes involved the trial-to-trial variability rather than the absolute value (Hillyard, 1969). Various extraneous conditions such as distractors or response set can affect CNV without a comparable effect on RT (Kirst, 1975). The type and modality of the imperative stimulus in a CNV paradigm can differentially affect RT and CNV (Rebert, 1972). In view of these and many other factors, it is now generally felt that the CNV and RT are independent, and reflect the activity of different psychological processes (Rebert and Tecce, 1973). Further, it is felt that, with most subjects, there is no relationship between CNV and RT, except that the slowest RTs are associated with small CNVs (Rebert, 1974) (see additional caution regarding CNV discussed under "Attention" in the next section).

ATTENTION AND VIGILANCE

Beyond the operations involved in thought processes, *per se*, the human engineer must also be vitally concerned with what might be called the tonic level of cognitive processes. The person's moment-to-moment readiness to process information and respond (attention) or the long-term maintenance of such readiness (vigilance) is as much a determinant of final system performance as any design parameter. Thus, in the person/system interface, it is important to consider designs which foster attention and vigilance as well as to develop good indices of the individual's level of these cognitive characteristics.

There has been no lack of behavioral study in these areas (Mackworth, 1970a). Synthetic vigilance tasks, and paradigms for interpreting results on them have produced a vast amount of data. These provide good insights into the mechanisms underlying such factors, and give valuable clues and principles for design. Again, however, as in the study of thought processes, most of these synthetic tasks are intrusive with respect to the primary task. Thus, though they may be useful in the laboratory, they would have limited field use for on-line evaluation of a system or subject.

Physiological measures have, by and large, attempted to supplement behavioral techniques, either by providing a correlate of performance, or by providing insight into underlying mechanisms. In these goals, some considerable success has been achieved. In this section, representative examples of measures of attention and vigilance, especially eye movements, EKG, and EEG measures will be presented. The two topics will be discussed separately, although the distinction is, of course, somewhat arbitrary. In addition, the related topic of motivation will be discussed under this heading because of its obvious impact on attentional behavior.

ATTENTION

We wish to define attention rather loosely and superficially here as the short-term readiness to respond to a pre-designated task. In this sense, attentional behavior is focused, by definition, and we are interested in evaluating the quality of that focus. Many techniques have been used in the attempt to do this. Gaarder (1966) postulated that fine eye movements would be different during attentive and inattentive states due to a feedback control system. During attention, a closed-loop feedback was assumed to produce a stable system. During inattention, the loop was assumed to be open, and instability would appear in the fine eye movements. Some evidence for this type of control has recently been presented by Bahill and Stark (1979). Using a mechanical model and studying saccadic eye movements of various types, it was found that pulsed messages sent by specific motoneurons produce different kinds of saccades. The end-point of the movement is coded as a firing level, and if the pulse overshoots or undershoots, different types of corrections can be instituted. The frequency of these correction types has been found to change with disease, fatigue, and other states. Thus, it would be expected that more detailed analysis of the types of fine eye movements being made could index attention. As mentioned previously, Stern and his colleagues (Stern, Beideman, and Chan, 1976) have used sensitive measures of saccade velocity to do just this. Their results are most encouraging, and indicate that periods of "blanking" or "drop out" occur in certain conditions (see p. 21). If confirmed and related more fully to real-world effects, this measurement could provide an easy, non-obtrusive laboratory or even field approach to monitoring attention.

Galvanic Skin Response has also been suggested as a measure of attention, since it appears related to overall arousal (Raskin, 1973). However, most studies of this have tended to indicate that the GSR is too slow a response to measure moment-to-moment fluctuations. Lee (1969), confirming many previous indications, found that subjects who showed many spontaneous GSR fluctuations tended to show more impulsive types of

responses (e.g., responses to irrelevant stimuli, utilizing fewer cues before responding) than low GSR subjects. Thus, their attentional level may have been over-tuned. Similarly, Greene (1977) has shown that GSR responses become more frequent and larger as a discrimination task becomes more difficult. These overall indications of attention, while informative, offer little hope for a truly usable measure in operational environments.

A fascinating series of experiments carried out in France has raised the possibility that spinal reflexes may index attention. It was first noted (Bathien, *et al.*, 1967) that as attention diminished, there was a graded series of physiological effects which included changes in respiration, evoked potentials, heart rate, GSR, EMG, and spinal reflex sensitivity. Studying these reflexes further (Bathien and Hugelin, 1969) the authors found a stereotyped pattern of motor effects with attention. For instance, during attention, the soleus tendon reflex and tendon reflex of the biceps femoris were increased. However, the polysynaptic reflex of the biceps produced by stimulation of certain afferents was inhibited. These observations were confirmed and extended in other studies (Bathien, 1971; Bathien and Morin, 1972) to differentiate between intensive and selective attention. These observations raise important questions about the origin and maintenance of attention (and relate directly to the closed-loop eye-movement theory discussed above). It would seem desirable to further elaborate on these observations, and if they are confirmed, to develop a usable technique for assessing spinal reflexes.

The dependence of attention on interhemispheric balance was pointed out by Beck, Dustman, and Sakai (1969). Varying levels of attention were shown to be accompanied by changes in the minor hemisphere of the brain. Generally, increasing attention was found with increased amplitude evoked responses from the right side (although much of the evidence for this was clinically based and rather inferential). Later, Kinsbourne (1970; 1973; 1975) extensively studied the hemispheric activity accompanying attention to the left or right visual field. It was postulated that increased activity of one hemisphere biases attention to the contralateral side. Evidence for this position was developed mainly by asking subjects to perform tasks which were assumed to activate one hemisphere or the other. A task was then presented in one of the two halves of the visual field, and the side showing performance advantage was determined. These and certain clinical studies indicated that when two stimuli compete for attention, the relative degree of activation of the two cerebral hemispheres is an important determinant of the outcome. Obviously, this hypothesis holds considerable interest from an applied view in proposing a means to assess the moment-to-moment optimum location for presentation of information. However, the problems associated with utilizing interhemispheric measures must be remembered (Donchin, Kutas, and McCarthy, 1976). With such cautions in mind, however, it appears that this is a fruitful area to pursue.

Heart rate concomitants of attention have been observed for some time. Typically, when an individual is attending to a particular stimulus, the heart rate decreases, although a fairly complex pattern can emerge (Hassett, 1978). Spence and Lugo (1973) have shown that this decrease is specific to the information input stage, and is less evident during decisions. In general, however, the heart rate measure appears more valuable in assessing somewhat longer-term behavior, and will be discussed in greater detail under 'Vigilance'.

Electroencephalographic Measures. As early as 1960, it was suggested that a bioelectric scale of human alertness could be derived by using the EEG (Burch and Greiner, 1960). In a series of studies, Mulholland and Runnals (1962a; 1962b) established that alpha blocking indexed at least one kind of attention, i.e., transitory alerting to an external signal, and developed a feedback device for monitoring this. Later, however, Mulholland came to believe that alpha blocking was primarily an epiphenomenon due mainly to processes of eye fixation, lens accommodation, and pursuit tracking (Mulholland and Peper, 1971). Nevertheless, it was still felt that the alpha measure (no matter what its origin) could be used to control presentation of information in a visual display (Mulholland, 1970). Further, alpha, as controlled by the subject through voluntary "attentive looking", could be used to modify the sensory characteristics of a display (Mulholland, 1973). For instance, when alpha was suppressed by attentive looking, information presentation could be accelerated. Conversely, if attention lagged, the display could be made to "magnify". Unfortunately, adequate tests of these hypotheses have not been carried out, and in their absence, it is impossible to tell whether the obvious benefits of such a system are feasible. The variability of alpha in many people tends to argue against its easy utilization, but the fact that it is so easy to measure and that it has shown up as an index of arousal in so many laboratory studies suggest that it should be given an adequate test.

Another phenomenon of considerable interest is the occasional mental "block" which interferes with performance. The EEG during such blocks (not to be confused with alpha blocking) has been studied (Baumler and Klausnitzer, 1972). Visually scoring records from the right occipital area, it was found that the EEG just before the block shows a small increase in beta (13-30 Hz) activity. During the block, beta remains high, and all other EEG bands decrease. Again, this is a preliminary set of observations which, if confirmed, could have significant utility in predicting, monitoring, and perhaps preventing lapses in attention. Schmidtke (1972) has proposed a more ambitious use of the cortical EEG. Using derivative statistical analyses of the raw signal, a few parameters were derived to predict, within 2.3 seconds prior to a critical stimulus, whether a response would be made. Apparently, this technique has not been further developed, and no subsequent reports of use for this purpose were found (but see Bottge and Holock, 1973, discussed under the next section on vigilance).

Evoked Response and Attention. Study of the effect of attention on the cortical evoked response was undertaken early. This was done as much to determine how necessary it was to control for attentional variables in the evoked response as to study attention itself (Schechter and Buchsbaum, 1973). However, the studies produced a highly consistent (but not universal) set of results (Tecce, 1970). When attention is directed to the stimulus, the evoked response is usually enhanced. This is even more frequently true when attention is directed toward performance of a psychomotor task. Inattention, whether by distraction or by habituation, reduces the evoked response. The majority of the components of the evoked response are enhanced by attention, but several are not (Ciganek, 1969). Generally, for an auditory stimulus, the negative component of the response, peaking at 80 to 110 msec, shows greatest enhancement from selective attention (Hillyard, *et al.*, 1973).

Hink and Hillyard (1976) demonstrated an impressive potential application of this technique. They dichotomically presented two messages to the subject. One message was to be attended to, and the other ignored. Probe stimuli in the form of vowel sounds were used to generate evoked responses from each ear. The amplitude of the response to the probes in the "attended" ear was significantly larger than those in the unattended ear. Thus, the evoked response, in a totally non-intrusive task, was able to index the overall level of attention. It would be most interesting to determine if, using single trial evoked response techniques, the moment-to-moment attentiveness of the individual could be monitored. If this could be done, long messages could be monitored for the receivers' attention level, and "missed" parts ("mental blocks") could be repeated.

Another, more recent application has been made of the evoked response technique for assessing attention (Schaffer, 1977). Since the evoked response is an averaged waveform, it can be recorded in situations where the subject is totally preoccupied with something else. The brain's activity to the primary focus of attention will be randomly generated, and therefore will be averaged out of the response. Capitalizing on this, Schaffer recorded evoked responses from subjects while they watched TV programs of high and low interest content. The evoked response was generated by a flicker in the TV set, hardly noticeable to the subject. In three separate experiments, the late components to programs of high interest (e.g. M.A.S.H.) were larger than those to programs of low interest (e.g., Meet the Press).

It was hypothesized that the late components reflect the workings of an active attentional process within the brain, and that the evoked response technique permits study of this process under more real-life conditions than were previously possible. The sensitivity and further validity of this approach obviously must be established. However, if even part of the possibilities raised by this study are realized, the impact on the assessment of attention and interest will be considerable.

The Contingent Negative Variation and Response Readiness. Although it can be used to measure many things, the Contingent Negative Variation (CNV) is essentially an "expectancy" wave (see previous discussion of CNV form and paradigm). As such, it should be able to index the degree to which the subject expects, or is ready for, a stimulus, especially if a preparatory stimulus can be defined. However, the early excitement that the CNV would serve as an easy, useful index of readiness or expectancy has not proven justified. The phenomenon soon became mired in controversy concerning its origins, functional significance, and even its validity (Donchin, et al, 1973). However, McCallum (1969) and others have concluded, after reviewing a large quantity of evidence, that the CNV does in fact relate to attention, and that it indexes moment-to-moment changes in "conscious attention". Donchin (1973) points out that much CNV work has been done without necessary controls, and cautions that research on CNV is "like playing the recorder, an instrument that is extremely easy to play at an elementary level, but which is exceedingly difficult to play well".

After eliminating a great deal of the often contradictory early work on CNV, it appears possible to make the following general observations. The occurrence of a CNV is not dependent on a motor response (Donchin, Gerbrandt, Leifer, and Tucker, 1972). The CNV is probably not identical with the P3, and their distributions are topographically different, probably representing functionally distinct cortical mechanisms (Donchin, et al, 1975). However, some stimulus conditions cause both CNV and P3 to co-vary (Donchin, et al, 1976). Further, under some conditions, the phenomenon originally labelled CNV may consist of two or more processes. Jarvilehto and Fruhstorfer (1970) identify three groups of slow negative potentials associated with voluntary movements: (1) a 'readiness potential' which is centrally dominant; (2) a frontally dominant potential related to the subject's uncertainty, and (3) the actual CNV which may be a summation of the other two. Otto, et al (1977) have also investigated these separate components, and have developed a model which postulates that the various kinds of slow potentials summate linearly to produce the final slow potential seen on the scalp. Their data indicate that the CNV and 'readiness potential' contribute differently to the final response, and probably reflect different generator mechanisms.

In terms of the actual psychological meaning of the CNV, Tecce (1972) has proposed that the CNV magnitude may ultimately be determined by two interacting factors: attention may have a direct monotonic relationship to amplitude, while arousal level may be related in an inverted U fashion. Resolution of the CNV, after the imperative stimulus is presented, is felt by some to reflect selective attention paid to the stimulus (Wilkinson and Ashby, 1974). Overall, after nearly 20 years experience, the problems as well as the possibilities of the CNV continue to taunt the researcher. These have been well documented by McCallum and Knott (1974), and anyone contemplating work with the CNV should certainly read this review before attempting research or, especially, interpretation (see also McAdam, 1974).

Measures of Preparation for Voluntary Motor Acts. In 1964, it was found that the brain showed a slow negative potential just prior to a voluntarily initiated motor movement (Kornhuber and Deecke, 1965). This wave showed many characteristics in common with the CNV, but was not initiated by an external warning stimulus. It was designated the Bereitschaftspotential (BP), or readiness potential (RP) and has been studied under a number of conditions (e.g., Becker, et al, 1976; McAdam and Rubin, 1971; McAdam and Seales, 1969). The magnitude of the BP increases with motivation and decreases with inattentiveness, carelessness or decreased motivation (Deecke, 1973). The BP can also be recorded prior to the initiation of voluntary eye movements (Becker, et al, 1973). Because of the difficulties associated with recording and analyzing any slow activity on the scalp, this potential probably will be of limited value to the applied researcher until technology makes it more manageable. At the present time, it is simply too difficult to isolate the BP from the other sources of slow activity, and too difficult to interpret it when it is isolated. The fact that it precedes a response, especially an eye movement, however, makes it an extremely attractive measure for the applied researcher looking for predictive measures. For this reason, basic research into this area should certainly be encouraged.

A somewhat different observation concerning the period immediately preceding voluntary motion was made by Hazemann and Lille (1976). These authors noted that the somatosensory evoked response, generated by peripheral stimulation of a muscle, was attenuated by the temporal proximity of a motor act. The authors speculate that this might reflect a divided attention effect, with decreased orientation to the somatosensory stimulus. In any case, except in an extremely unusual circumstance, it is unlikely that such a technique would have applied implications. There may, however, be considerable applied value in the general observation that responses from other modalities appear to be attenuated in preparation for a

motor behavior.

VIGILANCE

In the previous section, attention was limited to short-term readiness to respond. In this section, vigilance is considered to be that same readiness continued over a long period of time. Of course, as noted earlier, the two are not mutually exclusive, and techniques discussed in the previous section may certainly be applicable here. However, the distinction is not a trivial one from the applied researchers point of view. Essentially, vigilance research attempts to assess and predict tonic changes or states in the individual. However, by the nature of a vigil, changes in vigilance must ultimately be validated by looking at performance on a discrete, usually rare event. Thus, there is considerable room for confusion. A vigilance measure should, strictly speaking, assess the overall response readiness of the person. As such, it must not be too sensitive to momentary fluctuations in attention. Thus, the measures will often be different.

Galvanic Skin Response measures and their utilization provide a case in point. As an index of moment-to-moment attention, the GSR has not proven extremely valuable. However, the slow, somewhat stable characteristics of the GSR are, in many ways, ideal for studying the long-term attentional capability of the individual. For example, Verschoor and vanWieringen (1970) found that skin conductance of "good detectors" in a vigilance task remained constant, while that of "poor detectors" showed a drop. More generally, the GSR-determined characteristics of personality referred to as "labile" and "stable" have been found to correlate with vigilance performance. Crider (1972) established two extreme groups based on whether the subject showed rapid or slow GSR habituation to serially presented tones. Such GSR habituation would be similar to the "orienting response" habituation which Mackworth (1970) predicted would be involved in poor vigilance performance. "Labiles" were defined as slow habituators, and "stables" were fast habituators. It was found that labiles showed a high and sustained level of performance, while stables showed an initial deficit which increased with time on task. Labiles had more spontaneous GSRs during the task, but no group differences in absolute conductance level appeared. In a subsequent study (Crider and Augenbraum, 1975) these results were confirmed. However, it was determined that the performance differences were due to group differences in the response criterion level rather than to differences in the rate of attentional decrement. Further elaboration has been provided by Hastrup (1977). Using both the frequency of spontaneous fluctuations in GSR and the orienting response habituation speed as determinants of lability, it was found that GSR measures predicted vigilance decrement over time for a difficult vigilance task, but not for an easy task. Thus, GSR measures may provide a good index and predictor of vigilance performance, particularly on difficult tasks where response criterion levels can differ appreciably.

Cardiac Measures and Vigilance. The mean heart rate for an individual, either at rest or during the task, does not appear to correlate with vigilance performance. However, several studies have implicated heart rate variability in such performance. Thackray, Jones, and Touchstone (1974) found a significant relationship between HR variability and performance decrement. Similarly, Wieringen (1975) found a rather tenuous indication that increased sinus arrhythmia may predict good vigilance performance. These results are reminiscent of the relationship found between heart rate variability and reaction time (see previous section on reaction time) and may be related to the same factor. They also complement the picture of the good performer in vigilance task which is given by the GSR measures; a reactive person showing considerable physiological lability. Together, such observations could be tested and developed into a prediction index for operators in vigilance situations.

The Electroencephalogram and Vigilance. Since vigilance is a long-term phenomenon, it is not surprising that many attempts to use the EEG to measure vigilance have used broad units of analysis such as alpha percentage or EEG abundance. The underlying premise is that a particular level of brain activity in any of the well-accepted bands of the EEG adequately reflects the activation level of the brain, and that this level will be reflected in performance. Many would argue with this view of the present time, since we now know that the EEG has much more complexity and specificity than such epoch analysis permits. Nevertheless, it appears that in many cases, relatively crude period analysis reveals respectable correlations with vigilance performance. More sophisticated statistical treatment can increase the power of these techniques even further.

