Aviation Research Lab
Institute of Aviation

University of Illinois
at Urbana-Champaign
1 Airport Road
Savoy, Illinois 61874

EFFECTS OF FRAME OF REFERENCE AND VIEWING CONDITION ON ATTENTIONAL ISSUES WITH HELMET MOUNTED DISPLAYS

Michelle Yeh, Christopher D. Wickens,
and F. Jacob Seagull

Technical Report
ARL-98-1/ARMY-FED-LAB-98-1

January 1998

Prepared for

U.S. Army Research Laboratory
Interactive Displays Federated Laboratory
Aberdeen Proving Ground, MD

Contract DAAL 01-96-2-0003

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **JAN 1998**	2. REPORT TYPE	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Effects of Frame of Reference and Viewing Condition on Attentional Issues with Helmet Mounted Displays**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army Research Laboratory,2800 Powder Mill Road,Adelphi,MD,20783-1197**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**75**	

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

Presentation of information using an helmet mounted display (HMD) allows users to view the world through a visor or eyepiece, on which additional data relevant to the task performed, can be superimposed onto the forward field of view. In the research presented here, the issues of **frame of reference** and **viewing condition** (i.e., one eye versus two) are examined in order to determine their effects on tasks of focused and divided **attention**.

Superimposing information from the near domain onto the far domain eliminates scan time and prevents eye accommodation when switching between the near and far domains, but these benefits may be offset by the cost of increasing the amount of clutter in the forward field of view. For HMDs, the issue of frame of reference involves a comparison of world-referenced displays, in which information is displayed so that it is slaved to the momentary orientation of the head, with screen-referenced, such that the location of objects on the display is based on a pre-determined set of x- and y- coordinates, independent of head movement. Additionally, HMDs can be configured so that information is displayed to monoscopically or stereoscopically to one eye or two.

In the current experiment, sixteen subjects (8 civilian, 8 military) viewed static two-dimensional renderings of three-dimensional images depicting hilly terrain, in which targets consisting of tanks, soldiers, land mines, and nuclear devices were hidden. The experiment was conducted in an immersed virtual reality environment known as the CAVE using head-tracked shutter glasses. Subjects were asked to detect, identify, and give location information for targets in the far domain while performing a monitoring task in the near domain. The CAVE presented a field of view of 270° surrounding the subject. Symbology consisting of a cueing arrow, lock-on reticle, and heading tape were presented to one eye or monoscopically to both eyes. Cueing symbology, presented for half the targets, consisted of an arrow pointing in the direction of the target object, which turned into a lock-on reticle once the target was present in the forward field of view. In the world-referenced configuration, the cueing arrow was presented in the periphery of a 40° field of view and the lock-on reticle was superimposed over the target. In the screen-referenced display, the cueing arrow was always in the center of the field of view; the reticle appeared in the same position as the cueing arrow. A nuclear device was sometimes present in the environment as an “unexpected” target. Subjects were instructed that reporting the device took precedence over standard target detection.

The results showed that subjects were more likely to detect targets if they were expected, regardless of priority. The presence of cueing aided the target detection task for expected targets but by drawing attention to the expected target, the cue also drew attention away from the detection of unexpected targets in the environment. This cognitive tunneling effect was somewhat mediated by frame of reference such that unexpected targets were detected more often when subjects searched with the world-referenced display than the screen-referenced display.

Cued targets were detected faster when the symbology was presented to two eyes monoscopically rather than only one, possibly due to the fact that the symbology was more intense when viewed with two eyes. Thus, subjects detected the presence of a cueing arrow – and the target – faster in the two eyed condition than the one eyed condition. There were no significant effects of display and task variables in performance on either the monitoring task or the global positioning task, but there were differences attributable to subject population for the latter task. Military subjects were more accurate and more confident in their recollection of target position within the environment than the civilians. The results are examined within the framework of an information processing model.

1. Introduction

The need to present complex information has led to advancements in technology which attempt to display data more effectively while making the interface more invisible. At the same time, the desire for constant information has caused designers to search for solutions for users who perform tasks in a rapidly changing environment, such as the pilot searching for enemy fighters or the land soldier searching the battlefield for enemy positions (National Research Council, 1997). One option currently being examined is the use of a helmet-mounted display (HMD), which allows users to view the world through a visor or eyepiece that superimposes in the users' forward field of view additional information relevant to the task performed.

This technology is a child of virtual reality, "displays which have been enhanced by special processing … to convince users they are immersed in a synthetic space" (Ellis, 1994, p. 17). The difference however is that the user wearing an HMD is performing tasks in the *real world* with a head mounted "guide" as opposed to being immersed and interacting within a *simulated* environment. HMD technology is similar in many respects to that of head-up displays (HUD) which have been used in aircraft and automobiles to present imagery within the operator's field of view. In fact, the motivations for the development of the HMD are almost identical to those predicted for the HUD twenty years ago (Bailey, 1994).

1.1 Motivations for HMDs

The challenge in the design of helmet-mounted displays is support of operator performance while taking into account operators' talents as well as the constraints due to their presence within unnatural human environments. HMD systems attempt to aid user interaction in the following ways:

1. *Reduce the amount of time spent head down by placing the necessary information in the user's forward field of view*. Valuable time can be saved if one's eyes are aimed continuously at the outside environment. Creating an artificial world in the near domain, that which is presented on the HMD, and superimposing it onto the actual world would reduce scanning time between monitoring information on displays, usually presented at a head-down location, and searching for hazards in the outside world.

2. *Prevent eye accommodation when switching focus between the near domain (symbology) and far domain (the world)*. The benefit is time savings for tasks requiring the user to access information presented at different distances. In aviation, the pilot must divide attention between the instrument panel in the cockpit and the horizon line in the outside world; when driving, the driver must monitor speed on the dashboard of the car but must look outward at the far domain to scan for hazards on the roadway. To bring objects at different distances into focus on the retina, the muscles in the eye adjust the shape of the lens. This process may take up to four seconds (Larry and Elworth,1972).

3. *Present conformal information so that objects in the outside world have a corresponding position on the display*. Presentation of a virtual copy of objects in the far domain on the HMD, i.e., conformally, allows the user to better integrate information between two different worlds without the need for additional eye or head movement. Information in the near domain is linked to objects in the far domain, enabling the user to collect and combine natural cues present in the environment in order to complete his task (McCann and Foyle, 1995). In aviation, conformal symbology has often been used to present a horizon line or virtual runway symbol to the pilot using HUD technology. The imagery serves as a guide to the pilot and allows him to better scan between information in the near domain and the actual objects in the world (Wickens and Long, 1995).