Caille (1964) reported that the duration of alpha increased in all subjects from the first to the third watch in a monotonous vigilance situation. However, performance did not show a corresponding decrease. Daniel (1967) similarly failed to find EEG correlates of detection failures or errors, but confirmed the progressively decreasing arousal throughout the session as indexed by the brain waves. At these levels of vigilance, therefore, it appears that the decrease in EEG indices over time do not correlate with performance.

Overall EEG activation level, however, seems to produce a more positive picture. It has been found that good vigilance performance correlates positively with production of alpha frequencies in the low range and negatively with high alpha frequencies (Becker-Carus, 1971). Correspondingly, the amount of alpha activity is negatively correlated with performance. These, in turn, correlate with a number of personality variables such as neuroticism and rigidity, so it is perhaps more likely that these global characteristics were responsible for the observed effects rather than the alpha or brain activity itself. Of course, this cannot be inferred from the above correlations. Nevertheless, if the correlations are reliable, they would have utility in predicting vigilance performance. One group (Bottge and Holloch, 1973) has gone so far as to develop a 10-point identification system for EEG measures of vigilance. This system uses spectral analysis of the raw EEG, along with several derivative measures. It was able to define four states of vigilance, depending on eye status and attention to the signal. Results with initial trials of this approach were encouraging for some subjects, although not for others (Holloch, 1972). These should be further pursued. Such preliminary successes in utilizing neurophysiological measures have been obtained by others establishing laboratories to study performance, and it has been recommended that greater use be made of computerized neurophysiological evaluation in the assessment of vigilance (Monesi and Ravaccia, 1976; Offenloch, 1977).

O'Hanlon and Beatty (1977) have carried out such an assessment in subjects performing a simulated

sea-surveillance radar monitoring task. EEG measures of alpha, theta, and beta abundance showed a consistent and significant relationship to performance in the expected direction. These results provide evidence on the mechanism of an effect shown by Beatty, *et al.*, (1974). These investigators used biofeedback techniques to train subjects to suppress theta activity during performance of the same vigilance task as above. Although the percentages involved in both theta control and performance decrement were small, significant enhancement of monitoring efficiency was found during theta suppression, with the opposite being true during theta enhancement.

It appears that overall EEG measures have much to contribute to the prediction of vigilance performance, particularly if a global approach is taken. Sophisticated analysis procedures appear likely to add significant precision to these techniques. However, Gale (1977) has pointed out that one reason for the lack of short-term predictive success of EEG in vigilance research may be the fact that the brain reacts differently to short-term memory or response competition than it does to vigilance requirements. Confounding by these demands during the vigil may lead to momentarily contradictory indications. Such factors must certainly be considered in designing any EEG index of vigilance.

A series of studies by Dimond and Beaumont (1971; 1973) raise interesting speculations concerning the organization of vigilance behavior in the brain. It was observed that signals presented separately to the two brain hemispheres did not result in more detections on one side or the other. However, there were more false positives found when stimuli were in the left hemisphere than in the right as the task progressed. On the other hand, the left hemisphere showed better detection and gave fewer false positives early in the task. It was proposed (Dimond, 1977) that there are two different hemisphere vigilance systems in the brain. Based on the observation that totally split-brain patients show gross failures in vigilance, while those with the splenium of the corpus callosum spared do not, a new model of vigilance is proposed. A primary vigilance system in the left hemisphere shows high initial efficiency but decrements with time. A secondary system on the right side shows no decrement with time, but is less efficient. The splenium integrates the two and is essential for long-term vigilance performance. Such a view has great heuristic value, and would significantly impact design and scheduling if true. For these reasons, it should be investigated in much greater detail, and possible conflicts with the interhemispheric views expressed in the previous section of this AGARDograph (Beck, Dustman, and Sakai, 1969; Kinsbourne, 1970; 1973; 1975) should be reconciled.

The Evoked Response and Vigilance. Since the evoked response is a "discrete" measure which, nonetheless, seems related to the overall activation level of a subject, it holds considerable interest to the vigilance researcher. If some technique such as the evoked response could be used in a non-obtrusive way to index the subjects' level of responsivity, applications would range from laboratory design criteria to in-flight on-line monitoring. Haider was one of the first to establish that as vigilance fluctuated and waned over the course of a task, the amplitude and latency of certain components of the evoked response showed corresponding variations (Haider and Groll, 1964; Haider, Spong, and Lindsley, 1964a; 1964b). The drop in alertness correlated positively (.75) with the reduction in evoked response amplitude, and negatively (-.75) with the increase in latency. This effect seemed to be strongest to nonsignal (or irrelevant) stimuli over the detection period than to signal stimuli requiring a response. These effects were further studied by Fruhstorfer and Bergstrom (1969) who found that the vigilance-related decrease in amplitude was related to three prominent early components of the evoked response (N_{1a} , N_{1b} , and P_{2b}) but not the earliest positive peak (P_1). Storm (1970) later localized these changes most powerfully at the occiput (for visual input) and showed a negatively accelerated decline in evoked response amplitude with a parallel deterioration in performance within a session. Further, signal probability and not total stimulation was the major determinant of this amplitude decrease. Finally, the change in amplitude with time in the session was confirmed by Harkin (1975) who, however, did not find changes in evoked response latency or amplitude related to detection decrement.

These and other studies relating evoked responses to vigilance performance have been reviewed by Davies and Parasuraman (1977). From this review and from their own work, these authors conclude that both late component amplitude and latency measures are related to performance changes during a vigil, and to the effect of stimulus and response variables on latency of the response. Thus, it would appear that, from an applied practical point of view, the evoked response can supply a non-obtrusive way to monitor on-going vigilance. Irrelevant signals (or perhaps those requiring a minimum of decision) could be used during the vigil to generate evoked responses. Further research is, of course, needed to fully define the parameters which would affect such responses and to develop criteria for their interpretation with respect to vigilance level. However, the experiments necessary to produce such data are straightforward. In view of the consistency of effects found, the effort seems fully worthwhile.

Multiple Physiological Indices of Vigilance. Many studies have attempted to utilize multiple physiological measures, often biochemical, to differentiate between individuals who perform well or poorly in vigilance situations, or to monitor vigilance decrement. In one early study, O'Hanlon and Horvath (1969) monitored adrenaline, noradrenaline, free fatty acids, glucose, heart rate, respiratory rate, GSR, and neck muscle EMG during basal conditions and various kinds of vigil. It was concluded that adrenaline, heart rate variability, respiratory rate, GSR, and EMG were related to vigilance performance, indicating that monotonous tasks elicit widespread physiological reactions. Pooch, Tuck and Tinsley (1969) found a significant multiple correlation between visual monitoring performance, skin temperature, and systolic blood pressure. In another study, Tinsley (1969) found the same relationship, but no correlation between performance, skin resistance, heart rate, or pulse pressure.

However, as with any attempt to relate a large number of possibly interconnected measures to an elusive variable such as performance, results are not always consistent. Other studies have found performance to be related to neck muscle tension and sinus arrhythmia (Innes, 1973), heart-rate variability (Thackray, Bailey, and Touchstone, 1977), and a variety of biochemical factors. Carriero (1977) points out that complex designs and statistical transforms may be necessary to isolate the subtle interactions between all these factors. He believes, however, that ultimately it may be possible to develop an "alertness indicator" from such measures. Obviously, vigilance requirements initiate from generalized psychophysiological response in the individual. It may be possible to tap these in some meaningful way to assess alertness. It may also be possible to utilize only one of these indices to indicate the status of the entire system.

At the moment, this later approach seems to hold greater attraction in view of the complexity of the multiple measure task and the recent success of single measures.

MOTIVATION

The relationship between attentional behavior and motivation is easy to see. It has not proven easy, however, to measure motivation, either behaviorally or psychophysiological. Although the amount of work on motivation is voluminous, no single coherent theory yet dominates. Similarly, no single behavioral technique for assessing motivation is universally used. Essentially, motivation is usually inferred from good performance or from instructions.

Psychophysiological attempts to infer motivational states fall in three main areas: EMG, EKG, and EEG. None of these have been exhaustively studied and except for a few random clues, not a great deal of progress has been made in developing an assessment for motivation. Bartoshuk (1955) discusses indications that the slope of a progressive increase in muscle potentials (an EMG gradient) may be indicative of the degree of motivation on such tasks as mirror tracing and attentive listening. In a mirror tracing experiment, this gradient slope (especially for the right forearm extensor) was directly related to speed and accuracy of performance. This and other evidence showing gradient changes with incentive led the author to conclude that gradient slope is a direct function of strength of motivation to perform a given task. No other attempts to extend these results were found. It would appear that such a measure, if confirmed for other tasks and over a precise range of motivational levels, would be most valuable.

The effects of motivation on the cardiac cycle were studied by Levison and Fenz (1971). No differences in the cardiac cycle itself were found for different motivational conditions. However, increased amplitude of the beats was found with higher motivational states. Douglas (1972) also found a lack of effect of motivation on heart beat, but found increased sinus arrhythmia with increased motivation. Again, beat-to-beat variability appears to be a better measure of subject involvement than other cardiac measures.

The Contingent Negative Variation (CNV), for all its difficulties, appears clearly related to motivational level. In a series of studies, Irwin, *et al* (1966a; 1966b) found larger CNVs in conditions assumed to generate higher motivation (shock, manual response, variable effort, etc) than with low motivation. The authors interpret these findings as indicating that conditions which increase "energizing factors" in behavior also increase CNV magnitude. These results have essentially been confirmed by other investigators (Waszak and Obrist, 1969), but have not been used in any systematic attempt to apply the observation.

In general, then, attempts to measure motivation psychophysiological have met with considerable laboratory success, but almost no application. This may be due to the intrinsic difficulty of defining motivation behaviorally, or to the fact that, outside the laboratory, these measures become difficult to tie down to an amorphous concept such as this. In any case, the topic is certainly of high interest, and the few studies which have been done produced optimistic enough results that the effort should be re-instituted.

WORKLOAD AND FATIGUE

The final set of questions dealing with cognitive assessment relates to the breakdown of performance due to two interrelated factors. Workload is an abused term, frequently used in the sense of "overwork" or excessive loading. Thus, much workload research has attempted to pinpoint the breaking point of the individual rather than describing the amount of effort required by a given system at all levels of load. Rather than assume this end-point view here, we choose to look at workload assessment as an attempt to index a dynamic process. Thus, any "work" being done by the individual, even doing nothing, is a workload which ultimately will have its effect on performance. Thus, while we are certainly interested in measuring the activity level of the person at a given point in time, we are also interested in assessing the gradient of buildup in fatigue, distraction, etc., which will ultimately result in person/system failure. For this reason, measures which can be used to evaluate this continuous, long-term buildup are emphasized below. Those which assess the moment-to-moment load on the individual will also be presented, but more briefly.

WORKLOAD

The problems of measuring workload physiologically largely reflect the same difficulty encountered in measuring amorphous variables such as motivation. If the human responds to varying workloads with a physiological response, it is probably a complex and multi-faceted interaction between many components. These may or may not reflect an overall energy mobilization controlled by some major system operating as a unit. Work output is probably a function of the momentary mobilization of energy, the momentary and previous conditions of the circulatory and neuromuscular systems, and the momentary receptivity of the subject to further stimulation (Geldreich, 1953). Being so multiply determined, it is difficult to obtain either a simple behavioral or physiological index of the overall activity level of the person. Nevertheless, such efforts have been made, with varying success, in the past. In fact, so much work has been done that it will not be possible to summarize it all here. Representative examples will be given, with emphasis given to those measurement techniques which have received most attention. Among these, the study of heart rate, and particularly heart rate variability, stands out most dramatically. Other measures such as pupil size, EEG, GSR, and EMG will also be surveyed.

Cardiac Measures of Workload. As early as 1963, Kalsbeek (1963) recognized that the heart beat irregularity seen in normal healthy subjects sitting at rest was suppressed with increasing difficulty of a task, particularly a mental or perceptual task. A simple scoring system was proposed to measure this sinus arrhythmia, and this was used in a task where the number of binary choices per minute was increased as a method of inducing workload. The irregularity of the heart beat was diminished in direct relationship with the increase in workload (Kalsbeek, 1968). These effects were so strong that the relationship between sinus arrhythmia and mental load was quantified and presented as a scale by Kalsbeek and Ettema (1968). Danev and deWintar (1970) also found suppression of sinus arrhythmia during performance of a serial 8-choice

reaction task. However, they found that there were short term decelerations of the heart rate level after mistakes, while after correct responses there was an accelerating tendency. They warned that attention and errors have to be taken into account when suppression of sinus arrhythmia is used for assessing the level of mental load. Other investigators also find that, though sinus arrhythmia scores differentiate significantly between several levels of mental load, other heart frequency measures can be of equal value in some cases (Blitz, Hoogstraten, and Mulder, 1970). Considerable disagreement was therefore generated over the use of heart beat irregularity as a measure of workload.

The controversy generated by the sinus arrhythmia question continued into the 1970s. The validity of the measure in a visual inspection task, and its superiority to diastolic and systolic blood pressure were confirmed by Parikh (1971). Sayers (1971; 1973; 1975) analyzed the interval variability of the heart. In laboratory and field studies, the effects of mental workload on the cardiac sequence were confirmed by this author, and results suggested that consistent changes in the interval spectrum, mainly centered in the 0.1 Hz region, occurred under workload. It was suggested that these changes may originate with changes in the pattern of respiration. Further confirmation came from studies by Strasser (1973) and Schacke and Weitowitz (1971). Boyce (1974) pointed out that both heart rate and sinus arrhythmia increase for an increase in physical load. It was concluded that changes in heart rate and sinus arrhythmia are best regarded as generalized responses to the imposition of a load. In another effort, paced choice reaction tasks produced decreased heart rate variability with higher loads (Mulder, and Mulder-Hajonides van der Meulen, 1973). In addition, spectral analysis of heart rate variability revealed the existence of a 0.1 Hz component, strongly correlated with respiration. Finally, Hyndeman and Gregory (1975) proposed a technique for digitally processing cardiac intervals to present the necessary information for determining cardiac arrhythmia. This scoring technique was used in a decision making task, and was shown to give a reliable indication of mental loading.

As a result of all the work done on relating sinus arrhythmia to mental workload, an entire issue of *Ergonomics* (1973, 16, 1-112) was devoted to this topic. In this, Kalsbeek (1973) reviewed the problems raised by the use of heart rate irregularity as a dependent variable. It was argued that mental load should be considered as a multiply determined concept, and that sinus arrhythmia may, in fact, measure important components of that concept. However, care must be exercised in using the terms "sinus arrhythmia" or "mental load" as if they were unitary. The future of heart rate variability measures in field applications was viewed optimistically, although the need for care in such applications was pointed out.

In general, it would appear that sinus arrhythmia is able to index certain kinds of mental workload with considerable precision. Sensitivity has been shown across a wide variety of tasks, and an adequate range of loads. However, it should be noted that not all attempts to utilize this measure have been successful. Sherman (1973) failed to find a systematic change in heart variability as the difficulty of the sonar doppler identification task was increased. Similarly, Hicks and Wierwille (in press) found that heart rate variability was insensitive to a driver task. In spite of these occasional failures, the attractiveness and simplicity of a heart rate variability measure make it desirable to continue attempts to apply this measure. Although cautions, as expressed by Kalsbeek, must be taken into consideration, there is reasonable hope that the task will be worthwhile.

Several attempts to measure other characteristics of heart signals have been made. Spyker *et al* (1971) found particular characteristics of the vector cardiogram which correlated with workload. These included the standard deviation of the T-wave amplitude, the standard deviation of the R-R interval, and the T-wave amplitude itself. Similar EKG measures were later found to be significant predictors of workload in a helicopter simulation and flight tests (Stackhouse, 1973; 1976). Hasbrook, and Rasmussen (1970) obtained heart rates from experienced pilots flying 10 simulated ILS approaches in a single engine General Aviation aircraft. Heart rate increases were found during each approach averaging 5.2 beats per minute, while the overall mean heart rate level decreased on successive approaches. The authors interpret these results in terms of the demands of the task.

In summary, it would appear that heart rate variability, of all the measures used, most often appears able to assess specific questions of mental workload. While other measures appear occasionally sensitive to certain aspects of the workload situation, the variability measure seems to encompass more specific components of workload than other cardiac techniques. The cautions pointed out by Kalsbeek must certainly be considered. Further, Luczak and Laurig (1973) have demonstrated that only certain measures of variability will produce statistically significant changes with operator loading. However, keeping these questions in mind, it should be possible to utilize variability measures productively, perhaps even in field settings.

Pupillary Dilation and Workload. As early as 1966, it was shown that during the short term memory task, pupil diameter dilates as material is presented, and constricts during report (Kahmeman and Beatty, 1966). Subsequently, this phenomenon has been confirmed for a number of tasks. Westbrook, Anderson, and Pietrzak (1966) found pupil diameter increase in a pilot as the difficulty level of a tracking task increased. The Naval Postgraduate School, at Monterey, California, produced two studies confirming the validity of the pupillary measure. In one, (Hope, 1971) a correlation was shown between the difficulty of a multiplication problem and changes in pupil size. With difficult problems, there was an increase in pupil size for correctly answered problems, and a decrease for wrong replies. Edwards (1972) showed that pupil diameter increased for a one-bit and two-bit information assessing task. This increase reached a maximum at maximum information assessing capacity, and the pupil rapidly constricted as this capacity was exceeded. This same type of phenomenon was found by Poock (1973) who reported that when subjects were required to process information at 75 to 100 per cent of their maximum capacity, pupil diameter increased. However, when the information capacity was exceeded, the pupil constricted significantly to below the base-line diameter. It was concluded that pupil diameter may be able to identify points in time where mental overload occur (see also Poock and Noel, 1975). Gardner, Beltramo, and Krinsky (1975) found that pupil dilation reflects mental activity involved in the storage and retrieval of information, more than the actual mental workload. This does not preclude the use of the pupillary dilation as a measure of workload, however. Beatty (1978) in a powerful demonstration, has shown that many studies using pupillary dilation as an index of workload have essentially come up with the same results. In fact, it is possible, by adjusting the various demands placed on the subject and time measurements, to inter-relate

the various studies. When this is done, a remarkable consistency appears, and it appears possible that pupillary dilation could be used to quantify mental workload in a precise way.