4. *Allow for more freedom in movement than is possible with the HUD, which is rigidly attached to some surface, whether it be the windshield of a cockpit or automobile*. The advantage of the HMD over the HUD is its flexibility of rotation, which allows the user to receive data updates as he moves around in the world. Since the

> HMD can be directly attached to the user, it allows the user to receive more information by simply turning the head as opposed to changing vehicle heading as with a HUD. Osgood and Wells (1991) showed that the use of a head-tracked HMD had a performance advantage over the HUD for target detection since pilots could rotate their head to search the terrain for cues. Once the target was located, head movements could be made to continuously track the target.

Ironically, some research has suggested that HMDs and HUDs may actually hinder performance of the very tasks they are intended to improve: by combining fields of information to reduce visual scanning time between two visually distant domains, more information will need to be presented to the user at once. Since each additional piece of information adds to the clutter and confusion on the display, superimposing symbology on the forward field of view may increase the difficulty in finding information necessary for the task at hand (Teichner and Mocharnak, 1979). These tasks would consist of focusing attention on either the near domain (i.e. symbology) or the far domain (i.e. real world) or dividing attention between the two domains (e.g. looking for information in the symbology but monitoring for targets in the real world). To perform the first two tasks, the user would need to be able to attend to information in one domain without interference from the second domain. In the divided attention task, users need to either integrate information from each domain or be able to monitor events in both domains simultaneously.

A second problem to HMD use is the limited field of view; the field of view of a person with normal vision is approximately 200°, with 120° of binocular overlap, but so far, no HMDs are available which allows this range. In fact, with some HMDs, the field of view is as small as 52° with only 18° of binocular overlap (Klymento and Rash, 1995), and consequently, the limited field of view for HMDs has been associated with poorer performance in target detection. Hettinger, Nelson, and Haas (1994) examined this issue in a comparison of an opaque HMD and a dome display. The field of view for the HMD was 60°, created by overlapping two monocular views by 50% to form a 20° binocular center. A projection room was used for the dome display condition. Subjects were tested in pairs; one in the HMD condition, the other in the dome condition. The target detection task required each subject to locate the other person's aircraft (depicted with gray circles), considered the enemy, and respond before being detected by their partner. Greater detection accuracy and lower detection times were found with the dome display. These problems with HMDs will be examined in greater detail in order to determine whether the advantages of an HMD outweigh the costs.

1.2 Configuration and Uses of HMD Technology

The technology driving the HMD consists of an image source, possibly a cathode ray tube (CRT) mounted on the helmet which converts information into a video display. The optics used to present the image consist of mirrors which magnify and transmit the image towards a collimating lens, projecting the object to virtual or optical infinity at the distance of the real world (e.g. the horizon). Another mirror serves as a combiner to superimpose the collimated image onto the outside environment (Barfield, Rosenberg, and Lotens, 1995; Bellenkes, 1988). Additionally, HMDs can take advantage of special hardware which uses sensors to track head or body movements and hence display information in what the user perceives to be world-referenced coordinates. This provides the simulation with a sense of reality by updating the display correspondingly (Ellis, 1994).

HMDs can be configured so that the outside environment is visible, i.e. the display is transparent, or hidden, i.e., the display is opaque. The use of a see-through HMD presents an image similar to that obtained using a HUD; for example, aircraft flight path information can be presented on the display and superimposed on the outside world. In this case, imagery can be conformal, non-conformal, or some combination of the two. Already, see-through HMDs have stimulated excitement in the field of medical imaging by allowing doctors to view a virtual image of the patient's internal organs superimposed on the patient's body (Rolland, et al., 1995). On the other hand, interaction using an opaque display on an HMD is like that of a computer display. The user is unable to see objects in the outside world through the display, much as a monitor obscures objects placed directly behind it.

Whether see through or opaque, the presentation of information can be displayed to one eye or two. A *monocular* presentation displays the image to only one eye with the other eye having an unaided, dark-adapted view of the surrounding environment. *Biocular* viewing presents the same image to both eyes so that the resulting view is a two-dimensional planar or three-dimensional perspective view. Finally, the *binocular* format displays a slightly offset view to each eye allowing the user to perceive the image with stereopsis depth cues.

Applications for HMD technology range from the medical domain for the viewing of ultrasound data (Bajura, Fuchs, and Ohbuchi, 1992) to aviation displays for the presentation of tactical information (Geiselman and Osgood, 1995). For military tasks, navigational information for orienteering can be viewed using HMD maps; tactical information regarding friend and foe positions could be presented to a platoon leader; and checklists of necessary equipment or a plan of attack could be viewable using an HMD (Sampson and Warren, 1992; National Research Council, 1997).

1.3 Display Taxonomy

In all of the activities mentioned, the benefit of the HMD is that it gives the user access to all types of imagery without re-focusing the eyes. Information can be presented so that it is oriented to head position or not. Data can be presented using FLIR night-vision goggles (NVGs) or computer generated imagery (CGI) using monocular, biocular, or binocular configurations. Thus, the various display configurations lead to the creation of a taxonomy in which tasks completed with HMDs can be categorized. The display taxonomy is presented in Figure 1.3.1.

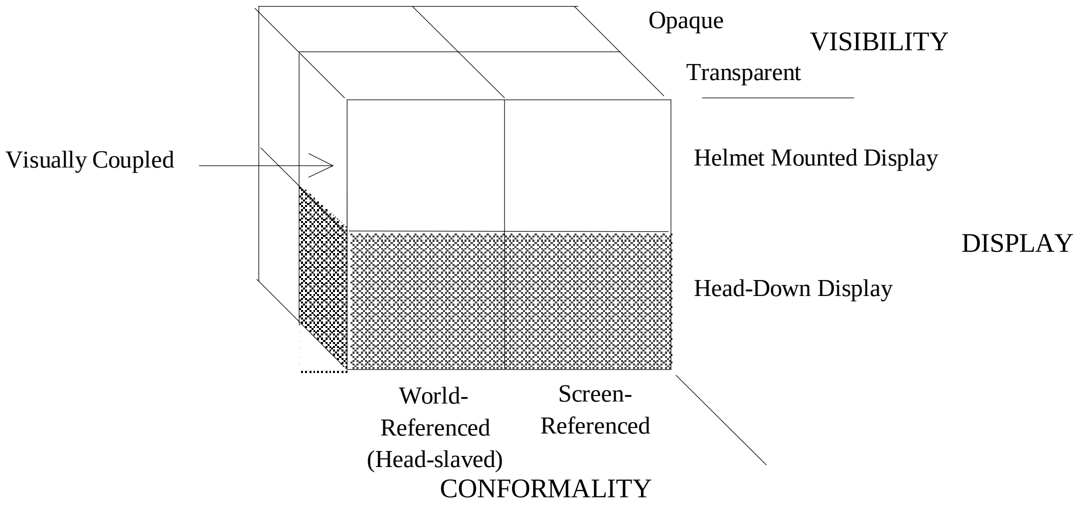


Figure 1.3.1. Display Taxonomy

The taxonomy shows the need to evaluate tasks based on *display* – helmet mounted or head down. This examination would consist of comparisons of performance using a head-down display (HDD), e.g. a vehicle dashboard, airplane instrument panel, or hand held display, versus using an HMD in which information that would normally have been presented head down is superimposed on the forward field of view. The second dimension defines the *visibility* of the world, the difference between opaque and transparent displays described above. The comparison leads into the dimension of *conformality*, or how information should be presented as the head moves. Three choices are available: (1) world-referenced (or head-slaved), such that the user is immersed in the environment, and the data displayed is dependent or slaved to the current orientation of the head; (2) screen-referenced, in which the displayed location of the information is based on a pre-determined set of x- and y- coordinates independent of head movement; or (3) eye-slaved (not included in Figure 1.3.1), such that information is based on the orientation of the eye – this display will not be discussed in the current literature review (National Research Council, 1995; Wickens, 1996). Figures 1.3.2 and 1.3.3 present examples of a world-referenced display and a screen-referenced display, respectively.

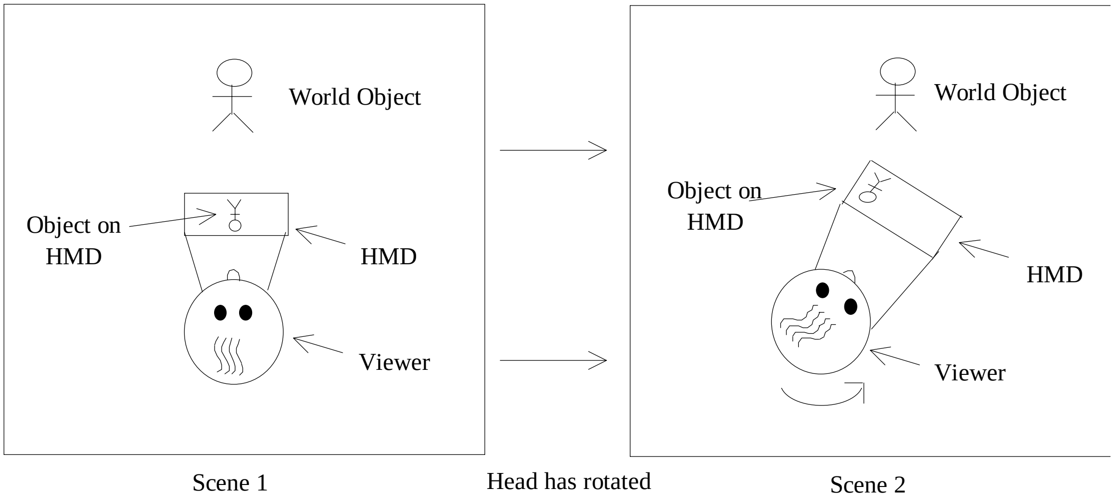


Figure 1.3.2. World-referenced imagery.

As shown in Scene 1 of Figure 1.3.2, the image on the HMD is presented at the same location as the actual object in the world would be seen by the viewer without the HMD – in the center of the field of view. As the viewer turns his head (Scene 2), the location of the image on the HMD is updated to correspond to the location of the world object with respect to the viewer's current head orientation. The use of a world-referenced display on an HMD is similar to the use of conformal symbology in the HUD domain. In HMD systems, what the user sees in a visually coupled system is dependent on head orientation measured using a head-tracking system, allowing the HMD user to be immersed in the virtual world (Fisher, 1982). For example, the user might see a pointer on the display designating the location of the target in the far domain. Other terms to describe this system are head-coupled, visually coupled, or viewpoint-dependent imaging.

Virtual environments generally presented on opaque displays traditionally require world-referenced imagery, so that the image on the display remains fixed in world coordinates as the head moves, thus giving "…users the illusion of displacement to another location" (Ellis, 1994, p. 17). The world-referenced displays afford active exploration within a virtual environment, which has been shown to aid spatial knowledge acquisition similar to that achieved when interacting with the real world (Arthur, Hancock, and Telke, 1996). The various uses of head-coupling in VR applications range from games (Kaplan and Brown, 1997) to simulating the laws of physics (Dede, Salzman, and Loftin, 1996) to knowledge based maintenance assistance (Feiner, MacIntyre, and Seligmann, 1993). Hodges et al. (1996) are creating a virtual environment simulation of an airplane to help people overcome the fear of flying. Dede, Salzman, and Loftin (1996) are using color, stereoscopic, head-tracked HMDs in conjunction with auditory and haptic input devices to teach students about science by allowing them to become a part of science. For example, to learn more about the interaction of mass, velocity, and energy, users in the virtual world can throw and catch balls of different masses through a virtual corridor; they can even "become" one of the balls bouncing through the corridor.

In contrast to world-referenced imagery, Figure 1.3.3 depicts an example of screen-referenced displays.

As shown in Figure 1.3.3, the location of the image is pre-determined, and the image itself may have no real world referent, i.e. its location on the screen has no one-to-one correspondence with the location of objects in the real world. Thus, the screen-referenced configuration presents information in a non-conformal manner. This type of display may be used in HUDs for presenting heading information; the display updates to reflect the current heading of the aircraft but the location of the display itself does not change as the aircraft's heading changes.