It should be noted, however, that for operational situations the problems introduced because of pupillary dilation due to other sources such as eye movements, ambient changes in illumination, and even physiological changes generated internally by the eye musculature are virtually insurmountable. For laboratory situations, pupillary dilation appears as one of the most stable and productive indices of mental workload available. As long as all extraneous factors can be closely controlled, this measure is to be highly recommended.

Electroencephalographic Measures of Workload. The EEG has not been extensively utilized as a measure of workload. Spyker, *et al* (1971) used evoked responses, among other physiological measurements, and found two features correlated with the difficulty of a performance task. These features consisted of the amplitude of the P2 peak, and the overall maximum power in the evoked response. Defayolle, Dinand, and Gentil (1971) believe that such sensitivity may be due to the fact that evoked response changes reflect differences in the way the operator approaches the task. Thus, evoked response measures may be an indirect measure of the operator's view of the primary task. Donchin *et al* (1973) demonstrated that the amplitude of the P3 peak is a graded function of the complexity of the information processing required from the subject. However, this was true only when the subject was not under a time/accuracy pressure. Under such pressure, other factors appeared to dominate the P3.

Some evidence has emerged to indicate that during processing requiring simple recognition and discriminative responses, increased processing load generates larger late positive components in the evoked response (Poon, Thompson, and Marsh, 1976). In addition, the right hemisphere showed a large P2 component during simple recognition, and this asymmetry was enhanced during more complex processing. In an information processing task, Gomer, Spicuzza and O'Donnell (1976) found graded changes in the amplitude of the P3 component as a function of the number of letters that the subject was required to remember and discriminate. These differences however, were significant primarily for the non-relevant letters (those not in the memory set). The latency of the P3 showed a regular increase with increasing memory load (Figure 24). It can be seen that the increase in P3 latency with increasing memory load was extremely regular (approaching 99% linearity). This was much more consistent than the increase in reaction time with increasing memory load.

Recently, Wickens, *et al* (1976; 1977) have demonstrated a most ingenious technique for assessing tracking workload. These investigators presented an auditory task to the subject during the course of a manual tracking task. The secondary auditory task consisted simply of counting tones of a certain frequency. Two frequency levels were used, and the counted tones were presented less frequently than the non-counted tones. Results demonstrated a dramatic reduction in evoked response amplitude with the imposition of the tracking task. Simple analysis of evoked response P3 amplitudes, however, did not correlate with the level of difficulty of the task. More complex analysis, based on the sequential dependency of the P3 (Squires, Wickens, Squires, and Donchin, 1976 - see previous discussion of EEG measures under 'thought processes') provided a graded measure of operator loading in a 2-axis tracking situation. These results are most stimulating, since they suggest that a non-obtrusive secondary auditory task is able to index a visual workload of the subject, and to do so in a graded way. If validated, this technique will have a significant impact on the assessment, perhaps even on-line assessment, of operator workload.

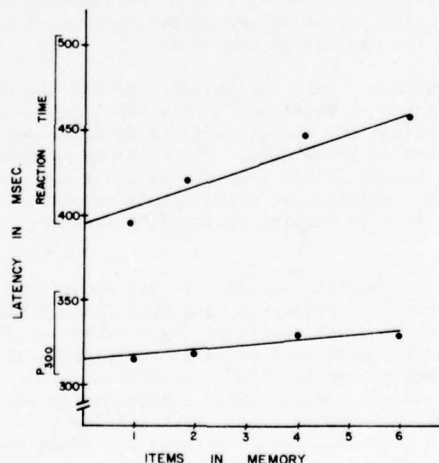


Figure 24. Reaction time and P300 latency related to memory load.

Voice Analysis and Workload. When an individual is under stress, changes in the fine musculature of the vocal cords or auxiliary apparatus cause slight differences in the frequency composition of the voice. In most cases, these changes serve to reduce the amount of physiological tremor, so that the voice has less frequency modulation. The theory with respect to workload would be that as the load reached a stressful limit, these changes would be seen.

One of the earliest utilizations of this phenomenon employed a commercial device, the "Psychological Stress Evaluator" (PSE) (see Wierwille and Williges, 1978) in police lie detection efforts. Considerable success and user acceptance is reported. More recently, at the Royal Aircraft Establishment, Farnborough, voice patterns of pilots flying British Airways Aircraft were analyzed (Cannings, *et al*, 1977). Although some success was achieved in analyzing both pitch and formant information, no workload analyses have been reported. Other attempts to utilize this measure have had limited success, and investigators have cautioned that problems can occur which make the procedure extremely difficult to use (Older and Jenney, 1975; Siminov and Frolov, 1977). Harris, North, and Owens (1977) have reported a successful use of voice recognition to aid in workload/stress analysis, but caution that the equipment itself can be a major source of error. In general, then, it appears that this technique may be in a primitive stage of development. It apparently suffers from a number of methodological problems which will have to be worked out. It is an attractive technique, however, with much to recommend it to the applied researcher in workload. It will be fascinating to see if its face validity is confirmed in subsequent research.

Other Techniques for Assessing Workload. At one time or another, virtually every psychophysiological technique has been tried in the attempt to assess workload. Most have been unsuccessful. None, other than those described above, have shown exceptionally great promise. A few of the techniques which have at least been successful in some laboratory situations will be surveyed here.

Galvanic skin response (GSR) was used by Harding, Stevens, and Marston (1973) who presented an information reduction task to subjects and took GSR measures in such a way that they could be related to response necessity, motivation, skeletal response conflict, or rate of information transmission. This last factor accounted for 95 per cent of the variance in the GSR, indicating that GSR could index mental load. Helander (1976) similarly, found that GSR seems to be "an efficient indicator of the mental effort involved in driving". However, responses had to be ensemble averaged over drivers in order to obtain the effect. Spyker, *et al* (1971) failed to find any of 27 GSR-related measures clearly related to workload.

Critical flicker frequency (CFF) was found to be reported in only one study of workload. Jenney, Older, and Cameron (1972) took before and after measures of CFF in a simulation of air traffic control tasks. They found decreases in CFF for low workload but less decrease for high workload. This somewhat paradoxical result could reflect an activation factor which is indexed by CFF, since before and after measures could also index fatigue or boredom (see discussion of CFF and fatigue in the next section).

Electromyographic (EMG) measures have been used in a number of workload contexts. Laville and Wisner (1965) report that the EMG of neck muscles correlated with subjective stress in a demanding, precise task better than either posture or heart rate. Wisner (1973) again showed the same effect from neck muscles. Although Spyker, *et al* (1971) failed to find EMG measures correlated with workload, Stackhouse (1976) using similar procedures found EMG from the forehead and forearm to be correlated with workload. Grip pressure has been found to increase in high workload tracking tasks (Smith, 1972; Hickok, 1973). However, grip pressure also increases as a function of the operator's effective gain.

Other psychophysiological measures of workload have included respiration and even handwriting analysis (see Wierwille and Williges, 1978). While the above techniques may sometimes show correlations, especially when used in multivariate studies such as those of Spyker (1971) and Stackhouse (1976) they have not shown the consistency necessary to consider them likely candidates for workload assessment. It will be a pleasant surprise if one of them becomes a useful, valid adjunct in the area.

FATIGUE.

The last topic to be considered involves the entire complex of changes which occur in the individual after long periods of work or wakefulness -- fatigue. Obviously, most of the measurement techniques discussed previously will be used in assessing fatigue. It is considered separately here more because it can be assessed in field situations more easily than many other questions. In fact, driving simulators and instrumented cars have been used extensively. It therefore provides a model for the utilization of the psychophysiological techniques so far discussed. In the following section, therefore, the techniques of assessment specific to fatigue study will be discussed briefly. Then, a final section will be devoted to field measurement of fatigue, emphasizing both methods and results.

Specific Measures of Fatigue. As noted above, many physiological measures can be used to assess fatigue. Heart changes, biochemical analysis, respiration, EEG, etc., can all be used, and have often been employed. Two measures should be noted briefly because they have been considered to be especially appropriate in measuring fatigue. Critical Flicker Fusion (CFF) has been discussed extensively here. Rey (1971) further presents evidence indicating that CFF changes may have their origin centrally. Further, Grandjean, *et al* (1977) report evidence indicating that CFF values show a correlation with subjective and objective indices of "fatigue" and "sleepiness". Thus, CFF appears to be a good measure for revealing the "state of fatigue due to an excessive demand on the central nervous system". For these reasons, the use of CFF in studies of fatigue should be encouraged.

The second specific technique involves measures of eye activity, particularly blink analysis and saccade velocity. Stern (1972; Stern, Beideman, and Chen, 1976) has pointed out that the form of the eyeblink changes under certain conditions affecting central alertness. Other changes occur in the movement speed and trajectory of the eyes. It would appear highly desirable to develop these measures more fully and to implement them into fatigue studies on a more routine basis.

Field Measurement of Fatigue. As noted above, fatigue has been studied in actual field situations perhaps as often as any other question of operational psychological interest. For the most part, these

studies have used automobiles, although air traffic controllers and others have been used. Emphasis in the present case is more on techniques and results than on specific applications.

EKG changes over time were found in drivers travelling continuously for 200 to 700 miles (Burns, et al, 1966). These were so atypical that the authors state that had they been seen in response to other stress, they would have been considered abnormal. Sugarman and Cozad, (1972) ran a series of road and simulator tests to determine the effects of driving time, acoustic noise, and task complexity on driver performance. EEG and heart rate measures were taken, and the car (as in most of these studies) was instrumented to record steering actions, lane position, and control activations. EEG alpha increased with time, and correlated with road position error. Small steering wheel reversals (2 degrees or less) decreased. Heart rate and theta EEG changes were used as evidence that the use of the automatic speed controller fostered decreases in alertness. The authors felt that the measures in these conditions were stable enough that a multiple regression analysis could be used to develop an index of driver alertness.

O'Hanlon (1971) actually produced such an index, based entirely on heart rate variability. This "experimental alertness indicator" (EAI) was used in a series of road tests and indicated significant heart rate variability effects as a function of driving time. In another road test, O'Hanlon and Kelley (1974) found heart rate slowing and EEG alpha slowing correlated with progressive deterioration in road tracking and vehicle control. Finally, Riemersma, et al, (1977) found progressive decrements in performance in eight hours of driving, along with decreased heart rate variability.

These studies illustrate the advantages of using field environments to generate the performance metrics which are correlated with physiological measures. At their present state of development, psychophysiological indices are in need of field validation if they are to realize their full potential. The above studies demonstrate that, even with an amorphous topic like fatigue, it is possible to design studies utilizing real systems in ways which allow for control of major variables. These systems can be employed in such a way that the condition of interest can be expected to occur (e.g., fatigue, workload, stress, etc.) and performance metrics which are direct measures of the effects of those conditions can be taken from the system. Physiological correlates then have real interpretability, and can later stand on their own right. In applications to aircraft design, especially in assessing cognitive function, such real-world tests of psychophysiological measures are becoming essential.

THE FUTURE ROLE OF PSYCHOPHYSIOLOGICAL MEASUREMENT

In the preceding pages, the power as well as the limitations of psychophysiological measurement approaches to human engineering problems have been presented. Since this is a general survey of possible contributions as well as existing capabilities, some attempt was made to present as many positive developments and prospects as possible. It is time now to take a realistic look at the future of psychophysiological measurement techniques in applied settings. Such a look will reveal that these techniques are certainly not as good, or as powerful, or as simple as we would like. For many applications, there is no reason to resort to the procedural and technical sophistication required to utilize physiological measures. They are, for many specific applications, not as good as existing procedures. On the other hand, a critical look will also reveal that these techniques, for a large number of questions facing the human engineer, are better than anything available at the present time. For such applications psychophysiological techniques require, at worst, some additional standardization and validation. At best, they offer the very high prospect of permitting objective answers at a level of specificity so far unrealizable with behavioral metrics.

The disappointing early history of psychophysiological measurements can be viewed as being due principally to attempts to utilize it too soon and too fast. We now realize, with the advantage of hindsight, that theories were much too simplistic to permit the elegant and detailed predictions and interpretations which were made. Even more importantly, technology stood on the threshold of major breakthroughs. Before these breakthroughs occurred, it was somewhat naive to expect that the kind of signals being generated could be recorded reliably. Even if they could be recorded, such signals could not yet be analyzed. Only recently emerging from the era of single variable statistics and extremely simplistic experimental designs, psychology and medicine were unable to comprehend the incredible complexity of the physiological signals being recorded. Under such conditions, it is not remarkable that physiological measurement techniques enjoyed limited success and were of limited value to the design engineer and other applied specialists. Investigators who sensed the power of these techniques are not to be disparaged for their attempts. Rather, they are to be admired for the courage and scope of insight they demonstrated, and for their ingenuity in generating even the small amount of useful data which was produced. Like a dog walking on two legs, the remarkable thing is not that he does it badly, but that he does it at all.

What should be learned from this is that application of psychophysiological techniques cannot be presented as general, universal answers to applied questions without sound theoretical and technical bases. Early researchers suffered from inadequate theory upon which to base such generalizations. Present day researchers, it must be recognized, are only slightly better off. No totally unified theory of physiological function has yet emerged which allows one to fit each measurement technique into a total model of the human. Lacking this, it would be naive and almost certainly counterproductive to present psychophysiological techniques in any single, unified framework.

On the other hand, this does not mean that such techniques do not have extensive applicability. The preceding chapters have indicated that it is now possible, using current techniques of data acquisition and analysis, to establish very high correlation and, in some cases, even causal relationships between specific behaviors or sets of behaviors and particular psychophysiological techniques. This has certainly been true in the realm of sensory function, and can be argued forcefully for cognitive function. When care is taken to establish the functional relationship between a particular psychophysiological measurement technique and well defined behavioral correlates, the measure then becomes operationally useful. If a further step can also establish the physiological basis of the psychophysiological measurement (as in the brain stem evoked response, where each peak is identified with a physical structure in the brain) then further interpretation of changes in the psychophysiological measurement is possible. When this can be done, the value of the technique is increased immeasurably. Further applications can then be hypothesized on the bases of theoretical relationships between the brain structures involved and the behavior carried out. These can then be tested empirically. Thus, it is seen that if the application of the psychophysiological technique is viewed from a specific enough level, the lack of broad theoretical base is not an insurmountable obstacle.

It should be clear from the kinds of data presented in preceding chapters that many of the techniques currently available can be used in the above way to answer questions of aircraft design and human engineering interests. These cannot be applied blindly. Simply because one measures change in the evoked response under two conditions, it is not necessary that such a change be meaningful, or that such a change has more validity than verbal reports. Only if the meaning of the change has been established empirically, either in a predictive or a causal sense, is the change of any value to the engineer. Even then, it must be established that the magnitude of this change represents a meaningful component of the total variance in behavior. Lacking this, there is no reason to recommend the use of psychophysiological techniques.

With such cautions in mind, however, we believe that a broad range of the techniques discussed in this AGARDograph have reached the above level of empirical validity. While recognizing the obvious danger of too early and too general application of basic research techniques, we believe that the human engineering community has been too timid and too slow to try these techniques in applied settings. We believe also that there has been a general lack of creativity in defining the potential utilizations of basic laboratory techniques in applied settings. Throughout this work, we have been impressed by the fact that specific techniques which demonstrated very good reliability and which could be consistently related to behaviors in the laboratory have not had adequate field trials in order to determine whether the laboratory validity extended to the real world. In some cases, there is a very high probability that such extensions will be found, and there have even been scattered examples of such applications. In other cases, this is clearly a research question which should be actively pursued by the human engineering community. In any case, it is apparent that this community can no longer ignore psychophysiological measures as applied techniques. It is probably inevitable that the use of such measures will increase dramatically. It is up to us to see that such use is carried out with an optimum balance of scientific caution and creativity.