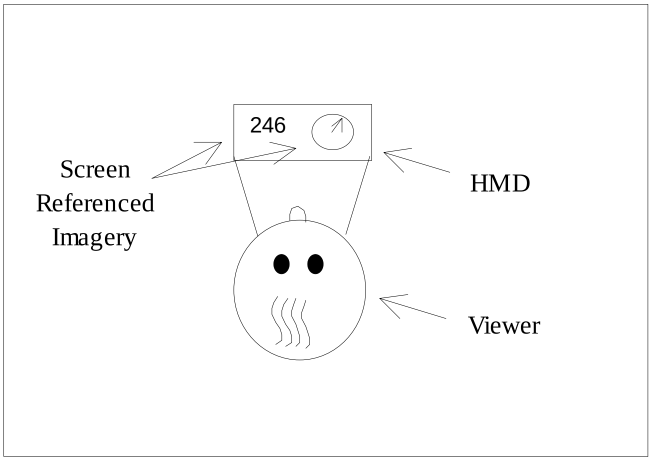


Figure 1.3.3. Screen-referenced Configuration

The issue of conformality is linked to the second dimension of the taxonomy, that of *visibility*. World-referenced displays can be presented on see through or non-see through HMDs. If the HMD is transparent, the use of world-referenced symbology will result in the presentation of information conformally, such that objects on the display need to have a reference in the outside world. The advantage to this type of presentation is information integration between *near and far domains*. On the other hand, the use of an opaque HMD immerses subjects in a virtual world; interaction in this case is similar to that of virtual reality applications such that users act and react within the simulated world (Dede, et al., 1996; Pausch, Proffitt, and Williams, 1997; Pausch, Shackelford, and Proffitt, 1994). Although the HMD would be lighter due to the less complex optics in an opaque display, the inability to see beyond the display would totally isolate one from the surrounding world (National Research Council, 1997). Note also that an opaque monocular view can be used in a manner that will be discussed later in this chapter.

The research described in this literature review will be examined in terms of the framework presented in Figure 1.3.1. The issue of **frame of reference** (FOR) is linked closely to that of *conformality*, thus studies will be described concerning the use of world-referenced displays as well as conformal vs. non-conformal imagery. This leads into the issue of **attention**; that is, whether the HMD user will be able to effectively divide his attention between information on the display and information in the world so that attention can be focused specifically on one domain for one task but divided efficiently between both domains if needed. When using an HMD, does the time savings in reduced visual scan between the near and far domains outweigh the cost of additional clutter in the forward field of view? This issue is especially important for tasks of land navigation in which the HMD user must not only monitor information on the display and objects in the world directly in front of him but also obstacles in the world surrounding him. Finally, *display* comparisons will examine the presentation of HMD information to **one eye versus two**; this issue concerns the costs and benefits of monoscopic and stereoscopic viewing and comparisons of monocular, biocular, and binocular display configurations.

The current literature review will use the taxonomy presented above in order to evaluate the issues previously mentioned in tasks of HMD interaction, i.e. frame of reference, attention, and one eye versus two eyed presentation of information. Although many of the studies reported have been conducted using HUD, rather than HMD, technology, given the close similarity of

HMDs to HUDs, we believe that the results will be generalizable to the HMD domain. The literature review will be followed by an overview of the experiment to be performed – an examination of these issues in the context of land navigation.

2. Frame of Reference (FOR)

Previous research with head-down displays indicates that the optimal FOR is task dependent; that is, the nature of the task serves to define which FOR will support it best. Visual search has been shown to benefit from the exocentric FOR, allowing the user to examine the whole environment at once without the need to change orientation (McCormick and Wickens, 1995; McGovern, 1991; Miller, 1988). On the other hand, travel benefits are achieved with the egocentric FOR which provides a natural compatibility between perception and control such that the display viewpoint is identical to the axis of control. Wickens (1997b) reports data from numerous experiments which showed that travel toward the viewpoint with an exocentric perspective increased travel time by about 10% as compared to the egocentric FOR. Finally, tasks involving the understanding of objects' locations within the environmental space benefit from the use of the exocentric viewpoint (Arthur, Hancock, and Chrysler, 1993; Barfield, Rosenberg, and Furness, 1995; McCormick and Wickens, 1995; Wickens and Prevett, 1995). This result is not surprising since viewing the immersed display requires mentally piecing together various "snapshots" of the environment taken from different perspectives of virtual space to form a "big picture" whereas the exocentric display presents one global view from one "permanent" angle. Ellis, Tharp, Grunwald, and Smith (1991) hypothesized that the use of exocentric displays could prevent misjudgment of viewing location or direction since the viewpoint would be from a "bird's eye" view of the world.

For HMDs, the issue of FOR involves a comparison of world-referenced to screen-referenced display imagery. These two displays have complements in the HUD domain – those of conformal and non-conformal displays, respectively. Since few comparisons of world-referenced to screen-referenced FORs have been made in the HMD domain, the use of conformal versus non-conformal displays in the HUD domain will be discussed. Research examining the addition of world-referencing in various virtual reality technology applications will then be described.

2.1 Conformal vs. Partially Conformal vs. Non-Conformal Symbology

The presentation of information head-up reduces visual scanning between the outside world and the displays but imposes a cost as a result of the additional information on the forward field of view – that is, the amount of clutter. Although decluttering the display has been examined in aviation (Mykityshyn, Kuchar, and Hansman, 1994; Hofer, Palen, and Possolo, 1993), its use must be considered within the context of the task; Dudfield, Hardiman, and Selcon (1995) warn of using decluttering in high workload situations in which the user may not have the time to add or remove symbology from the display. One way to reduce clutter in the forward field of view is to present the image conformally or in world-referenced coordinates. Since the objects on the display have a one-to-one relationship with objects in the outside world, the amount of clutter presented head-up is reduced. The exception would be when the alignment of the display to the world is not adequate. One type of conformal symbology is virtual conformality, which overlays information on a physical position in space rather than on the far domain objects themselves. For example, a virtual conformal display of a desired flight path might show a tunnel in the sky along that flight path, even though such a tunnel is not actually visible in the far domain. Similar to the conformal display, McCann and Foyle (1995) propose replacing conventional HUD symbols with scene-linked objects that appear to be part of the virtual world, e.g. the addition of virtual buildings to the image scene to indicate height. In aircraft displays, conformal imagery is aircraft referenced – as the pilot moves through the airspace, the HUD updates its information to correspond with new information in the forward field of view in the outside world based on the orientation of the plane.

In contrast, non-conformal symbology is symbology whose location in the display is generally predetermined and not influenced by momentary changes in heading or orientation. Its main advantage over a head-down presentation of the same information is the reduction of visual scanning distance between the information and the outside world. However, since the information is placed at x- and y- positions on the screen that are independent of vehicle or head orientation, that is the symbology has no isomorphic relationship with the objects in the world, a cost to this type of display is the generally the addition of clutter. The question that must be addressed is whether the presentation of information head-up and non-conformally in a transparent HMD will aid performance relative to the head-down conditions.

Experiments comparing performance between the conformal and non-conformal HUDs in an aircraft landing task have shown an advantage for conformal displays (Martin-Emerson and Wickens, 1997; May and Wickens, 1995; McCann and Foyle, 1995; Wickens and Long, 1995). Martin-Emerson and Wickens (1997) compared the two display configurations by presenting pictorial information consisting of head-up flight path guidance symbology conformally on half the trials and non-conformally on the other half. Subjects executed the landing in varying levels of visibility. The results showed no changes in tracking error with visibility when subjects used the conformal symbology but as visibility decreased, error increased when subjects used the non-conformal symbology.

Similarly, Wickens and Long (1995) showed performance benefits for conformal displays when subjects needed to access information from both near and far domains. In their experiment, pilots flew along a pre-specified flight path maintaining vertical and lateral tracking before breaking out of the clouds, seeing the runway, and executing a landing task. The two symbology sets were identical to those of Martin-Emerson and Wickens (1997); however, the presentation differed such that the guidance information (a virtual runway or non-conformal ILS) was presented head-up or head-down. The results showed an advantage for conformal symbology especially when presented head up and when the outside world was visible (e.g. overlapping HUD and far domain imagery). In contrast, the use of non-conformal displays resulted in a trade-off for the HUD as the benefits of reduced scanning were offset by the additional clutter of the overlapping imagery.

In both HUD and HMD domains, information must be extracted from cluttered displays. The use of partially conformal displays, in which information presented relates to objects in the far domain but is not superimposed on them, has been examined cursorily. Thus, their benefits are still unclear. May and Wickens (1995) compared the use of virtually conformal, partially conformal, and non-conformal symbology in a flight task performed under varying weather conditions, which were simulated by varying the intensity in the far domain. The difference between conformal and partially conformal symbology was created by condensing the heading scale in the latter condition. Subjects were asked to maintain flight based on heading, altitude, and airspeed information presented in the near domain while monitoring for aircraft in the far domain. The results showed no difference in performance between the conformal and partially conformal displays.

Collectively, the results of these aviation-based HUD studies indicate an advantage for conformal, or scene-linked, symbology and offer encouragement for the use of partially conformal symbology. Findings of the experiments show that although displaying information head-up has a disadvantage in the increase of clutter on the display, the cost can be minimized by the use of conformal symbology (Wickens and Long, 1995). Additionally, the benefits of reduced visual scan generally outweigh the clutter costs imposed by the non-conformal symbology (May and Wickens, 1995). One final advantage for conformality may be the prevention of binocular rivalry when using binocular HMDs, as will be discussed later. Applying these findings to HMDs suggests that the use of world-referenced displays to present spatial information would prove advantageous over screen-referenced displays; that is, a benefit is expected for HMD symbology which is updated as the head turns, identical to the manner in which conformal HUD symbology changes as the aircraft adjusts its pitch or yaw.

2.2 Visually Coupled Systems

The presentation of data with a world-referenced display is based on viewpoint location and perspective calculated by measuring head movements. Initial results show enthusiastic subjective response to this viewpoint especially in improving the sense of presence in three-dimensional scenes (Arthur, Booth, and Ware, 1993; Hendrix and Barfield, 1995; Hendrix, Brandt, and Barfield, 1995; McKenna, 1992). However, current limitations in head-tracking technology result in a time delay between head movement and the updating of the display, which may cause dissociations between visual and vestibular processing, resulting in negative interaction within the virtual environment (Dudfield, Hardiman, and Selcon, 1995). Thus, the benefits of world-referenced imagery would be expected to be contingent upon how rapidly the image is updated when the head turns.

Spatial awareness of three dimensional environments and the perceived realness of the interaction increase with the use of world-referencing. Two studies by Pausch and his colleagues demonstrated the advantage of head-coupled versus hand coupled guidance in aiding the sense of presence in virtual environments. In Pausch, Shackelford, and Proffitt (1994), subjects' ability to locate a target within a virtual room with head-tracked and hand-tracked displays was examined. Subjects were asked to perform a visual search task for targets (a two digit number) in a scene (a synthetic room with numbers on the walls) using a head-tracked biocular opaque HMD or virtual reality environment viewed through a desktop environment. The results showed that subjects using head tracked displays performed the tasks 42% faster than those performing the tasks with the hand-tracked display. An added benefit for using head-tracked displays for training was found, such that subjects performed the search task faster with the hand tracked display if they had initially used the head-tracked configuration. In Pausch, Proffitt, and Williams (1997), the authors attempted to further quantify the concept of immersion, or "being there," by asking subjects to perform a visual search task not only for targets in the scene, as in Pausch, Shackelford, and Proffitt (1994) but also for targets not present in the scene. The latter task was included in order to determine the time that subjects needed to confidently search the scene in a virtual environment versus a desktop environment. The equipment used was the same as that used in the previous experiment. Subjects searched for a target letter in a virtual room containing 170 letters displayed on the walls, ceiling, and floor. Results showed that users of the HMD head-tracked display were faster to indicate when no target was present. That is, the use of a head-tracked display gave subjects a better mental FOR for the virtual room. Subjects navigating with a joystick needed to double check areas in order to confirm the absence of a target.

Other research has also supported the value of head tracking. Hendrix, Brandt, and Barfield (1995) discovered that the use of a head-tracker encouraged subjects to interact with objects more realistically within the simulated environment. In other words, the visually coupled system improved the fidelity of interaction within the virtual world. McKenna (1992) examined the effects of viewpoint control in a task requiring subjects to align two cubes presented on a monoscopic monitor display. The use of a head-tracked control and mouse to change the viewpoint orientation was compared to a viewpoint independent condition. The findings showed better performance when the viewpoint could be changed. Subjective ratings showed a preference for using the head-tracked display over mouse control; subjects' comments indicated that head movement improved the sense of depth within the space.