REFERENCES

- Adey, W. R., Kado, R. T., and Walter, D. O. Computer Analysis of EEG data from Gemini flight GT-7. Aerospace Medicine, 1967, 38, 345-359.
- Andreassi, J.L., Stern, M., and Okamura, H. Visual cortical evoked potentials as a function of intensity variations in sequential blanking. Psychophysiology, 1974, 11, 336-345.
- Armington, J.C. Relations between electroretinograms and occipital potentials elicited by flickering stimuli. Doc. Ophthalmology, 1964, 18, 194-206.
- Armington, J.C. Spectral sensitivity of simultaneous electroretinograms and occipital responses. Vision Research, 1966, Supl 1, 225-233.
- Armington, J.C. The electroretinogram, the visual evoked potential, and the area-luminance relation. Vision Research, 1968, 8, 263-276.
- Armington, J.C. Simultaneous electroretinograms and evoked potentials. American Journal of Optometry, 1970, 47, 450-458.
- Armington, J.C. The electroretinogram. New York: Academic Press, 1974.
- Armington, J.C., Corwin, T.R., and Marsetta, R. Simultaneously recorded retinal and cortical responses to patterned stimuli. Journal of the Optical Society of America, 1971, 61, 1514-1521.
- Aunon, J.I. and McGillem, C.D. Detection and processing of individual components in the VEP. Psychophysiology, 1979, 16, 71-79.
- Bahill, P., Clark, M., and Stark, L. Dynamic overshoot in saccadic eye movements is caused by neurological control signal reversal. Experimental Neurology, 1975, 48, 107-122.
- Barber, H.O., and Stockwell, C.W. Manual of electronystagmography, St Louis: Mosby Co., 1976.
- Barlow, J.D. Pupillary size as an index of preference in political candidates. Perceptual and Motor Skills, 1969, 28, 587-590.
- Bartley, S.H. The Psychophysiology of Vision. In S.S. Stevens (Ed.). Handbook of Experimental Psychology. John Wiley & Sons, Inc. New York, 1951.
- Bartoshuk, A.K. Electromyographic gradients as indicants of motivation. Canadian Journal of Psychology, 1955, 9, 215-230.
- Bathien, N. Reflexes spinaux chez l'homme et niveaux d'attention. [Spinal reflexes and attention level in man.] Electroencephalography and Clinical Neurophysiology, 1971, 30, 32-37.
- Bathien, N., Bertrand, F., Gautier, H., Hugelin, A., and Vincent, J. Sensibilite comparee de tests sensitifs, moteurs, vegetatifs et respiratoires au cours de l'attention. [Comparative sensitivity of sensory, motor, vegetative and respiratory tests to levels of attention.] Journal de Physiologie (Paris), 1967, 59, 342. (abstract)
- Bathien, N., and Hugelin, A. Reflexes monosynaptiques et polysynaptiques de l'homme au cours de l'attention. [Monosynaptic and polysynaptic reflexes in man during attention.] Electroencephalography and Clinical Neurophysiology, 1969, 26, 604-612.
- Bathien, N., and Morin, C. Variations comparees des reflexes spinaux au cours de l'attention intensive et selective. [Comparing variations of spinal reflexes during intensive and selective attention.] Physiology and Behavior, 1972, 9, 533-538.
- Baumler, G., and Klausnitzer, J. Zur EEG-symptomatik von mentalen blockierungen. [EEG changes during mental blocking.] Zeitschrift fur Experimentelle und Angewandte Psychologie, 1972, 19, 196-225.
- Baumstiller, Y., and Parrot, J. Stimulus generalization and spontaneous blinking in man involved in a voluntary activity. Journal of Experimental Psychology, 1971, 88, 95-102.
- Beach, F.W. Neural and chemical regulation of behavior. In Harlow, H.F., and Woolsey, N. (Eds.), Biological And Biochemical Bases of Behavior. Madison: University of Wisconsin Press, 1958.
- Beatty, J., Greenberg, A., Diebler, W., and O'Hanlon, J. Operant control of occipital theta rhythm affects performance in a radar monitoring task. Science, 1974, 183, 871-873.
- Beatty, J. and Kaheman, D. Pupillary changes in two memory tasks. Psychonomic Science, 1966, 5, 371-372.
- Beck, E. C. Electrophysiology and behavior. In M.R. Rosenzweig and L.W. Porter (Eds.) Annual Review of Psychology, 1975, 26, 233-262.
- Beck, E.C., Dustman, R.E., and Sakai, M. Electrophysiological correlates of selective attention. In C.R. Evans and T. B. Mulholland (Eds.), Attention in Neurophysiology. London: Butterworths, 1969.
- Becker, W., Hoehne, O., Iwase, K., and Kornhuber, H.H. Cerebral and ocular muscle potentials preceding voluntary eye movements in man. Electroencephalography and Clinical Neurophysiology, 1973, Supplement 33, 99-104.

- Becker, W., Iwase, K., Jurgens, R., and Kornhuber, H.H. Bereitschaftspotential preceding voluntary slow and rapid hand movements. In W.C. McCallum and J.R. Knott (Eds.), The responsive brain. Proceedings of the Third International Congress on Event-Related Slow Potentials of the Brain. Bristol, England, 1973. Bristol: John Wright, 1976.
- Becker-Carus, C. Relationships between EEG, personality and vigilance. Electroencephalography and Clinical Neurophysiology, 1971, 30, 519-526.
- Begleiter, H., and Porjesz, B. Evoked brain potentials as indicators of decision-making. Science, 1975, 187, 754-755.
- Beh, H.C., and Nix-James, D.R. The relationship between respiration phase and reaction time. Psychophysiology, 1974, 11, 400-402.
- Bekesy, G.V. A new audiometer. Acta Oto-Laryngologica, Stockholm, 1947, 35, 411-422.
- Bergman, S. A., Wolthuis, R.A., Hoffler, G.W., and Johnson, R.L. Analysis of phonocardiographs from Apollo crews: a pre- and post-flight evaluation during LBNP. 1973 Aerospace Medicine Meeting. Proceedings, pp. 205-206.
- Bergum, O., and Lehr, J. Prediction of stimulus approach: CORE measures experiment II. Research report R66-6. Rochester, N.Y.: Xerox Corp., 1966.
- Bernstein, A.S., Taylor, K.W., and Weinstein, E. The phasic electrodermal response as a differential complex reflecting stimulus significance. Psychophysiology, 1975, 12, 158-169.
- Bickford, R.G. Properties of the photomotor response system. Electroencephalography and Clinical Neurophysiology, 1964, 17, 456.
- Bickford, R.G. Effect of facial expression on the averaged evoked response to light in man. Electroencephalography and Clinical Neurophysiology, 1967, 23, 78-79.
- Bickford, R.G., Jacobson, J.L., Cody, D.T.R. Nature of average evoked potentials to sound and other stimuli in man. Annals of the N.Y. Academy of Sciences, 1964, 112, 204-223.
- Biederman-Thorson, M., Thorson, J., and Lange, G.D. Apparent movement due to closely spaced sequentially flashed dots in the human peripheral field of vision. Visual Research, 1971, 11, 889-903.
- Billo, A.G. The influence of muscular tension on the efficiency of mental work. American Journal of Psychology, 1927, 38, 227-259.
- Blackwell, H.R., Pritchard, B.S., and Ohmart, J.G. Automatic apparatus for stimulus presentation and recording in visual threshold experiments. Journal of the Optical Society of America, 1954, 44, 322-326.
- Blakemore, C., Nachmias, J., and Sullon, P. The perceived spacial frequency shift: evidence for frequency selective neurones in the human brain. Journal of Physiology, 1970, 210, 727-750.
- Blakemore, C., and Campbell, F.W. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. Journal of Physiology, 1969, 203, 237-260.
- Blitz, P.S., Hoogstraten, J., and Mulder, G. Mental load, heart rate and heart rate variability. Psychologische Forschung, 1970, 33, 277-288.
- Boddy, J. The relationship of reaction time to brain wave period: A re-evaluation. Electroencephalography and Clinical Neurophysiology, 1971, 30, 229-235.
- Boddy, J. The psychological refractory period and vertex evoked potentials. Quarterly Journal of Experimental Psychology, 1972, 24, 175-192.
- Boiko, E.I. (Ed.) Studies in Higher Neurodynamics as Related to Problems of Psychology. Academy of Pedagogical Sciences of the RSFSR. Institute of Psychology, USSR, 1957. Published by the Israel Program for Scientific Translations and the Department of Health, Education, and Welfare, USA. Available from the office of Technical Services, U.S. Department of Commerce, Washington 24, DC.
- Bongard, M.M. Visual colorimetry by the method of replacement. Biofizika, 1957, 2, 119-123.
- Bostock, H., and Jarvis, M.J. Changes in the form of the cerebral evoked response related to the speed of simple reaction time. Electroencephalography and Clinical Neurophysiology, 1970, 29, 137-145.
- Bostrum, C., Keller, E.L., and Marg, E. A reconsideration of visual evoked potentials for fast automated ophthalmic refraction. Investigative Ophthalmology, 1978, 17, 182-185.
- Bottge, H., and Holoch, J. Ein Erkennungssystem für spektrale Muster in spontanen Elektroenzephalogramm und seine Anwendung in der Wachsamkeitsforschung. [An identification system of spectral patterns in a spontaneous electroencephalogram and its application in vigilance research.] Psychologische Beiträge, 1973, 15, 341-374. (Abstract)
- Boyce, P.R. Sinus arrhythmia as a measure of mental load. Ergonomics, 1974, 17, 177-183.
- Boyle, R.H., Dykman, R.A., Ackerman, P.T. Relationships of resting autonomic activity motor impulsivity and EEG tracings in children. Archives of General Psychiatry, 1965, 12, 314-323.

- Briggs, G.E., Fisher, R.P., Greenberg, S.N., Lyons, J.J., Peters, G.L., and Shinar, D. Multi-Task Time Sharing Requirements. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio 45433. AMRL-TR-71-105, 1972.
- Brown, A.E. Measurement of auditory thresholds. Journal of the Acoustical Society of America, 1965, 38, 86-92.
- Brown, B. Dynamic visual acuity, eye movements and peripheral acuity for moving targets. Vision Research, 1972, 12, 305-321.
- Brown, J.L. Visual sensitivity. In P. H. Mussen and M. R. Rosenzweig (Eds.), Annual Review of Psychology, Vol. 24. Palo Alto: Annual Reviews, Inc., 1973.
- Burch, N.R., Breiner, T.H., and Correll, E.G. Automatic analysis of electroencephalogram as an index of minimal changes in human consciousness. Federation Proceedings, 1955, 14, Part 1, 23.
- Burch, N.R., and Breiner, T.H. A bioelectric scale of human alertness: Concurrent recordings of the EEG and GSR. Psychiatric Research Reports, 1960, 12, 183-193.
- Burns, N.M., Baker, C. A., Simonson, E., and Keiper, C. Electrocardiogram changes in prolonged automobile driving. Perceptual and Motor Skills, 1966, 23, 210.
- Butler, S.R., and Glass, A. Interhemispheric asymmetry of contingent negative variation during numeric operations. Electroencephalography and Clinical Neurophysiology, 1971, 30, 366.
- Cael, W.W., Nash, A., and Singer, J.J. The late positive components of the human EEG in a signal detection task. Neuropsychologia, 1974, 12, 385-387.
- Caille, E. J. Psychophysiologie de la veille-radar. [Psychophysiology of radar-viewing.] Travail Humain, 1964, 28, 143-144. (Abstract)
- Callaway, E., and Halliday, R.A. Evoked response variability: Effects of age, amplitude, and methods of measurement. Electroencephalography and Clinical Neurophysiology, 1973, 34, 125-133.
- Campbell, F.W., Atkinson, J., Francis, M.R., and Green, D.M. Estimation of auditory thresholds using evoked potentials. In J.E. Desmedt (Ed.), Auditory Evoked Potentials in Man. Basel: Karger, 1977.
- Campbell, F.W., and Green, D.G. Optical and retinal factors affecting visual resolution. Journal of Physiology, 1965, 187, 427-436.
- Campbell, F.W., and Maffei, L. Electrophysiological evidence for the existence of orientation and size detectors in the human visual system. Journal of Physiology, 1970, 207, 635-652.
- Cannings, R., Borland, R.G., Hill, L.E., and Nicholson, A.N. Pitch and formant analysis of the voice in the investigation of pilot workload. In K. E. Klein (Ed.), Methods to assess workload. AGARD Conference Preprint, No. 216, 1977. (Available from: National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22151.)
- Cannon, W.B. The James-Lange theory of emotions: a critical examination and an alternative. American Journal of Psychology, 1927, 39, 106-124.
- Cannon, W.B. Bodily changes in pain, hunger, fear and rage. New York, Appleton, 1929.
- Caperell, C., and Shucard, D.W. The visual evoked potential as an index of cerebral lateralization of cognitive processing. Psychophysiology, 1977, 14, 90-91. (Abstract)
- Carmichael, L.V., and Dearborn, W.F. Reading and visual fatigue. New York: Houghton-Mifflin, 1947.
- Carriero, N.J. Physiological correlates of performance in a long duration repetitive visual task. In R. R. Mackie (Ed.), Vigilance: Theory, operational performance, and physiological correlates. New York: Plenum Press, 1977.
- Carriero, N.J., Cameron, M.W., Craig, J.D., DeManss, J.L., and Fite, J. An Annotated Bibliography of the Literature Dealing with the Physiological Correlates of Cognitive Performance. U.S. Army Engineering Laboratory, Aberdeen Proving Ground, MD, USA, 21005, June, 1978.
- Carter, B.E., Henning, G.B. The detection of gratings in narrow-band visual noise. Journal of Physiology, 1971, 219, 355-365.
- Chapman, R.M. Evoked responses to relevant and irrelevant visual stimuli while problem solving. Proceedings of the 73rd Annual Convention of the American Psychological Association, 1965, 177-178. (Summary).
- Chapman, R.M. Evoked potentials of the brain related to thinking. In F. J. McGugian and R. A. Schoonover (Eds.), The psychophysiology of thinking. New York and London: Academic Press, 1973.
- Chiles, W.D. Objective Methods. In A. H. Roscoe (Ed.), Assessing Pilot Workload. AGARD-AG-233, 1978.
- Ciganeck, L. Visually evoked potential correlates of attention and distraction in man. Psychiatria Clinica, 1969, 2, 95-108.

- Cohen, G.H., and Brown, W.K. Changes in ECG contour during prolonged $+G_z$ acceleration. Aerospace Medicine, 1969, 40, 874-879.
- Cohen, J. Cerebral Psychophysiology: the contingent negative variation. In R.F. Thompson and M.M. Patterson (Eds.), Bioelectric Recording Techniques. Part B. Electroencephalography and Human Brain Potentials. New York: Academic Press, 1974.
- Cohn, R. Differential cerebral processing of noise and verbal stimuli. Science, 1971, 172, 599-601.
- Cornsweet, T.N. Visual Perception. New York: Academic Press, 1970.
- Cox, S.G. Development of an eye point of regard recording system. Royal Aircraft Establishment, Technical Report 73-120, Farnborough, Hants, U.K., 1973.
- Cox, S.G. Memorandum on the use of the eye point of regard recording system in flight and other experiments. Royal Aircraft Establishment Technical Memorandum EP571, Farnborough, Hants, U.K., 1974.
- Cox, S.G. Flight trials to monitor pilot eye fixations during ILS approaches. Royal Aircraft Establishment Tech. Memo. FS 30, Farnborough, Hants, U.K., 1975.
- Crider, A. Electrodermal lability and vigilance performance. Psychophysiology, 1972, 9, 268. (Abstract)
- Crider, A., and Augenbraun, C.B. Auditory vigilance correlates of electrodermal response habituation speed. Psychophysiology, 1975, 12, 36-40.
- Danev, S.G., and de Winter, C.R. Expectation, attention and heart rate deceleration; some complications in the use of heart rate variability for assessing mental loads. Leiden, the Netherlands: Netherlands Institute of Preventive Medicine, TNO, 1970.
- Daniel, R.S. Alpha and theta EEG in vigilance. Perceptual and Motor Skills, 1967, 25, 697-703.
- Darrow, C.W. The electroencephalogram and psychophysiological regulation of the brain. American Journal of Psychiatry, 1946, 102, 791-798.
- Darrow, C.W. The rationale for treating the change in galvanic skin response as a change in conductance. Psychophysiology, 1964, 1, 31-38.
- Davidson, M.L. Perturbation approach to spatial brightness interaction in human vision. Journal of the Optical Society of America, 1968, 58, 1300-1309.
- Davies, D.R., and Parasuraman, R. Cortical evoked potentials and vigilance: A decision theory analysis. In R.R. Mackie (Ed.), Vigilance: Theory, Operational Performance, and Physiological Correlates. New York: Plenum Press, 1977.
- Davis, H. Mechanism of the inner ear. Annals Otology, Rhinology and Laryngology, 1968, 77, 644-655.
- Davis, H. Enhancement of evoked cortical potentials in humans related to a task requiring a decision. Science, 1964, 145, 182-183.
- Davis, H. Principles of Electric Response Audiometry. St. Louis: Beltone Institute for Hearing Research, 1976.
- Davis, H., Hirsh, S.H., Shelnutt, J., and Bowers, C. Further validation of evoked response audiometry (ERA). Journal of Speech and Hearing Research, 1967, 10, 717-732.
- Davis, J.F. Manual of Surface Electromyography. WADC Technical Report 59-184, Dec., 1959.
- Davis, R.C. Factors affecting the galvanic, reflex. Archives of Psychology of New York, 115, 1930.
- Davis, R.C. The functional significance of imagery differences, Journal of Experimental Psychology, 13, 630-661, 1932.
- Davis, R.C. Modification of the galvanic reflex by daily repetition of a stimulus. Journal of Experimental Psychology, 1934, 17, 504-535.
- Davis, R.C. Methods of measuring muscular tension. Psychological Bulletin, 1942, 39, 329-346.
- Decker, T.A., Williams, R.E., Kuether, C.L., Logar, N.D., and Wyman-Cornsweet, D. The Mark III haploscope. NASA-NGR-44-012-099 and NASA NGR-44-003-057, Baylor College of Medicine, Texas, 1975.
- Deecke, L., Becker, W., Grozinger, B., Scheid, P., and Kornhuber, H. Human brain potential preceding voluntary limb movements. Electroencephalography and Clinical Neurophysiology, 1973, Supplement 33, 87-94.
- Defayolle, M., Dinand, J.P., and Gentil, M.T. Averaged evoked potentials in relation to attitude, mental load and intelligence. In W. Singleton, J. Fox and D. Whitfield (Eds.), Measurement of Man at Work, London: Taylor and Francis, 1971.
- Delfini, L.F., and Campos, J.J. Signal detection and the "cardiac arousal cycle." Psychophysiology, 1972, 9, 484-491.

Della Casa, F., and Birkhauser, R. On the official introduction of the dark adaptation test in traffic. Ophthalmologica, 1946, 111, 315-320.

Dement, W.C. The effect of dream deprivation. Science, 1960, 131, 1705-1707.

Denison, D.M. A non-invasive technique of cardiopulmonary assessment. AGARDograph AG-143-70, 1970.

DiGiovanni, C., and Chambers, R.M. Physiological and psychologic aspects of the gravity spectrum. New England Journal of Medicine, 270, 88-94, 134-139, 1964.

Dimond, S.J. Vigilance and split-brain research. In R.R. Mackie (Ed.), Vigilance: Theory, Operational Performance, and Physiological Correlates. New York: Plenum Press, 1977.

Dimond, S., and Beaumont, G. Hemisphere function and vigilance. Quarterly Journal of Experimental Psychology, 1971, 23, 443-448.

Dimond, S.J., and Beaumont, J.G. Differences in the vigilance performance of the right and left hemispheres. Cortex, 1973, 9, 259-265.

Dixon, N.F. Apparatus for continuous recording of the visual threshold by the method of "closed loop control." Quarterly Journal of Experimental Psychology, 1958, 10, 62-63.

Donchin, E. Average evoked potentials and uncertainty resolution. Psychonomic Science, 1968, 12, 103.

Donchin, E. Discriminant analysis in average evoked response studies: The study of single trial data. Electroencephalography and Clinical Neurophysiology, 1969 b, 27, 311-314.

Donchin, E. Methodological issues in CNV research. Electroencephalography and Clinical Neurophysiology, 1973, Supplement 33, 3-17.

Donchin, E. Brain electrical correlates of pattern recognition. In G. F. Inbar (Ed.), Signal Analysis and Pattern Recognition in Biomedical Engineering. New York: John Wiley, 1975.

Donchin, E. Brain electrical activity as an index of mental workload in man-machine systems. Proceedings of the conference on Man-System Interface: Advance in Workload Study sponsored by the Air Line Pilots Association, Washington, D.C., 1978.

Donchin, E., and Cohen, L. Averaged evoked potentials and intramodality selective attention. Electroencephalography and Clinical Neurophysiology, 1967, 22, 537-546.

Donchin, E., and Cohen, L. Anticipation of relevant stimuli and evoked potentials: A reply to Naatanen. Perceptual and Motor Skills, 1969a, 29, 115-117.

Donchin, E., and Cohen, L. Further reply to Naatanen. Perceptual and Motor Skills, 1969b, 29, 270.

Donchin, E., Gerbrandt, L.A., Leifer, L., and Tucker, L. Is the contingent negative variation contingent on a motor response? Psychophysiology, 1972, 9, 178-188.

Donchin, E., Gerbrandt, L.K., Leifer, L., and Tucker, L.R. Contingent negative variation and motor responses. Electroencephalography and Clinical Neurophysiology, 1973, Supplement 33, 187-190.

Donchin, E., and Herning, R.I. A simulation study of the efficacy of stepwise discriminant analysis in the detection and comparison of event related potentials. Electroencephalography and Clinical Neurophysiology, 1975, 38, 51-68.

Donchin, E., Johnson, R., Herning, R., and Kutas, M. Covariation of the magnitude of the CNV and P300 as a function of the subject's task. In W.C. McCallum and J.R. Knott (Eds.), The responsive brain. Proceedings of the Third International Congress on Event-Related Slow Potentials of the Brain. Bristol, England, 1973, Bristol: John Wright, 1976.

Donchin, E., Kubovy, M., Kutas, M., Johnson, R., Jr., and Herning, R. Graded changes in evoked response (P300) amplitude as a function of cognitive activity. Perception and Psychophysics, 1973, 14, 319-324.

Donchin, E., Kutas, M., and McCarthy, G. Electrocorical indices of hemispheric utilization. In S. Harnad et al. (Eds.), Lateralization in the nervous system. New York: Academic Press, 1976.

Donchin, E., and Lindsley, D. B. Cortical evoked potentials and reaction times. Electroencephalography and Clinical Neurophysiology, 1965, 18, 523. (Abstract)

Donchin, E., and Lindsley, D.B. Average evoked potentials and reaction times to visual stimuli. Electroencephalography and Clinical Neurophysiology, 1966, 20, 217-223.

Donchin, E., and Sutton, S. The "psychological significance" of evoked responses: A comment on Clark, Butler, and Rosner. Communications in Behavioral Biology, 1970, 5, 111-114.

Donchin, E., Tueting, P., Ritter, W., Kutas, M., and Heffley, E. On the independence of the CNV and the P300 components of the human averaged evoked potential. Electroencephalography and Clinical Neurophysiology, 1975, 38, 449-461.

Douglas, R.E. Information processing, sinus arrhythmia and academic achievement, Monterey, California: Naval Postgraduate School, March 1972. (NTIS NO. AD-743-732)

- Doyle, J.C., Ornstein, R., and Galin, D. Lateral specialization of cognitive mode: II. EEG frequency analysis. Psychophysiology, 1974, 11, 567-578.
- Duffy, E. Activation. In N.S. Greenfield and R.A. Sternback (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart, and Winston, 1972.
- Duffy, E., and Lacey, O.L. Adaptation in energy mobilization: changes in general level of palmar skin conductance. Journal of Experimental Psychology, 1946, 36, 437-452.
- Eason, R.G., and Dudley, L.M. Effect of stimulus size and retinal focus of stimulation on visually evoked cortical responses and reaction in man. Psychonomic Science, 1971, 23, 345-347.
- Eason, R.G., White, C.T., and Bartlett, N. Effects of checkerboard pattern stimulation on evoked cortical responses in relation to check size and visual field. Psychonomic Science, 1970, 21, 113-115.
- Eckstrand, G., and Gilliland, A.R. The psychogalvanometric method for measuring the effectiveness of advertising. Journal of Applied Psychology, 1945, 32, 415-425.
- Edelberg, R. Electrical Activity of the skin. Its measurement and uses in psychophysiology. In N.S. Greenfield and R.A. Sternback (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart, and Winston, 1972.
- Edelberg, R. Electrical properties of the skin. In C.C. Brown (Ed.), Methods in Psychophysiology. Baltimore: Williams and Wilkins, 1967.
- Edwards, B.R. Pupillary size changes during information processing overload. Unpublished master's thesis, Naval Postgraduate School, Monterey, Calif., 1972. (DDC No. AD-743-754)
- Farnsworth, D. The Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color-Vision. Journal of the Optical Society of America, 1943, 33, 568-578.
- Fitts, P.M. Engineering psychology and equipment design. In S.S. Stevens (Ed.), Handbook of Experimental Psychology. Wiley and Son, New York, 1951.
- Ford, J.M., Roth, W.T., and Kopell, B. S. Auditory evoked potentials to unpredictable shifts in pitch. Psychophysiology, 1976, 13, 32-39.
- Fortuin, M.J., Hood, W.P., Sherman, M.E., and Craige, E. Determination of left ventricular volumes by ultrasound. Circulation, 1971, 44, 375-384.
- Freeman, G. L. The Energetics of Human Behavior. Ithaca: Cornell University Press, 1948.
- Fruhstorfer, H., and Bergstrom, R.M. Human vigilance and auditory evoked responses. Electroencephalography and Clinical Neurophysiology, 1969, 27, 346-355.
- Gaarder, K. Fine eye movements during inattention. Nature, 1966, 209, 83-84.
- Galambos, R., and Hecox, K. Clinical applications of the brain stem auditory evoked potentials. In J.E. Desmedt (Ed.), Auditory Evoked Potentials in Man. Basel: Karger, 1977.
- Gale, A. Some EEG correlates of sustained attention. In R.R. Mackie (Ed.), Vigilance: Theory, operational performance, and physiological correlates. New York: Plenum Press, 1977.
- Gale, A., Jones, D.M., and Smallbone, A. Short-term memory and the EEG. Nature, 1974, 248, 439-440.
- Gale, A., Jones, D.M., and Smallbone, A. Short term recall of 9-digit strings and EEG. 1. Correlates of serial position, practice, within-subject and between-subject differences in recall (auditory presentation). British Journal of Psychology (awaiting decision).
- Galin, D., and Ornstein, R. Lateral specialization of cognitive mode: An EEG study. Psychophysiology, 1972, 9, 412-418.
- Galin, D., Ornstein, R., Kocel, K., and Merrin, E. Hemispheric localization of cognitive mode by EEG. Psychophysiology, 1971, 8, 248-249. (Abstract)
- Gardner, R.M., Beltramo, J.S., and Krinsky, R. Pupillary changes during encoding, storage, and retrieval of information. Perceptual and Motor Skills, 1975, 41, 951-955.
- Geldreich, E. W. Some physiological concomitants of mental work. Psychological Monographs, 1953, 67 (8, Whole No. 358).
- Giltrow-Tyler, J.F., Crews, J., and Drasdo, N. Electoretinography with noncorneal electrodes. Investigative Ophthalmology, 1978, 17, 1124-1127.
- Ginsburg, A. Applying basic research in vision to operational Air Force problems: Explaining how we detect and identify complex objects using biologically derived filters. Air Force Office of Scientific Research. Contractors Conference Proceedings, Dayton, Ohio, 1978 (Abstract)
- Ginsburg, A. Visual information processing based upon spatial filters constrained by biological data. Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio 45433. Technical Report in press.
- Goldbeck, R.A., and Charlet, J.D. Task parameters for predicting panel layout design and operator performance. Philco-Ford Corp. Rpt WDL-TR-5480, 1974.

- Goldbeck, R.A., and Charlet, J.D. Prediction of operator work station performance. Aeronutronic-Ford Corp. rpt WDL-TR-7071, 1975.
- Goldstein, I.B. Electromyography: A measure of skeletal muscle response. In N.S. Greenfield and R.A. Sternbach (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart, and Winston, 1972.
- Goldwater, B.C. Psychological significance of pupillary movements. Psychological Bulletin, 1972, 77, 340-355.
- Gomer, F.E., and Bish, K. Evoked potential correlates of display image quality. Human Factors (In press).
- Gomer, F.E., Spicuzza, R. J., and O'Donnell, R.D. Evoked potential correlates of visual item recognition during memory-scanning tasks. Physiological Psychology, 1976, 4, 61-65.
- Graham, C.H. Visual perception. In S.S. Stevens (Ed.), Handbook of Experimental Psychology. New York: John Wiley and Sons, 1951.
- Grandjean, E., Baschera, P., Martin, E., and Weber, A. The effects of various conditions on subjective states and critical flicker frequency. In R.R. Mackie (Ed.), Vigilance: Theory, operational performance, and physiological correlates. New York: Plenum Press, 1977.
- Green, D.G. Testing the vision of cataract patients by means of laser-generated interference fringes. Science, 1970, 168, 1240-1242.
- Greene, R.L., Kille, S., and Hogan, F.A. Electrodermal measures of attention and effort to stimulus onset and offset. Psychophysiology, 1977, 14, 79. (Abstract.)
- Grossman, S.P. A Textbook of Physiological Psychology. New York: Wiley, 1967.
- Gunn, C.G., Wolf, S., Block, R.T., and Person, R.J. Psychophysiology of the cardiovascular system. In N.S. Greenfield and R.A. Sternbach (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart, and Winston, 1972, 457-482.
- Gutmann, H.E., Easterling, R.G., and Webster, R.G. The effects of flicker on performance as a function of task-loading. Albuquerque, New Mexico: Sandia Laboratories, SC-TM-72 0617, November, 1972.
- Haider, M., and Groll, E. Vigilanzleistungen, vigilanzausfalle und ihre neurophysiologischen korrelate. [Levels of vigilance, drops in vigilance levels, and related neurophysiological correlates.] Proceedings of Second International Congress on Ergonomics, Dortmund, 1964, 221-222.
- Haider, M., Spong, P., and Lindsley, D. Attention, vigilance, and cortical evoked-potentials in humans. Science, 1964 a, 145, 180-182.
- Haider, M., Spong, P., and Lindsley, D.B. Cortical evoked potentials during visual vigilance task performance. Electroencephalography and Clinical Neurophysiology, 1964 b, 17, 714.
- Hall, R.J., and Cusack, B.L. The measurement of eye behavior: critical and selected reviews of voluntary movement and blinking (Technical Memorandum 18-72). Aberdeen Proving Ground, MD: Human Engineering Laboratory, 1972. (AMCMS Code 501B.11.84100).
- Halliday, AM., and Michael, W.F. Changes in pattern evoked responses in man associated with the vertical and horizontal meridians of the visual field. Journal of Physiology, 1970, 208, 499-593.
- Harding, G.B., Stevens, E.E., and Marston, P.T. Rate of information processing and skin resistance. Journal of Experimental Psychology, 1973, 99, 306-313.
- Hardy, L. Single judgment test for red-green discrimination. Journal of the Optical Society of America, 1943, 33, 152-154.
- Harkins, S.W. Aspects of the psychobiology of attention: Visual evoked potentials and performance in a task demanding sustained attention (Doctoral dissertation, The University of North Carolina at Chapel Hill, 1974). Dissertation Abstracts International, 1975, 36, 471B-472B. (University Microfilms No. 75-17, 646.)
- Harris, J.D. Audition. In Paul H. Mussen and Mark R. Rosenzweig (Eds.), Annual Review of Psychology, 1972, 23, 313-346.
- Harter, M.R. Evoked cortical responses to checkerboard patterns: effect of check-size as a function of retinal eccentricity. Vision Research, 1970, 10, 1365-1376.
- Harter, M.R. Evoked cortical responses to the on-and off-set of patterned light in humans. Vision Research, 1971, 11, 685-695.
- Harter, M.R., and Suitt, C.D. Visually evoked cortical responses and pattern vision in the infant: a longitudinal study. Psychonomic Science, 1970, 18, 235-237.
- Harter, M.R., and White, C.T. Evoked cortical responses to checkerboard patterns: effect of check-size as a function of visual acuity. Electroencephalography and Clinical Neurophysiology, 1970, 28, 48-54.
- Hasbrook, A.H., and Rasmussen, P.G. Pilot heart rate during in-flight simulated instrument approaches in a general aviation aircraft (AM 70-7). Oklahoma City, Oklahoma: Office Aviation Medicine, Civil Aero-medical Institutue, April 1970. (NTIS No. AD-711268)

- Hassett, J. A primer of psychophysiology. San Francisco, W. H. Freeman and Co., 1978.
- Hastrup, J.L. Electrodermal lability, introversion, and vigilance decrement. Psychophysiology, 1977, 14, 111. (Abstract)
- Hayes, K.J. Wave analysis of tissue noise and muscle action potentials. Journal of Applied Physiology, 1960, 15, 749-752.
- Hazemann, P., and Lille, F. Somatosensory evoked potentials, attention and voluntary self-paced movements. In W. C. McCallum and J. R. Knott (Eds.), The responsive brain. Proceedings of the Third International Congress on Event-Related Slow Potentials of the Brain. Bristol, England, 1973. Bristol: John Wright, 1976.
- Hecht, S., Ross, S., and Muller, C. The visibility of lines and squares at high brightness. Journal of the Optical Society of America, 1947, 37, 500-507.
- Hecht, S., Schlaer, S., and Pirenne, M.H. Energy, quanta, and vision. Journal of General Physiology, 1942, 25, 819-840.
- Helander, M. Vehicle control and driving experience, a psychophysiological approach. Proceedings of the 6th Congress of the International Ergonomics Association, 1976, 335-339.
- Heron, W. The pathology of boredom. Scientific American, 1957, 196, 52-56.
- Hess, E.H. Altitude and pupil size. Scientific American, 1965, 212, 46-54.
- Hess, E.H. Pupillometrics. In F.M. Bass, C.W. King, and E.A. Pessemier (Eds.), Application of the Sciences in Marketing Management. New York: Wiley, 1968.
- Hess, E.H. Pupillometrics. In N.S. Greenfield and R.A. Sternbach (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart, and Winston, 1972.
- Hess, E. The Tell-Tale Eye. New York: Van Nostrand Runhold, 1975.
- Hess, E.H., and Polt, J. M. Pupil size as related to interest value of visual stimuli. Science, 1960, 132, 349-350.
- Hess, E.H., and Polt, J.M. Changes in pupil size as a measure of taste differences. Perceptual and Motor Skills, 1966, 23, 451-455.
- Hickok, J. H. Grip pressure as a measure of task difficulty in compensatory tracking tasks. Master's thesis, Naval Postgraduate School, Monterey, California, September, 1973.
- Hillyard, S.A. Relationships between the contingent negative variation (CNV) and reaction time. Physiology and Behavior, 1969, 4, 351-357.
- Hillyard, S.A. The CNV and the vertex evoked potential during signal detection: A preliminary report. In E. Donchin and D. B. Lindsley (Eds.), Average evoked potentials Methods, results and evaluations. Washington, DC: US Government Printing Office, 1969.
- Hillyard, S.A. The CNV and human behavior: A review. Electroencephalography and Clinical Neurophysiology, 1973, Supplement 33, 161-171.
- Hillyard, S.A. Methodological approach in CNV research. In R. F. Thompson and M. M. Patterson (Eds.), Bioelectric Recordings Techniques. Part C. Electroencephalography and Human Brain Potentials. New York: Academic Press, 1974.
- Hillyard, S.A., Hink, R.F., Schwent, V.L., and Picton, T.W. Electrical signs of selective attention in the human brain. Science, 1973, 182, 177-180.
- Hillyard, S.A., Squires, K.C., Bauer, J.W., and Lindsay, P.H. Evoked potential correlates of auditory signal detection. Science, 1971, 172, 1357-1360.
- Hink, R.F., and Hillyard, S.A. Auditory evoked potentials during selective listening to dichotic speech messages. Perception and Psychophysics, 1976, 20, 236-242.
- Holmberg, L. The distribution and reliability of certain color vision data as measured by the Pickford-Nicholson anomaloscope. Psychological Research Bulletin, 1963, 3, 7.
- Holoch, J. Wachsamkeitsprognose mittels Computeranalyse des Spontanen Elektroenzephalogrammes. [Vigilance prognosis through computer analysis of the spontaneous electroencephalogram.] Doctoral dissertation, Technische Universitaet Muenchen, 1972.
- Hope, R.E. Pupillary size change as correlated with mental activity. Unpublished master's thesis, Naval Postgraduate School, Monterey, Calif., 1971 (DDC No. AD-720697).
- Hubel, D.H., and Wiesel, T.N. Receptive fields of single neurons in the cat's striate cortex. Journal of Physiology, 1959, 148, 574-591.
- Hyndman, B. W., and Gregory, J.R. Spectral analysis of sinus arrhythmia during mental loading. Ergonomics, 1975, 18, 255-270.