However, rotating around an environment using a head-tracked display may impose a disadvantage on judgments of spatial position. Henry and Furness (1993) compared performance on spatial tasks within a museum environment using four different experimental conditions. Subjects could tour the actual place, interact with a computer model of the museum displayed on a monitor, view the computer model using an opaque screen-referenced stereoscopic HMD, or view the model with an opaque head-tracked stereoscopic HMD. A spaceball was used for viewpoint rotation and movement control for the monitor and screen-referenced HMD conditions. In the head-tracked condition, the spaceball was used only for forward movement as viewpoint rotation could be accomplished by simply turning the head. After a 15 minute tour, subjects were asked to perform tasks requiring judgments of spatial perception and orientation by estimating the dimension of the various rooms in the museum and indicating the placement of objects they contained during the tour. The results showed poor estimation of dimensions for all three simulation conditions; this difficulty was most prevalent with the head-tracked display. The authors hypothesized that this result may have been due to the fact that subjects not only turned their heads to change the view as they would in real space but also moved their eyes to the edges of the HMD where the distortion of images was greatest, thus resulting in estimation errors. There was no significant difference between the display types for the orientation task. Subjective results evaluating the sense of presence within the space showed similarities in ratings between interaction with the head-tracked display and interaction within the actual environment.

Benefits have been found for the addition of an HMD in the airplane cockpit as it allows the pilot more freedom of head movement for detecting events in the world than is possible with HUDs or other traditional cockpit displays – these advantages are a direct result of the ability to present data continuously in the pilot's forward field of view. Osgood, Geiselman, and Calhoun (1991) and Geiselman and Osgood (1994) both found improved target detection and tracking in air search and attack scenarios when pilots viewed information with a conjunction of HUD and HMD symbology versus HUD alone. Results of the experiments showed that presentation of data with the HMD allowed the pilots to look farther off boresight for longer periods of time than with HUD only presentation.

In the aviation domain, the *conformality* dimension of the HMD taxonomy as shown in Figure 1.3.1 includes three display types: world-referenced, aircraft referenced, and screen-referenced. The world-referenced and screen-referenced configurations were described earlier in the chapter. The aircraft referenced configuration is similar to the world-referenced configuration in that the location of the image on the HMD is updated to correspond to the location of the world object but this is accomplished with respect to aircraft orientation rather than the viewer's current head orientation. The presentation of data would be identical to that displayed on conformal HUDs.

At times, a combination of display types may need to be considered; the literature suggests that the advantages of head-tracked displays may be determined by the task to be completed. Haworth, Sharkey, and Lee (1995) examined the combinations of world-referenced, aircraft referenced, and screen-referenced symbology when using HMDs to present pilots with flight information. The far domain was a green monochrome scene, over which one of five different symbol sets, ranging in display perspective, could be presented:

(1) screen-referenced presentation of flight path information (e.g. altitude, attitude, and heading).
(2) aircraft referenced presentation of flight path information in addition to a compass rose, pitch lines, conformal horizon line, and ownship symbol. A border was drawn around the symbology simulating the outline of a HUD, and the display of information did not exceed the borders regardless of head position.
(3) screen-referenced flight path information, an uncompressed pitch line, and ownship symbol.
(4) same as (2) except that the conformal horizon could be viewed off axis from the center of the aircraft and was therefore no longer restricted within the outline of the HUD.
(5) screen-referenced flight path information with the addition of world-referenced pitch lines and horizon lines.

Helicopter pilots were asked to complete a nap-of-the-earth flight, determine whether the slope of the surface was acceptable for landing, and perform a landing task. The results of all the three tasks showed benefits for a combination of a world-referenced horizon line and screen-referenced presentation of flight path information which allowed data to be displayed continuously in the pilot's forward field of view. In fact, in the nap-of-the-earth flight, subjects made more head motion reversals when the symbology was confined to the HUD. Subjective ratings of awareness for current and upcoming events for the nap-of-the-earth flight task were higher and decisions regarding whether or not to land based on the slope of the terrain were more accurate with symbol sets (4) and (5), i.e. when the artificial horizon and elevation lines were presented conformally and other symbology displayed non-conformally but constantly in the forward field of view. Control on the landing task was improved with the presentation of flight path data using a head-tracked display; subjective ratings for the task also showed that displays in which information remained in the field of view, i.e. screen-referenced [symbol sets (1), (3), and (5)], reduced workload. Thus, the results overall point to an advantage of conformal, world-referenced symbology.

Geiselman and Osgood (1995) examined the orientation of HMD symbology, i.e. world-referenced versus screen-referenced, for presenting both target location and fly-to information on a transparent HMD using a search and attack scenario. The search task required subjects to locate a target (aircraft), and the attack phase consisted of intercepting the target by bringing it within weapon limit parameters and then tracking it until a cue to shoot was presented (usually after 10 seconds). Four conditions were examined:

- The control condition was a HUD-only presentation of altitude information, target locator information when the target was outside the HUD field of view, target designator box when the target was within HUD field of view, and target range and closure rate scale.
- Another condition was a world-referenced presentation of target location information on an HMD in conjunction with the HUD symbology used in the control condition. Target range and closure rate information was duplicated from the HUD and presented on the HMD.
- A third condition displayed aircraft-referenced fly-to information showing the most efficient path to the target on the HMD in addition to the HUD-only symbology; target range and closure rate information was duplicated from the HUD and presented on the HMD.
- A combination of world-referenced target location data and aircraft-referenced fly-to information was presented in a fourth condition in conjunction with the HUD-only symbology.

The results for the second experiment showed that there was no difference in performance time or viewing angle for the search task. The results of the attack phase showed that the amount of time spent off-boresight was greater for all three HMD conditions compared with the HUD only condition; the angle viewed off-boresight was greater with the presentation of fly-to and both fly-to and look-to information than the HUD only condition, but there was no difference between the performance with HMD look-to information and the other three displays. The results led Geiselman and Osgood (1995) to conclude that the benefits of HMD presentation of information was in the reduction of time spent maneuvering and the increased time spent looking off-boresight to intercept the target and bring it within firing distance.

The results of studies by Haworth, et al. (1995) and Geiselman and Osgood (1995) stress that the benefit for world-referenced or screen-referenced presentations on an HMD may be the continuous availability of information in the forward field of view allowing pilots to look further off-boresight for a greater amount of time to search for objects in the outside world. Thus, there are advantages for both screen-referenced and world-referenced display configurations. The benefit to screen-referenced displays over a HDD is the continuous presentation of objects directly in the user’s forward field of view which reduces visual scan between the near and far domains. The world-referenced displays includes the advantage of the screen-referenced display and also creates an increased sense of presence allowing for more realistic interaction within the environment. Given the success of the conformal display in the HUD domain, we predict a similar trend for the world-referenced HMD.