- Innes, H.E. Subjective and physiological indicators of fatigue in a vigilance task. Unpublished master's thesis, Naval Postgraduate School, Monterey, Calif., 1973. (NTIS No. AD-761 503)
- Irwin, D.J. Characteristics of the Average Human Visual Evoked Response. Air Force Institute of Technology Report No GE/BE/74-7, Wright-Patterson AFB, Ohio 45433, 1974. AD-779 079.
- Irwin, D.A., Knott, J.R., McAdam, D.W., and Rebert, C.S. Motivational determinants of the "contingent negative variation." Electroencephalography and Clinical Neurophysiology, 1966, 21, 538-543.
- Irwin, D.A., Rebert, C.S., McAdam, D.W., and Knott, J.R. Slow potential changes (CNV) in the human EEG as a function of motivational variables. Electroencephalography and Clinical Neurophysiology, 1966, 21, 412-413. (Abstract)
- Jacobs, G., Color Vision. Annual Review of Psychology. Annual Reviews Inc. Palo Alto, California, 1976, 27, 63-89.
- Jacobson, E. Muscular tension and the estimation of effort. American Journal of Psychology, 1951, 64, 112-117.
- Janisse, M.P. Pupil eye and effect: A critical review of the literature since 1960. The Canadian Psychologist, 1973, 14, 311-329.
- Janisse, M.P., and Peavler, W.S. Pupillary research today: Emotion in the eye. Psychology Today, February 1974, 7, 60-63.
- Jarvilehto, T., and Fruhstorfer, H. Differentiation between slow cortical potentials associated with motor and mental acts in man. Experimental Brain Research, 1970, 11, 309-317.
- Jasper, H.H. Report of committee on methods of clinical examination in EEG: Appendix: The ten-twenty electrode system of the International Federation. Electroencephalography and Clinical Neurophysiology, 1958, 10, 371-375.
- Jenney, L.L., Older, H.J., and Cameron, B.J. Measurement of operator workload in an information processing task. Washington, DC: National Aeronautics and Space Administration, Contractor's Report NASA CR-2150, December, 1972.
- Johnson, E.P. The electrical response of the human retina during dark adaptation. Journal of Experimental Psychology, 1949, 39, 597-609.
- Johnson, L.C. A psychophysiology for all states. Psychophysiology, 1970, 6, 501-516.
- Jouvet, M. Neurophysiology of the states of sleep. Physiological Review, 1967, 47, 117-177.
- Kahneman, D., and Beatty, J. Pupil diameter and load on memory. Science, 1966, 154, 1583-1585.
- Kahneman, D., and Beatty, J. Pupillary responses in a pitch discrimination task. Perception and Psychophysics, 1967, 2, 101-105.
- Kahneman, D.; and Beatty, J., and Pollack, I. Perceptual deficit during a mental task. Science, 1967, 157, 218-219.
- Kalafat, J.D. The relationship between signal detection and magnitude of heart rate deceleration (Doctoral dissertation, University of Colorado, 1970). Dissertation Abstracts International, 1971, 32, 1242B-1243B. (University Microfilms No. 71-21, 598.)
- Kalsbeek, J.W.H. Measurement of mental work load and of acceptable load: Possible applications in industry. International Journal of Production Research, 1968, 7, 33-45.
- Kalsbeek, J.W.H. Do you believe in sinus arrhythmia? Ergonomics, 1973, 16, 99-104.
- Kalsbeek, J.W.H., and Ettema, J.H. Continuous recording of heart rate and the measurement of perceptual load. Ergonomics, 1963, 6, 306-307. (Abstract)
- Kalsbeek, J.W.H., and Ettema, J.H. L'arythmie sinusale comme mesure de la charge mentale. [Sinus arrhythmia as a measure of mental load.] Actes du Troisieme Congres de la Societe d'Ergonomie de Langue Francaise, Bruxelles: Presses Universitaires de Bruxelles, 1968, 251-258.
- Karlin, L., Martz, M. J., Brauth, E., and Mordkoff, A.M. Auditory evoked potentials, motor potentials and reaction time. Electroencephalography and Clinical Neurophysiology, 1971, 31, 129-136.
- Karlin, L., Martz, M.J., and Mordkoff, A. M. Motor performance and sensory-evoked potentials. Electroencephalography and Clinical Neurophysiology, 1970, 28, 307-313.
- Keesee, R.L. Prediction of modulation detectability thresholds for line-scan displays. Aerospace Medical Research Laboratory, TR-76-38, Wright-Patterson Air Force Base, Ohio 1976.
- Keighley, G., Clark, W.G., and Drury, D.R. Critical flicker frequency measurements on man subjected to positive acceleration on a human centrifuge. Journal of Applied Physiology, 1951, 4, 57-62.
- Kelly, D.H. Spatio-temporal frequency characteristics of color-vision mechanisms. Journal of the Optical Society of America, 1974, 64, 983-990.

- Kennedy, J.L. A possible artifact in electroencephalography. Psychological Review, 1959, 66, 347-352.
- Kinney, J.A.S., McKay, C.L., Mensch, A.J., and Luria, S.M. Techniques for analyzing differences in VERs: Colored and patterned stimuli. Vision Research, 1972, 12, 1733-1747.
- Kinney, J.A.S., and McKay, C.L. The visual evoked cortical potential as a measure of stress in naval environments: (3) the response to pattern and color. Naval Submarine Medical Research Laboratory Report Number 778, Groton, Conn, 1974(a).
- Kinney, J.A.S., and McKay, C.L. Test of color defective vision using the visual evoked response. Journal of the Optical Society of America, 1974 (b), 64, 1244-1250.
- Kinsbourne, M. The cerebral basis of lateral asymmetries in attention. Acta Psychologica, 1970, 33, 193-201.
- Kinsbourne, M. The control of attention by interaction between the cerebral hemispheres. In S. Kornblum (Ed.), Attention and performance IV. New York and London: Academic Press, 1973.
- Kinsbourne, M. The mechanism of hemispheric control of the lateral gradient of attention. In P.M.A. Rabbitt and S. Dornic (Eds.), Attention and performance V. New York: Academic Press, 1975.
- Kirst, S.P. Human information processing and EEG correlates: The effect of processing strategy and distraction on reaction time and the contingent negative variation. (Doctoral dissertation, University of California, Los Angeles, 1974.) Dissertation Abstract International, 1975, 36, 3112B. (University Microfilms No. 75-26, 965).
- Knox, G.W. Some effects of auditory stimuli on the perception of visual flicker. American Journal of Optometry, 1953, 30, 520-526.
- Kopriva, K., Horvath, M., and Stern, J.A. Eye movement recordings for the assessment of psychoactive compound: A preliminary report (unpublished manuscript, 1971).
- Kornhuber, H.H., and Deecke, L. Hirnpotentialänderungen bei Willkurbewegungen und passiven Bewegungen des Menschen: Bereitschafts potential und reafferente Potentiale. [Changes in human brain potentials following voluntary as well as passive movements: Readiness and afferent potentials.] Pfluegers Archiv fuer die Gesamte Physiologie des Menschen und der Tiere, 1965, 284, 1-17.
- Krauskopf, J. Contributions of the primary chromatic mechanisms to the generation of visual evoked potentials. Vision Research, 1973, 13, 2289-2299.
- Krippner, R.A., and Fenwick, C.A. A Digital Flight Management System Concept for Advanced Tactical Fighter Aircraft. Proceedings of the IEEE 1975 National Aerospace and Electronics Conference, Dayton, Ohio, 443-450.
- Lacey, J.I. Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome. In E. A. Rubinstein and M. A. Parloff (Eds.), Research in Psychotherapy. Washington DC: American Psychological Association, 1959.
- Lacey, J.I., and Lacey, B.C. The relationship of resting autonomic activity to motor impulsivity. In the Brain and Human Behavior. Association for Research in Nervous and Mental Disease. Vol. 36. Baltimore: Williams and Wilkins, 1958.
- Landis, C. An annotated bibliography of critical flicker-fusion phenomena. National Research Council, Vision Committee Secretariat, Ann Arbor, Michigan, 1953.
- Landis, C. Determinants of the critical flicker-fusion threshold. Physiological Review, 1954, 34, 259-286.
- Lang, P.J., Gatchel, R.J., and Simons, R.F. Electro-cortical and cardiac rate correlates of psychophysical judgment. Psychophysiology, 1975, 12, 649-655.
- Laville, A., and Wisner, A. Etude electromyographique des muscles de la nuque au cours d'une tache de precision. [An electromyographic study of the neck muscles during a precision task.] Journal de Physiologie, 1965, 57, 260.
- Laycock, J.A. Bibliography on eye movements. Royal Aircraft Establishment Tech. Memo HFG 127, Farnborough, Hants, U.K., 1974.
- Lazarus, R.S., and McCleary, R.A. Autonomic discrimination without awareness: a study of subception. Psychological Review, 58, 113-122, 1951.
- Lee, W.M. Spontaneous electrodermal fluctuations as an index of variability in attention and cue utilization (Doctoral dissertation, The Catholic University of America, 1969). Dissertation Abstracts International, 1969, 30, 1885B-1886B. (University Microfilms No. 69-16, 329)
- Levinson, J.B., and Fenz, W.D. The effects of motivation on cardiac activity. Journal of Psychology, 1971, 78, 11-26.
- Lewis, D.H., and Duane, T.D. Electroretinogram in man during blackout. Journal of Applied Physiology, 1956, 9, 105-110.
- Licklider, J.C.R. Basic Correlates of the Auditory Stimulus. In the S.S. Stevens (Ed.), Handbook of Experimental Psychology. John Wiley and Sons, Inc.; New York. Chapman and Hall, Ltd. London, 1951, 985-1039.

- Licklider, J.C.R., and Miller, George A. The Perception of Speech. In S. S. Stevens (Ed.), Handbook of Experimental Psychology, John Wiley and Sons, Inc., New York, Chapman and Hall, Ltd. London, 1951, 1040-1074.
- Lindsley, D.B. Emotion. In S.S. Stevens (Ed.), Handbook of Experimental Psychology, John Wiley, New York, 1951.
- Lindsley, D.B. Basic perceptual processes and the EEG. Psychiatric Research Reports, 1956, 6, 161-170.
- Lippold, D.C.J., and Novotry, G.E.K. Is alpha rhythm an artifact? Lancet, 1970, 1, 976-979.
- Lockhart, J.M. Ambient temperature and the flicker fusion threshold. Journal of Experimental Psychology, 1971, 87(3), 314-319.
- Lowenstein, O., and Loewenfeld, I.E. Electronic pupillography: a new instrument and some clinical applications. Archives of Ophthalmology, 1958, 59, 352-363.
- Lubin, A., Moses, J.M., Johnson, L.C., and Naitoh, P. The recuperative effects of REM sleep and stage 4 sleep on human performance after complete sleep loss: Experiment 1. Psychophysiology, 1974, 11, 133-146.
- Luby, E.D., Frohman, C.E., Grisell, J.L., Lenzo, J.E., and Gottlieb, J.S. Sleep deprivation: Effects on behavior, thinking, motor performance, and biological energy transfer systems. Psychosomatic Medicine, 1960, 22, 182-192.
- Luckiesh, M., and Moss, F.K. Reading as a Visual Task. New York: Van Nostrand, 1942.
- Luczak, H., and Laurig, W. An analysis of heart rate variability. Ergonomics, 1973, 16, 85-97.
- Mackworth, J.F. Vigilance and Attention: A Signal Detection Approach. Baltimore, MD.: Penguin Books, 1970.
- Malmo, R.B. Activation: A neurophysiological dimension. Psychological Review, 1959, 66, 367-386.
- Marg, E., Freeman, D.N., Peltzman, P., and Goldstein, P.J. Visual acuity development in human infants: Evoked potential measurements. Investigative Ophthalmology, 1976, 15, 150-153.
- Materson, R.B., and Berkley, M. Brain Functions. In P.H. Mussen, and M.R. Rosenzweig (Eds.), Annual Review of Psychology. Palo Alto, California, 1974.
- Max, L.W. An experimental study of the motor theory of consciousness. IV. Action current responses of the deaf during awakening, kinesthetic imagery, and abstract thinking. Journal of Comparative Psychology, 1937, 24, 301-344.
- McAdam, D.W. The contingent negative variations. In R.F. Thompson and M.M. Patterson (Eds.), Bioelectric recording techniques. New York: Academic Press, 1974.
- McAdam, D.W., and Rubin, E.H. Readiness potential, vertex positive wave, contingent negative variation and accuracy of perception. Electroencephalography and Clinical Neurophysiology, 1971, 30, 511-517.
- McAdam, D.W., and Seales, D.M. Bereitschaftspotential enhancement with increased level of motivation. Electroencephalography and Clinical Neurophysiology, 1969, 27, 73-75.
- McCallum, C. The contingent negative variation as a cortical sign of attention in man. In C.R. Evans and T.B. Mulholland (Eds.), Attention in neurophysiology. London: Butterworths, 1969.
- McDonald, D.G., Johnson, L.C., and Hord, D.J. Habituation of the orienting response in alert and drowsy subjects. Psychophysiology, 1964, 1, 163-173.
- McGinnies, E. Emotionality and perceptual defense. Psychological Review, 1949, 56, 244-251.
- McLaughlin, S.C., Jr. An automatic recording visual adaptometer. Journal of the Optical Society of America, 1954, 44, 312-314.
- Mecocci, L., and Spinelli, D. The effects of spatial frequency adaptation on human evoked potentials. Vision Research, 1976, 16, 477-479.
- Melzack, R. The Puzzle of Pain. New York: Basic Books, 1973.
- Merchant, J., Morrisette, R., and Porterfield, J.L. Remote measurement of eye direction allowing subject motion over one cubic foot of space. IEEE Transactions on Biomedical Engineering, 1974, BME-21, 309-317.
- Messulam, M.M., and Perry, J. The diagnosis of love-sickness: Experimental psychophysiology without the polygraph. Psychophysiology, 1972, 9, 546-551.
- Miller, E.F., II. Counterrolling of the human eyes produced by head tilt with respect to gravity. Acta Othlaryngologia, Stockholm, 1961, 54, 479-501.
- Miller, E.F., II. Ocular counterrolling. In R. J. Wolfson (Ed.), The vestibular system and its diseases. Philadelphia: University of Pennsylvania Press, 1966.

- Miller, E.F., II, and Graybiel, A. Otolith function as measured by ocular counterrolling. In: The Role of the Vestibular Organs in the Exploration of Space. NASA SP-77. Washington, DC, National Aeronautics and Space Administration, 1965.
- Miller, E.F., II, Graybiel, A., and Kellogg, R.S. Otolith organ activity within earth standard, one-half standard and zero gravity environments. Aerospace Medicine, 1966, 37, 399-403.
- Miller, H.L. Effect of auditory stimulation on critical flicker fusion frequency. Journal of Experimental Psychology, 1969, 81, 365-369.
- Miller, N.D. Evaluation of non-invasive physiologic measurement techniques. USAF School of Aerospace Medicine, Brooks AFB, Texas 78217, SAM-TR-76-34, 1976.
- Misiak, H., and Rizy, E.F. The effects of dextroamphetamine and phenobarbital on a simplified standardized CFF measure. Psychopharmacologia, 1968, 13, 346-353.
- Misiak, H., Zenhausen, R., and Salafia, W.R. Continuous temporal evaluation of the effect of meprobamate on critical flicker frequency in normal subjects. Psychopharmacologia, 1966, 9, 457-461.
- Mittleman, B., Wolff, H.G. Affective states and skin temperature: Experimental study of subjects with "cold hands" and Raynaud's syndrome. Psychosomatic Medicine, 1939, 1, 271-292.
- Moise, S. Neurophysiological metrics in human performance. Air Force Office of Scientific Research Contractor's Conference Proceedings, Dayton, Ohio, 1978. (Abstract)
- Monesi, F., Ravaccia, F. A study of behaviour during a trial of vigilance in non-piloting personnel. In B.O. Hartman (Ed.), Higher mental functioning in operational environments. AGARD Conference Proceedings, No. 181, 1976. (Available from: National Aeronautics Space Administration, Langley Field, Virginia 23365, Attn: Report Distribution and Storage Unit)
- Morgan, A.H., MacDonald, H., and Hilgard, E.R. EEG alpha: Lateral asymmetry related to task, and hypnotizability. Psychophysiology, 1974, 11, 275-282.
- Morrell, L.K. EEG frequency and reaction time--a sequential analysis. Neuropsychologia, 1966, 4, 41-48.
- Morris, C.J. Electroencephalographic and evoked potential correlates of reaction time and visual discrimination performance. Psychonomic Science, 1971, 23, 193-195.
- Moushegian, G. The frequency following potentials in man. In J.E. Desmedt (Ed.), Auditory Evoked Potentials in Man. Basel: Karger, 1977.
- Moushegian, G., Rupert, A.L., and Stillman, R.D. Scalp recorded early responses in man to frequencies in the speech range. Electroencephalography and Clinical Neurophysiology, 1973, 35, 665-667.
- Mulder, G., and Mulder-Hajonides van der Meulen, W.R.E.H. Mental load and the measurement of heart rate variability. Ergonomics, 1973, 16, 69-83.
- Mulholland, T.B. Automatic control of visual displays by the attention of the human viewer. In C.M. Williams and J.L. Debes (Eds.), First National Conference on Visual Literacy. New York: Pitman, 1970.
- Mulholland, T. Objective EEG methods for studying covert shifts of visual attention. In F. J. McGugian and R.A. Schoonover (Eds.), The Psychophysiology of Thinking. New York and London: Academic Press, 1973.
- Mulholland, T.B., and Peper, E. Occipital alpha and accommodative vergence, pursuit tracking, and fast eye movements. Psychophysiology, 1971, 8, 556-575.
- Mulholland, T., and Runnals, S. Increased occurrence of EEG alpha during increased attention. Journal of Psychology, 1962 a, 54, 317-330.
- Mulholland, T., and Runnals, S. A stimulus-brain feedback system for evaluation of alertness. Journal of Psychology, 1962 b, 54, 69-83.
- Newton, F.A., and Bird, B.L. High frequency EEG correlates of cognition. Psychophysiology, 1977, 14, 83-84. (Abstract)
- Nirenberg, L.M., Haber, J., and Moise, S. A high-speed algorithm for computing conditional probabilities of substrings of sequentially observed data. Behavioral Research Methods and Instrumentation, 1973, 5, 291-294.
- Norman, D.A. Memory and Attention: An Introduction to Human Information Processing. New York: J. Wiley and Sons, 1969.
- Northern, J.L., and Downs, M.P. Hearing in Children. Baltimore, MD.: Williams and Wilkins Co., 1974.
- Obrist, P.A. The cardiovascular-behavioral interaction as it appears today. Psychophysiology, 1976, 13, 95-107.
- O'Donnell, R. D. Physiological and performance measures in short-term sensory restriction. Unpublished Ph.D. dissertation. Fordham University, New York, New York, 1970.
- O'Donnell, R.D. Secondary Task Assessment of Cognitive Workload in Alternative Cockpit Configurations. Conference on Higher Mental Functioning in Operational Environments, AGARD-CPP/181, Ankara, Turkey, 1975.