2.3 Summary

The presentation of information from the near domain onto the far domain can be displayed non-conformally or conformally. The advantage of non-conformal symbology is the reduction of scan time between the display in the near domain and the outside world; the advantage of conformal symbology is not only the facilitation of scanning between the near and far domains but also the reduction of clutter in the forward field of view. The advantages of conformal imagery over non-conformal imagery have been noted in the aviation domain (Wickens and Long, 1995; Martin-Emerson and Wickens, 1997, May and Wickens, 1995).

In the HMD environment, conformal and non-conformal displays translate to world-referenced and screen-referenced configurations, respectively. No benefit has been found for one symbology type or the other. Rather experiments conducted in the aviation domain have shown that the main advantage to HMDs is the presentation of symbology continuously in the pilot's forward field of view. The use of a world-referenced display benefits the acquisition of knowledge regarding the location of objects in a virtual world relative to a monitor display (Pausch,et al., 1997) and improves the fidelity of the interaction within a virtual environment relative to non-head tracked displays (Hendrix,et al., 1995). However, spatial judgments made while interacting within a simulated world is often underestimated, possibly due to the reduced field of view, which deprives users of spatial information in the periphery (Henry and Furness, 1993).

3. Attention

As mentioned in the introduction, presenting information directly in the user's forward field of view on an HMD results in time savings, by eliminating scanning between the displayed data and outside world. Such presentation can also prevent the need for the eye to re-accommodate by displaying the information at a point close to virtual infinity.

The benefit attributable to the prevention of visual scan between the near and far domains was supported in an experiment by Andre and Cashion (1993) which examined the effects of display separation. Subjects were asked to perform a tracking task while simultaneously responding to a randomly appearing stimulus (a left or right arrow displayed within a circle). The task required subjects to reduce the "error" between a cursor represented by a cross and the target area. Tasks were completed using a workstation monitor. The results showed performance decrements in the tracking task when the display separation between the two sets of imagery required large head movements relative to a control condition in which the imagery was superimposed. No differences in performance were found for small display separations, i.e. less than 8 degrees.

The second benefit is the prevention of the need for eye accommodation by presenting near domain information, which is usually accessed head-down, in a collimated head up or helmet mounted format, superimposed on the outside world at virtual infinity. In an experiment conducted by Weintraub, Haines, and Randle (1985), subjects were asked to divide attention between flight instruments and the runway, presented so that it appeared far away. Subjects initially viewed the instrument panel and then shifted focus to a simulated runway, hidden at first by fog, to determine whether the runway was open or closed. The results showed that refocusing increased the time to make the decision as a result of accommodation between the near and far domains by up to 70ms per diopter of accommodative change.

Similarly, Fadden and Wickens (1997) concluded that there was a cost to presenting information at close depths to the pilot due to accommodation between the near and far domains. In their experiment, head-up presentation of flightpath information was compared to two head-down conditions: a HDD-near condition displayed flightpath information closer to the pilot than the far domain, and a HDD-far projected condition, which displayed the information at the same distance as the far domain. Subjects were asked to monitor for changes in airspeed in the near domain and for the appearance of a helicopter target in the far domain. The use of the HDD-near condition incurred a cost on the detection of far domain events while the use of the HDD-far condition incurred a cost on detection of near domain events, reflecting presumably these accommodative costs.

In HMDs, superimposing images from the near domain onto the far domain has the benefit of maintaining the high level of detail present in the real world and reduces the amount of computational resources which would otherwise be required to update computer-generated scenes of a complex outside world. However, the disadvantage to overlapping imagery is that the user may form two different mental models of the system: one of the computer imagery – the near domain – and the second of the real image – the far domain (Barfield, et al., 1995). The system designer must be able to integrate the two domains in such a way that the user views them as one, and must evaluate the presentation based on how well it lends itself to completing tasks of both focused attention in one domain – and thus filtering out irrelevant information in the second domain – and divided attention – e.g. tracking data in the near domain while monitoring for hazards in the world, or composing renderings of objects in both domains.

The ability to divide attention between information sources located at different areas on the display, and focus attention on one while ignoring the other can be predicted using two different theoretical conceptions of attention: object-based and space-based. Object-based theories of attention propose that the visual scene is separated into various objects that are processed sequentially. That is, processing occurs in stages, such that information is separated into perceptual groups before attention can be allocated to each (Duncan, 1984; Kramer and Jacobson, 1991; Yantis, 1992). If objects are, in part, defined by domains (near and far), then focusing attention on one domain thus reduces allocation of attention to the second domain. On the other hand, space based theories are based on the idea of a spotlight moving around the information to be processed so that items located close together within the space of the spotlight will be processed (Eriksen and Eriksen, 1974; LaBerge and Brown, 1989). In HMD design, the space based theories of attention would predict that when images are superimposed, information in the far and near domain located close together will be processed together, whether this is desirable for the task at hand (divided attention) or not (focused attention).

3.1 Focused Attention

The resulting visual clutter from two overlapping sources of information could interfere with tasks of focused attention. For example, superimposing information on the forward field of view may obscure events in the outside world. In their experiment described earlier, May and Wickens (1995) found that focusing attention in the near domain was difficult due to the low contrast between the HUD symbology and outside world in a task which required subjects to focus on flightpath information in the near domain and scan for aircraft in the far domain. Monitoring information on the HUD was impaired by the presence of objects in the outside world. It is not surprising though that detecting aircraft incursions in the far domain was faster when flightpath information (near domain) was displayed head-up due to the reduced scanning and elimination of changes in eye accommodation between the two data sources. Similarly, Wickens and Long (1995) found that focused attention on the far domain to detect runway incursions, i.e. an unexpected aircraft on the runway during the execution of a landing task, was impaired by the increased clutter resulting from the head-up presentation of information.

One technique for facilitating the automatic segregation of the search field is intensity coding or *lowlighting*, in which different domains of information are presented at varying levels of brightness. Items important to the task at hand can be presented in a higher intensity than those items which are currently inconsequential. Conversely, information which is not crucial to the current task can be presented in a lower intensity format which does not draw attention away from the task being performed. The use of lowlighting serves to visually segregate information on the display allowing the user to successfully focus attention on information presented at similar intensities as well as divide attention among objects presented at differing intensities (Martens and Wickens, 1995).