- O'Donnell, R.D., Chikos, P., and Theodore, J. Effect of carbon monoxide exposure on human sleep and psychomotor performance. Journal of Applied Physiology, 1971, 31, 513-518.
- Offenloch, K. Neurophysiological assessment of functional states of the brain. In K.E. Klein (Ed.), Methods to assess workload. AGARD Conference Preprint, No. 216, 1977. (Available from: National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22151.) (Abstract)
- O'Hanlon, J.F., Jr. Heart rate variability: A new index of driver alertness fatigue (Report No. 1712-1). Goleta, Calif.: Human Factors Research, Inc., 1971.
- O'Hanlon, J.F., Jr., and Beatty, J. Concurrence of electroencephalographic and performance changes during a simulated radar watch and some implications for the arousal theory of vigilance. In R.R. Mackie (Ed.), Vigilance: Theory, operational performance, and physiological correlates. New York: Plenum Press, 1977.
- O'Hanlon, J.F., Jr., and Horvath, S.M. Physiological reactions of men performing monotonous work. Physiologist, 1969, 12, 319. (Abstract)
- O'Hanlon, J.F., Jr., and Kelly, G.R. A psychophysiological evaluation of devices for preventing lane drift and run-off-road accidents (Report No. 1736-F). Goleta, Calif.: Human Factors Research, Inc., 1974.
- Ohzu, H., and Enoch, J.M. Optical modulation by the isolated human fovea. Vision Research, 1972, 12, 245-251.
- Ornstein, R.E. Right and left thinking. Psychology Today, 1973, 6 (12), 87-92.
- Osterhammel, P.A., Davis, H., Wier, C.C., et al. Adult auditory evoked vertex potentials in sleep. Audiology, 1973, 12, 116-128.
- Otto, D.A., Benignus, V.A., Ryan, L.J., and Leifer, L.J. Slow potential components of stimulus, response and preparatory processes in man. In J.E. Desmedt (Ed.), Progress in Clinical Neurophysiology. Vol. 1. Attention, voluntary contraction and event-related cerebral potentials. Basel, Switzerland: Karger, 1977.
- Paivio, A. Latency of verbal association and imagery to noun stimuli as a function of abstractness and generality. Canadian Journal of Psychology, 1966, 20, 378-387.
- Paivio, A., and Simpson, H.M. The effect of word abstractness and pleasantness on pupil size during an imagery task. Psychonomic Science, 1966, 5, 55-57.
- Parikh, N.I. A study of blood pressure and sinus arrhythmia as measures of mental load in a visual inspection task (Doctoral dissertation, The University of Oklahoma, 1971). Dissertation Abstracts International, 1971, 32, 311B-312B. (University Microfilms No. 71-17, 047)
- Perry, N.W., and Childers, D.G. The human visual evoked responses: Method and theory. Springfield, Ill.: Thomas, 1969.
- Pieron, H. Vision in intermittent light. In W.D. Neff (Ed.), Contributions to Sensory Physiology. Academic Press: New York, 1965.
- Pinson, E.A., and Chapanis, A. AML portable radium plaque night vision tester. Air Surg. Bulletin, 1945, 2, 285.
- Plutchik, R. The psychophysiology of skin temperature: a critical review. Journal of General Psychology, 1956, 55, 249-268.
- Pooch, G.K. Information processing vs pupil diameter. Perceptual and Motor Skills, 1973, 37, 1006-1002.
- Pooch, G.K., and Noel, C.E. Effects of layout of workplace on pupillary diameter. Perceptual and Motor Skills, 1975, 40, 304.
- Pooch, G.K., Tuck, G.A., and Tinsley, J.H. Physiological correlates of visual monitoring. Perceptual and Motor Skills, 1969, 29, 334.
- Poon, L.W., Thompson, L.W., and Marsh, G.R. Average evoked potential changes as a function of processing complexity. Psychophysiology, 1976, 13, 43-49.
- Porges, S.W. Heart rate variability as a predictor of attentional performance. Psychophysiology, 1971, 8, 273. (Abstract)
- Porges, S.W. Respiratory and heart rate indices of reaction time (Doctoral dissertation, Michigan State University, 1970). Dissertation Abstracts International, 19 593B-594B. (University Microfilms No. 71-18, 274)
- Porges, S.W. Heart rate variability and deceleration as indexes of reaction time. Journal of Experimental Psychology, 1972, 92, 103-110.
- Porges, S.W. Heart rate variability: An autonomic correlate of reaction time performance. Bulletin of the Psychonomic Society, 1973, 1, 270-272.
- Poulton, E.C., and Gregory, R.L. Blinking during visual tracking. Quarterly Journal of Experimental Psychology, 1952, 4, 57-65.

Pullen, R.L. (Ed.). Medical diagnosis. Philadelphia, Saunders, 1944.

Rapin, I., and Schimmel, H. Assessment of auditory sensitivity in infants and in uncooperative handicapped children by using the late components of the average auditory evoked potential. In J.E. Desmedt (Ed.), Auditory Evoked Potentials in Man. Basel: Karger, 1977.

Raskin, D.C. Attention and arousal. In W. Prokasy and D.C. Raskin (Eds.), Electrodermal Activity in Psychological Research. New York: Academic Press, 1973.

Rasmussen, C.T. Evoked potential correlates of accuracy and confidence in a visual detection task (Doctoral dissertation, University of Arizona, 1972). Dissertation Abstracts International, 1972, 33, 2842B. (University Microfilms No. 72-31, 847)

Reading, V.M. Analysis of eye movement responses and dynamic visual acuity. Pflugers Arch., 1972, 333, 27-34.

Rebert, C.S. The effect of reaction time feedback on reaction time and contingent negative variation. Psychophysiology, 1972, 9, 334-339.

Rebert, C.S. Further analysis of CNV and reaction time. Electroencephalography and Clinical Neurophysiology, 1974, 32, 435-436. (Abstract)

Rebert, C.S. Functional cerebral asymmetry and performance: I. Reaction time to words and dot patterns as a function of EEG alpha asymmetry. Behavioral Neuropsychiatry, 1977a.

Rebert, C.S. Functional cerebral asymmetry and performance: II. Individual differences in reaction time to word and pattern stimuli triggered by asymmetric alpha bursts. Behavioral Neuropsychiatry, 1977b.

Rebert, C.S., and Low, D.W. Differential hemispheric activation during complex visuomotor performance: PONG. Unpublished manuscript, Stanford Research Institute, 1977c.

Rebert, C.S., and Tecce, J.J. A summary of CNV and reaction time. Electroencephalography and Clinical Neurophysiology, 1973, Supplement 33, 173-178.

Rechtschaffen, A., and Kales, A. A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects. National Institute of Health Publication 204, 1968.

Regan, D. Some characteristics of average steady state and transient response evoked by modulated light. Electroencephalograph and Clinical Neurophysiology, 1966, 20, 238-248.

Regan, D. Evoked potentials in psychology, sensory physiology, and clinical medicine. London: Chapman and Hall, 1972.

Regan, D. Rapid objective refraction using evoked brain potentials. Investigative Ophthalmology, 1973, 12, 669-679.

Regan, D.M. Color coding of pattern responses in man investigated by evoked potential feedback and direct plot techniques. Vision Research, 1975a, 15, 175-183.

Regan, D.M. Recent advances in electrical recording from human brain. Nature, 1975b, 253, 401-407.

Regan, D. Steady-state evoked potentials. Journal of the Optical Society of America, 1977a, 67, 1475-1488.

Regan, D. New methods for neurological assessment: Overview, Proceedings of the San Diego Biomedical Symposium, Vol. 16, New York: Academic Press, 1977b.

Regan, D. Speedy assessment of visual acuity in amblyopia by evoked potential method. Ophthalmologica, 1977c, 175, 159-164.

Regan, D., and Beverly, K.I. Relation between the magnitude of flicker sensation and evoked potential amplitude in man. Perception, 1973, 2, 61-65.

Reger, S.N., and Voots, R.J. Design and construction of an automatic, self-testing recording, pulse-tone audiometer. USAF School of Aviation Medicine Report, 1957, No. 57-64, 20.

Rey, P. The interpretation of changes in critical fusion frequency. In W.T. Singleton, J. Fox and D. Whitfield (Eds.), Measurement of Man at Work. London: Taylor and Francis, 1971.

Rice, B. Rattlesnakes, french fries, and pupillometric oversell. Psychology Today, 1974, 7, 55-59.

Riemersma, J.B.J., Sanders, A.F., Wildervanck, C., and Gaillard, A. W. Performance decrement during prolonged night driving. In R.R. Mackie (Ed.), Vigilance: Theory, operational performance, and physiological correlates. New York: Plenum Press, 1977.

Riggs, L.A., and Wooten, B.R. In D. Jameson, and L.M. Hurvich (Eds.), Handbook of Sensory Physiology, Vol. 7, Part 4. Springer-Verlag, Berlin, 1972.

Rintelmann, W.F., and Harford, E.R. Type V. Bekesy pattern: Interpretation and clinical utility. Journal of Speech and Hearing Research, 1967, 10, 733-744.

Roman, J.A., Warren, B.H., and Graybiel, A. The sensitivity of the semicircular-canal during weightlessness. Aerospace Medicine, 1963, 34, 1085-1089.

Roman, J., Older, H., and Jones, W.L. Flight research program: VII. Medical monitoring navy carrier pilots in combat. Aerospace Medicine, 1967, 38, 133-139.

- Rowland, W.M., and Mandelbaum, J. Testing night vision. Air Surg. Bulletin, 1944, 1, No. 8, 14-15.
- Ruch, T.C. Motor systems. In S.S. Stevens (Ed.), Handbook of Experimental Psychology. New York: Wiley, 1951.
- Ryan, T.A., Bitterman, M.E., and Cottrell, C.L. Relation of critical fusion frequency to fatigue in reading. Illumination Engineering, N.Y., 1953, 48, 385-391.
- Sayers, B. McA. The analysis of cardiac interbeat interval sequences and the effects of mental work load. Royal Society of Medicine, 1971, 64, 707-710.
- Sayers, B. McA. Analysis of heart rate variability. Ergonomics, 1973, 16, 17-32.
- Sayers, B. McA. Physiological consequences of informational load and overload. In P.H. Venables and M.J. Christie (Eds.), Research in Psychophysiology, London: Wiley, 1975.
- Schacke, G., and Weitowitz, R. Zur Erfassung Psycho-Vegetativer Belastungen am Arbeitsplatz mit Hilfe der Radiotelemetrie. [Radiotelemetric measurement of psychovegetative stress at work.] Arbeitsmedizin-Sozialmedizin-Arbeitshygiene, 1971, 6, 259-262.
- Schaefer, H. Experiments on the scale of variations of the threshold of the human fovea for red light. Klin. Mbl. Augenheilk, 1949, 115, 690-699.
- Schafer, E. W.P. Brain responses to television reflect program interest. Psychophysiology, 1977, 14, 115. (Abstract)
- Schechter, G., and Buchsbaum, M. The effects of attention, stimulus intensity, and individual differences on the average evoked response. Psychophysiology, 1973, 10, 392-400.
- Shucard, D.W., Shucard, J.L., and Thomas, D.G. The auditory evoked potential and the assessment of cerebral lateralization of cognitive processing. Psychophysiology, 1977, 14, 91. (Abstract)
- Schultz, D.P. Sensory reaction. Effects on behavior. New York: Academic Press, 1965.
- Schwartz, G.E. Operant conditioning of human cardiovascular integration and differentiation through feedback and reward. Unpublished doctoral dissertation, Harvard University, 1971.
- Scow, J., Krasno, L.R., and Ivy, A.C. The immediate and accumulative effect on psychomotor performance of exposure to hypoxia, high altitude and hyperventilation. Journal of Aviation Medicine, 1950, 21, 79-81.
- Sears, R. Psychogalvanic responses in arithmetical work. Archives of Psychology of New York, No 155, 1933.
- Seashore, R.H. Work and motor performance. In S.S. Stevens (Ed.), Handbook of Experimental Psychology: Wiley and Son, New York, 1951.
- Sekuler, R., and Littlejohn, J. Tilt aftereffect following very brief exposures. Vision Research, 1974, 14.
- Sem-Jacobsen, C.W. Black out and unconsciousness revealed by airborne testing of fighter pilots. Aerospace Medicine, 1961, 32, 247.
- Shapiro, D., and Schwartz, G.B. Psychophysiological contributions to social psychology. In P.H. Mussen and M.R. Rosenzweig (Eds.), Annual Review of Psychology, Palo Alto: Annual Reviews, Inc., 21, 1970, 87-112.
- Simonov, P.V. and Frolov, M.V. Analysis of the human voice as a method of controlling emotional state: Achievements and goals. Aviation, Space, and Environmental Medicine, 1977, 48, 23-25.
- Simons, R.F., and Lang, P.J. Psychophysical judgment: Electro-cortical and heart rate correlates of accuracy and uncertainty. Biological Psychology, 1976, 4, 51-64.
- Simpson, H.M., and Hale, S.M. Pupillary changes during a decision-making task. Perceptual and Motor Skills, 1969, 29, 495-498.
- Simpson, H.M., and Paivio, A. Effects on pupil size of manual and verbal indicators of cognitive task fulfillment. Perception and Psychophysics, 1968, 3, 185-190.
- Sloan, L.L. Selection of color vision tests for the Army Air Forces. A summary of studies made at the Army Air Forces School of Aviation Medicine. Archives of Ophthalmology, Chicago, 1946, 36, 263-283.
- Sloan, L.L. Measurement of visual acuity. A.M.A. Archives of Ophthalmology, 1951, 45, 704-725.
- Sloan, L.L. New test charts for the measurement of visual acuity at far and near distances. American Journal of Ophthalmology, 1959, 48, 807-813.
- Smith, R.A. Adaptation of visual contrast sensitivity to specific temporal frequencies. Vision Research, 1970, 10, 275-279.
- Smith, R.A. Studies of temporal frequency adaptation in visual contrast sensitivity. Journal of Physiology, 1971, 216, 351-352.

- Smith, W. The measurement of emotion. London: Paul, 1922.
- Smith, W.S., Jr. Effects of neuromuscular tension in the use of an isometric hand controller. Monterey, California: U.S. Naval Postgraduate School, Master's thesis, December, 1972.
- Snyder, H.L. Prediction of modulation detectability thresholds for line-scan displays. Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio 45433. AMRL-TR-76-38, 1976.
- Sokol, S., and Riggs, L.A. Electrical and psychophysical responses of the human visual system to periodic variation of luminance. Investigative Ophthalmology, 1971, 10, 171-180.
- Spekreijse, H. Analysis of EEG responses in man evoked by sine wave modulated light. Hague, Netherlands: W. Junk, 1966.
- Spekreijse, H. Contrast Evoked Responses in Man. Vision Research, 1973, 13, 1577-1601.
- Spence, D.P., and Lugo, M. Cardiac change as an index of attention. Journal of Abnormal Psychology, 1973, 81, 289-295.
- Spicuzza, R.J., Pinkus, A.R., and O'Donnell, R.D. Development of performance assessment methodology for the digital avionics information system. Dayton, Ohio. Systems Research Laboratories, Final Report, August, 1974.
- Spyker, D.A., Stackhouse, S.P., Khalafalla, A.S., and McLane, R.C. Development of techniques for measuring pilot workload. Washington, D.C.: National Aeronautics and Space Administration, Contractor's Report NASA CR-1888, November, 1971.
- Squires, K.C., and Donchin, E. Beyond averaging: The use of discriminant functions to recognize event related potentials elicited by single auditory stimuli. Electroencephalography and Clinical Neurophysiology, 1976, 41, 449-459.
- Squires, K.C., Hillyard, S.A., and Lindsay, P.H. Cortical potentials evoked by confirming and disconfirming feedback following an auditory discrimination. Perception and Psychophysics, 1973, 13, 25-31.
- Squires, K.C., Petuchowski, S., Wickens, C., and Donchin, E. Sequential effects on event-related potentials: A comparison of auditory and visual modalities. Psychophysiology, 1977, 14, 96. (Abstract)
- Squires, K.C., Wickens, C., Squires, N.K., and Donchin, E. The effect of stimulus sequence on the waveform of the cortical event-related potential. Science, 1976, 193, 1142-1146.
- Squires, N.K., Duncan-Johnson, C.C., Squires, K.C., and Donchin, E. Cross modal sequential effects on the P300 component of the event-related potentials. Psychophysiology, 1977, 14, 96. (Abstract)
- Squires, R.D., Jensen, R.E., Sipple, W.C., and Gordon, J.J. Electroencephalographic changes in human subjects during blackout produced by positive acceleration. NADC-MA-6402, Naval Air Development Center, Johnstown, Pa., April, 1964.
- Stackhouse, S. Workload evaluation of LLNO display. Minneapolis, Minnesota: Honeywell, 7201-3408, October, 1973.
- Stackhouse, S.P. The measurement of pilot workload in manual control system. Minneapolis, Minnesota: Honeywell, Inc., F0398 FRI, January, 1976.
- Stennett, R.G. The relationship of performance level to level of arousal. Journal of Experimental Psychology, 1957, 54, 54-61.
- Stern, J.A. Eye movements, reading, and cognition. In J.W. Senders, D.F. Fisher, and R.A. Monty (Eds.), Eye movements and the higher psychological functions. Hillsdale: Lawrence Erlbaum and Assoc., 1978.
- Sternberg, S. The discovery of processing stages: Extension of Donder's method. In W.G. Koster (Ed.), Attention and Performance II. Amsterdam; North Holland, 1969.
- Sternberg, S. Recent research. Quarterly Journal of Experimental Psychology, 1975, 27, 1-32.
- Storm, W.F. Vigilance, habituation, and the cortical evoked response (Doctoral dissertation, University of Missouri, Columbia, 1970). Dissertation Abstracts International, 1970, 31, 3036B-3037B. (University Microfilms No. 70-20, 820)
- Strasser, H. Herzfrequenz- und Sinusarrhythmie-Messungen zur Frage der Objektivierung von Konzentration und Beanspruchung in Regelaufgaben. [Heart rate and sinusoidal arrhythmia in view of objectivating concentration and stress in control tasks.] Paper presented at the 14th meeting of the committee "Anthropotechnik" of the Deutsche Gesellschaft für Luft- und Raumfahrt, Munich, November 1973. [Also abstracted in Scientific and Technical Aerospace Reports (Report No. DLR-Mitt 74-19).]
- Sugarman, R.C., and Cozad, C.P. Road tests of alertness variables (ZM-5019-B-1). Buffalo, New York: Calspan Corp. November 1972. (NTIS No. PB-215 450)
- Surwillo, W.W. The relation of decision time to brain wave frequency and to age. Electroencephalography and Clinical Neurophysiology, 1964, 16, 510-514.
- Surwillo, W.W. Human reaction time and endogenous heart rate changes in normal subjects. Psychophysiology, 1971, 8, 680-682.

Surwillo, W.W. Reaction-time variability, periodicities in reaction-time distributions, and the EEG gating-signal hypothesis. Biological Psychology, 1975, 3, 247-261.

Suzuki, T. (Ed.), Electrophysiological measurements of human auditory function. Acta oto-larongol. Supplement 252 (1969).

Taval, M.E. Clinical phonocardiography and external pulse recording. Chicago: Yearbook Medical Publishers, Inc., 1968.

Tecce, J.J. Attention and evoked potentials in man. In D.I. Mostofsky (Ed.), Attention: Contemporary theory and analysis. New York: Appleton-Century-Crofts, 1970.

Tecce, J.J. Contingent negative variation (CNV) and psychological processes in man. Psychological Bulletin, 1972, 77, 73-108.

Tepas, D.I., and Armington, J.C. Electroretinograms from noncorneal electrodes. Investigative Ophthalmology, 1962, 1, 784-786.

Thackray, R.I., Bailey, J.P., and Touchstone, R.M. Physiological, subjective, and performance correlates of reported boredom and monotony while performing a simulated radar control task. In R.R. Mackie (Ed.), Vigilance: Theory, operational performance, and physiological correlates. New York: Plenum Press, 1977.

Thackray, R.I., Jones, K.N., and Touchstone, R.M. Personality and physiological correlates of performance decrement on a monotonous task requiring sustained attention. British Journal of Psychology, 1974, 65, 351-358.

Thomas, H.E. Handbook of biomedical instrumentation and measurement. Reston, Va.: Reston Publishing Co., Inc., 1974.

Thompson, L.W., and Botwinick, J. Age differences in the relationship between EEG arousal and reaction time. Journal of Psychology, 1968, 68, 167-172.

Thompson, L.W., and Botwinick, J. Stimulation in different phases of the cardiac cycle and reaction time. Psychophysiology, 1970, 7, 57-65.

Thompson, R.F. Foundations of Physiological Psychology. New York: Harper and Row, 1967.

Tinker, M.A. Recent studies of eye movements in reading. Psychological Bulletin, 1958, 55, 215-231.

Tinsley, J.H. An investigation of physiological correlates of vigilance performance. Monterey, California: Naval Postgraduate School, October 1969. (NTIS No. AD-706054)

Travis, R.C., and Kennedy, J.L. Prediction and automatic control of alertness. I. Control of lookout alertness. Journal of Comparative Physiological Psychology, 1947, 40, 457-461.

Travis, R.C., and Kennedy, J.L. Prediction and control of alertness. III. Calibration of the alertness indicator and further results. Journal of Comparative Physiological Psychology, 1949, 42, 42-57.

Tursky, B. The indirect recording of human blood pressure. In P.A. Obrist, A.H. Black, J. Brener, and L.V. DiCara (Eds.), Cardiovascular Psychophysiology. Chicago: Aldine, 1974a.

Tursky, B. Recording of human eye movements. In R.F. Thompson and M.M. Patterson (Eds.), Bioelectric Recording Techniques, Part C. New York: Academic Press, 1974b.

Tursky, B., Shapiro, D., and Schwartz, G.E. Automated constant cuff pressure system to measure average systolic and diastolic pressure in man. Transactions on Bio-Medical Engineering, 19, 271-275, 1972.

Tyler, D.B. The fatigue of prolonged wakefulness. Federation Proceedings, 1947, 6, 218.

United States Department of the Army. AGO. PRB. The development of experimental stress-sensitive tests for predicting performance in military tasks: Personnel Research Branch Technical Research Report, 1963, No 1079, 67.

VanCott, H.P., and Kinkade, R.G. Human Engineering Guide to Equipment Design (Rev. Ed.). New York: McGraw-Hill Co., 1972.

Van Ness, F.L., and Bouman, M.A. Spatial modulation transfer in the human eye. Journal of the Optical Society of America, 1967, 57, 401-406.

Vaughan, H.G. The relationship of brain activity to scalp recordings of event-related potentials. In E. Donchin and D.B. Lindsley (Eds.), Average Evoked Potentials: Methods, Results, and Evaluations. Washington, DC: NASA SP-191, 1969, 45-94.

Verschoor, A.M., and van Wieringen, P.C. Vigilance performance and skin conductance. Acta Psychologica, 1970, 33, 394-401.

Wald, G. The spectral sensitivity of the human eye. 1. A spectral adaptometer. Journal of the Optical Society of America, 1945, 35, 187-196.

Walraven, P.L. Color Vision. Annual Review of Psychology. Annual Reviews, Inc., Palo Alto, California, 1972, 23, 347-374.

- Walter, D.O. Spectral analysis for electroencephalograms: Mathematical determination of physiological relationships from records of limited duration. Experimental Neurology, 1963, 8, 155-181.
- Walter, W.G., Cooper, R., Aldridge, V.J., McCallum, W.C., and Winter, A.L. Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. Nature, 1964, 203, 380-384.
- Ward, B. Electroretinogram during 100 percent $+G_z$ gradient acceleration in dogs. Aerospace Medicine, 1968, 39, 468-474.
- Waszak, M., and Obrist, W.D. Relationship of slow potential changes to response speed and motivation in man. Electroencephalography and Clinical Neurophysiology, 1969, 27, 113-120.
- Watson, J.B. Psychology from the Standpoint of a Behaviorist. Philadelphia: Lippincott, 1919.
- Webar, A., Jermini, C., and Grandjean, E.P. Relationship between objective and subjective assessment of experimentally induced fatigue. Ergonomics, 1975, 18, 151-156.
- Wechsler, D. The measurement of emotional reaction. Archives of Psychology of New York, No. 76, 1925.
- Weidemann, W. A new and simple adaptometer. Klin. Mbl. Augenheilk, 1952, 121, 563-573.
- Wendt, G.R. An analytic study of the conditioned knee-jerk. Archives of Psychology, 1930.
- Westbrook, C.B., Anderson, R.O., and Pietrzak, P.E. Handling qualities and pilot workload. Wright-Patterson AFB, Ohio: AF Flight Dynamics Laboratory, AFFDL-FDCC-TM-66-5, September, 1966.
- Westheimer, G. Fourier analysis of vision. Investigative Ophthalmology, 1973, 12, 86-87.
- White, C.T., and Eason, R.G. Evoked cortical potentials in relation to certain aspects of visual perception. Psychological Monographs, 1966, 80, Whole No. 632.
- White, C.T., Kataoka, R. W., and Martin, J.I. Color evoked potentials: development of a methodology for the analysis of the processes involved in color vision. In J.E. Desmedt (Ed.), Visual evoked potentials in man. Oxford: Clarendon, 1977.
- White, W.J., and Jorve, W.R. The effects of gravitational stress upon visual acuity. Wright-Patterson Air Force Base, Ohio, WADC Technical Report 66-247.
- Wickens, C.D., Isreal, J. and Donchin, E. The event related cortical potential as an index of task workload. Proceedings of the 21st Annual Meeting of the Human Factors Society, San Francisco, California, 1977.
- Wickens, C., Isreal, J., McCarthy, G., Gopher, D., and Donchin, E. The use of event-related potentials in the enhancement of system performance. Proceedings of the 12th Annual NASA-University Conference on Manual Control, University of Illinois, May, 1976, 124-134. (NASA TM X-73, 70).
- Wieringen, P.C.W. Van. Arousal en motivatie by vigilantietaken. [Arousal and motivation in vigilance.] Doctoral dissertation, University of Amsterdam, 1975.
- Wilkinson, R.T., and Ashby, S.M. Selective attention, contingent negative variation and the evoked potential. Biological Psychology, 1974, 1, 167-179.
- Wilson, G., and O'Donnell, R.D. Steady state evoked responses: Interaction with cognitive task load. Administrative report, USAF-ASEE Summer Faculty Research Program, Wright-Patterson AFB. Available from Ohio State University Research Foundation, Columbus, Ohio 43212, November, 1978.
- Wisner, A. Electrophysiological measures for tasks of low energy expenditure. In W. T. Singleton, J.G. Fox, and D. Whitfield (Eds.), Measurement of Man at Work. London: Taylor and Francis, 1973, 61-73.
- Woodworth, R.S., and Schlosberg, H. Experimental Psychology. New York: Holt, Rinehart, and Winston, Inc., 1954.
- Wooton, B.R. Photopic and scotopic contributions to the visually evoked cortical potential. Vision Research, 1972, 12, 1647-1660.
- Wulff, B. Clinical electroretinodynamography. Doc. Ophthalmology, 1964, 18, 419-430.
- Zeaman, D., and Wegner, H. Cardiac reflex to tones of threshold intensity. Journal of Speech and Hearing Disorders, 1956, 21, 71-75.
- Zegers, R.T. Photo-sensitization in relation to mean and standard deviation values. Psychological Monographs, 1959, 73, 1-25.
- Zubek, J.P. Behavioral and EEG changes after 14 days of perceptual deprivation. Psychonomic Science, 1964, 1, 57-58.
- Zubek, J.P. Sensory Deprivation: Fifteen Years of Research. New York: Appleton-Century-Crofts, 1969.
- Zubek, J.P., Welch, G., and Saunders, M. EEG changes during and after 14 days of perceptual deprivation. Science, 1963, 139, 490-492.

REPORT DOCUMENTATION PAGE									
1. Recipient's Reference	2. Originator's Reference AGARD-AG-244	3. Further Reference ISBN 92-835-1325-8	4. Security Classification of Document UNCLASSIFIED						
5. Originator	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 rue Ancelle, 92200 Neuilly sur Seine, France								
6. Title	CONTRIBUTIONS OF PSYCHOPHYSIOLOGICAL TECHNIQUES TO AIRCRAFT DESIGN AND OTHER OPERATIONAL PROBLEMS								
7. Presented at									
8. Author(s)/Editor(s) R.D.O'Donnell	9. Date July 1979								
10. Author's/Editor's Address 6570 Aerospace Medical Research Laboratory Wright-Patterson Air Force Base Ohio 45433	11. Pages 88								
12. Distribution Statement	This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.								
13. Keywords/Descriptors	<table border="0"> <tr> <td>Aircraft</td> <td>Physiological psychology</td> </tr> <tr> <td>Design criteria</td> <td>Stress (psychology)</td> </tr> <tr> <td>Human factors engineering</td> <td>Safety devices</td> </tr> </table>			Aircraft	Physiological psychology	Design criteria	Stress (psychology)	Human factors engineering	Safety devices
Aircraft	Physiological psychology								
Design criteria	Stress (psychology)								
Human factors engineering	Safety devices								
14. Abstract	Physiological Psychology has always been the scientific approach most likely to provide in aircraft design the alleviation of environmental stress, the design of protective equipment, the planning of optimal cycles at work and rest, as well as the solution of many special problems relating to seats, lighting systems, safety devices, etc. With the dramatic conceptual changes intervening in aerospace systems, all performance measurements should indicate the manner in which the human is responding, in order to allow the interpretation of data obtained in laboratory conditions. The AGARDograph is a general survey of existing capabilities of, as well as the possible contributions to be made by, psychophysiological measurement techniques in applied settings. This AGARDograph was prepared at the request of the Aerospace Medical Panel of AGARD/ NATO.								

AGARDograph No.244 Advisory Group for Aerospace Research and Development, NATO CONTRIBUTIONS OF PSYCHOPHYSIOLOGICAL TECHNIQUES TO AIRCRAFT DESIGN AND OTHER OPERATIONAL PROBLEMS R.D.O'Donnell Published July 1979 88 pages Physiological Psychology has always been the scientific approach most likely to provide in aircraft design the alleviation of environmental stress, the design of protective equipment, the planning of optimal cycles at work and rest, as well as the solution of many special problems relating to seats, lighting systems, safety P.T.O	AGARD-AG-244 Aircraft Design criteria Human factors engineering Physiological psychology Stress (psychology) Safety devices	AGARDograph No.244 Advisory Group for Aerospace Research and Development, NATO CONTRIBUTIONS OF PSYCHOPHYSIOLOGICAL TECHNIQUES TO AIRCRAFT DESIGN AND OTHER OPERATIONAL PROBLEMS R.D.O'Donnell Published July 1979 88 pages Physiological Psychology has always been the scientific approach most likely to provide in aircraft design the alleviation of environmental stress, the design of protective equipment, the planning of optimal cycles at work and rest, as well as the solution of many special problems relating to seats, lighting systems, safety P.T.O	AGARD-AG-244 Aircraft Design criteria Human factors engineering Physiological psychology Stress (psychology) Safety devices
AGARDograph No.244 Advisory Group for Aerospace Research and Development, NATO CONTRIBUTIONS OF PSYCHOPHYSIOLOGICAL TECHNIQUES TO AIRCRAFT DESIGN AND OTHER OPERATIONAL PROBLEMS R.D.O'Donnell Published July 1979 88 pages Physiological Psychology has always been the scientific approach most likely to provide in aircraft design the alleviation of environmental stress, the design of protective equipment, the planning of optimal cycles at work and rest, as well as the solution of many special problems relating to seats, lighting systems, safety P.T.O	AGARD-AG-244 Aircraft Design criteria Human factors engineering Physiological psychology Stress (psychology) Safety devices	AGARDograph No.244 Advisory Group for Aerospace Research and Development, NATO CONTRIBUTIONS OF PSYCHOPHYSIOLOGICAL TECHNIQUES TO AIRCRAFT DESIGN AND OTHER OPERATIONAL PROBLEMS R.D.O'Donnell Published July 1979 88 pages Physiological Psychology has always been the scientific approach most likely to provide in aircraft design the alleviation of environmental stress, the design of protective equipment, the planning of optimal cycles at work and rest, as well as the solution of many special problems relating to seats, lighting systems, safety P.T.O	AGARD-AG-244 Aircraft Design criteria Human factors engineering Physiological psychology Stress (psychology) Safety devices

devices, etc. With the dramatic conceptual changes intervening in aerospace systems, all performance measurements should indicate the manner in which the human is responding, in order to allow the interpretation of data obtained in laboratory conditions. The AGARDograph is a general survey of existing capabilities of, as well as the possible contributions to be made by, psychophysiological measurement techniques in applied settings. This AGARDograph was prepared at the request of the Aerospace Medical Panel AGARD/NATO. ISBN 92-835-1325-8	devices, etc. With the dramatic conceptual changes intervening in aerospace systems, all performance measurements should indicate the manner in which the human is responding, in order to allow the interpretation of data obtained in laboratory conditions. The AGARDograph is a general survey of existing capabilities of, as well as the possible contributions to be made by, psychophysiological measurement techniques in applied settings. This AGARDograph was prepared at the request of the Aerospace Medical Panel AGARD/NATO. ISBN 92-835-1325-8
devices, etc. With the dramatic conceptual changes intervening in aerospace systems, all performance measurements should indicate the manner in which the human is responding, in order to allow the interpretation of data obtained in laboratory conditions. The AGARDograph is a general survey of existing capabilities of, as well as the possible contributions to be made by, psychophysiological measurement techniques in applied settings. This AGARDograph was prepared at the request of the Aerospace Medical Panel AGARD/NATO. ISBN 92-835-1325-8	devices, etc. With the dramatic conceptual changes intervening in aerospace systems, all performance measurements should indicate the manner in which the human is responding, in order to allow the interpretation of data obtained in laboratory conditions. The AGARDograph is a general survey of existing capabilities of, as well as the possible contributions to be made by, psychophysiological measurement techniques in applied settings. This AGARDograph was prepared at the request of the Aerospace Medical Panel AGARD/NATO. ISBN 92-835-1325-8

B258
4

AGARD

NATO  OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE
FRANCE

Telephone 745.08.10 · Telex 610176

DISTRIBUTION OF UNCLASSIFIED
AGARD PUBLICATIONS

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres, but if not may be purchased in Microfiche or Photocopy form from the Purchase Agencies listed below.

NATIONAL DISTRIBUTION CENTRES

BELGIUM

Coordonnateur AGARD – VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA

Defence Scientific Information Service
Department of National Defence
Ottawa, Ontario K1A 0Z2

DENMARK

Danish Defence Research Board
Østerbrogades Kaserne
Copenhagen Ø

FRANCE

O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92 Châtillon sous Bagneux

GERMANY

Zentralstelle für Luft- und Raumfahrt-
dokumentation und -information
c/o Fachinformationszentrum Energie,
Physik, Mathematik GmbH
Kernforschungszentrum
7514 Eggenstein-Leopoldshafen 2

GREECE

Hellenic Air Force General Staff
Research and Development Directorate
Holargos, Athens, Greece

ICELAND

Director of Aviation
c/o Flugrad
Reykjavik

ITALY

Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3, Piazzale Adenauer
Roma/EUR

LUXEMBOURG

See Belgium

NETHERLANDS

Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 126
Delft

NORWAY

Norwegian Defence Research Establishment
Main Library
P.O. Box 25
N-2007 Kjeller

PORTUGAL

Direcção do Serviço de Material
da Força Aérea
Rua da Escola Politécnica 42
Lisboa
Attn: AGARD National Delegate

TURKEY

Department of Research and Development (ARGE)
Ministry of National Defence, Ankara

UNITED KINGDOM

Defence Research Information Centre
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RE

UNITED STATES

National Aeronautics and Space Administration (NASA)
Langley Field, Virginia 23365
Attn: Report Distribution and Storage Unit

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW.

PURCHASE AGENCIES

Microfiche or Photocopy

National Technical
Information Service (NTIS)
5285 Port Royal Road
Springfield
Virginia 22161, USA

Microfiche

Space Documentation Service
European Space Agency
10, rue Mario Nikis
75015 Paris, France

Microfiche

Technology Reports
Centre (DTI)
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RF
England

Requests for microfiche or photocopies of AGARD documents should include the AGARD serial number, title, author or editor, and publication date. Requests to NTIS should include the NASA accession report number. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)

published by NASA Scientific and Technical
Information Facility
Post Office Box 8757
Baltimore/Washington International Airport
Maryland 21240, USA

Government Reports Announcements (GRA)

published by the National Technical
Information Services, Springfield
Virginia 22161, USA



Printed by Technical Editing and Reproduction Ltd
Harford House, 7-9 Charlotte St, London W1P 1HD

ISBN 92-835-1325-8

CONTRIBUTIONS OF DEVELOPMENTAL TECHNICAL INFORMATION TO