Ververs and Wickens (1996) conducted an experiment to determine whether intensity coding, or lowlighting, would help visually segregate information presented head-up. The authors used two levels of clutter in the experiment (high and low) presented under two different levels of intensity using two different display conditions (head up or head down). The task required pilots to fly along a specified path based on heading information (focused attention on the near domain) and detect aircraft events in the world (focused attention on the far domain). The results of the study showed an overall advantage to the HUD presentation of information for detection of both near and far domain events. As the amount of information on the display (i.e. clutter) increased, search time for near domain events increased and impaired detection of aircraft in the far domain. The addition of intensity coding served to visually reduce the amount of clutter in the forward field of view, reducing the search time for the relevant information without slowing down the processing of that information.

3.2 Divided Attention

The effectiveness of superimposing two fields of information was given support by the results of Naish (1964), who examined the overlaying of instrument displays onto the actual world both in real flight and simulation scenarios. Subjects in the flight experiment were required to perform two tasks: take-off, which required information in the outside world, and landing using an instrument approach, which required information on an electronic display presented in the pilot's forward field of view. Findings of successful performance by most of the subjects on the two tasks led Naish to conclude that it made no difference if the information to be focused on, i.e. the domain of interest, was in the near or the far domain. Similar results were obtained by subjects flying in the simulated world; information from both the display and simulation world was attended to successfully. In two separate laboratory experiments, Naish superimposed simple visual fields (numerals in the background for a monitoring task and a rotating helix used for a tracing task in the foreground) and complex fields (monitoring for the presence of a white light in the background while performing a tracking task in the foreground). The results of both experiments showed that subjects treated the superimposed fields of information as one source rather than two. That is, his data appeared to provide strong support for space-based theories of attention in x- and y- space, consistent with the use of a spotlight, which guides attention.

However, Neisser and Becklen (1975) showed that people have problems treating two overlapping sources of information as one. An experiment was conducted in which subjects were shown superimposed videos of people playing games and were asked to follow the action in one or both games. The results showed that subjects could easily focus their attention on only one game but were unable to pay attention to both games without practice. In fact, subjects were so focused on one game, that often they did not notice the figures in the other game performing actions which were unexpected and out of context. Becklen and Cervone (1983) re-investigated this issue and agreed that effectively dividing attention between the two games required extensive practice.

Hypotheses have been proposed as to what causes difficulty in dividing attention between two superimposed fields of information – i.e. *cognitive tunneling* in which one domain captures attention such that events in the second domain are missed or ignored. McCann and Foyle (1995) suggest that this phenomenon may simply be the result of Gestalt principles of grouping, such as color or motion differences between the two domains, which prevent the two domains from being parsed as one. Additionally, they hypothesize that the human visual system may encounter limitations in processing information from two domains. Larish and Wickens (1991), on the other hand, proposed that attentional tunneling may be due to a general narrowing of attention during a high workload task rather than being the result of focusing attention to one domain at the expense of the other. Finally, Foyle, Sanford, and McCann (1991) hypothesize that the difficulty in switching attention between the two domains may be due to a change in the FOR: symbology on the HUD is egocentric, i.e. referenced to the pilot, whereas the information outside the scene is exocentric, i.e. world-referenced. Thus, the problems in cognitive tunneling may result from the cognitive difficulty in switching between different frames of references for one task.

This difficulty in dividing attention between two tasks represented by overlapping imagery has been examined in the HUD and HMD domains (Larish and Wickens, 1991; Martin-Emerson and Wickens, 1992; Martin-Emerson and Wickens, 1997; McCann, Foyle, and Johnston, 1992; National Research Council, 1997; Wickens and Long, 1995); the results often show that near domain information presented on the display captures the attention of the user at the cost of detecting unexpected events in the visual scene. As an example of this cognitive tunneling phenomenon, Fischer, Haines, and Price (1980) and Wickens and Long (1995) showed that pilots failed to notice an airplane on the runway during a landing task when using a HUD display. The National Research Council (1997) describes a case reported in the Soldier Integrated Protective Ensemble (SIPE) in which soldiers were unable to see an ambush target even though the target presented itself upon numerous occasions. The squad was so confident in their ability to observe the kill zone using their HMDs that they actually positioned themselves further from the site. They hypothesized that they missed the target because they were so focused on the kill zone that had the target appeared in the periphery of their field of view, it may have been easily missed.

One solution to the problem of cognitive tunneling may be the use of conformal symbology. Atchley, Kramer, and Theeuwes (1997) showed that linking information displayed at two different depth planes with one object spanning the two planes aided divided attention tasks between near and far domains. Subjects wearing stereoscopic glasses were shown displays of two pipes, each of which could either be located at depth planes or both extend across two depth planes. Subjects were asked to search for two defects (small, green droplets) in the pipes, configured so that defects appeared on same or different pipes at same or different depths. The results showed that subjects were faster and more accurate at responding to the task when the defects were located on the same object and were slower when the defects were presented on pipes at different depths *and* on different objects. That is, the cost in dividing attention in depth was not present when the defects occurred on one object, which linked the near and far depths.

In the aviation domain, Foyle, McCann, and Shelden (1995) examined the use of scene-linking to reduce cognitive tunneling by asking subjects to maintain their altitude and follow a ground path in conditions in which altitude information was superimposed on a computer generated scene at fixed locations on the display (that is presented non-conformally), or conformally displayed along the flight path. A control condition was included which presented no altitude information on the display. The results showed that when non-conformal symbology was used, the availability of altitude information improved performance on the altitude maintenance task relative to the control condition but resulted in poorer performance in following the ground path. However, this trade-off in performance between the domains was not present for the conformal (or scene-linked) displays; rather performance on both tasks improved relative to the control condition.

Attentional factors in HMD use must be examined not only in the context of near versus far domain access of information but also in the context of performing dual tasks; in other words, whether one can attend to the symbology on the HMD and the outside world while simultaneously monitoring for information in the world in areas not currently present in the HMD field of view for obstacle detection. For example, consider the task of land navigation in which the HMD user would need to walk through unfamiliar territory and scan for obstacles in the path while monitoring for navigational information displayed with an HMD. This scenario was simulated by Sampson (1993), who asked subjects to stand or walk on a treadmill with or without obstacles while simultaneously performing a reaction time task using stimuli presented on monocular, opaque HMD. The primary task of obstacle avoidance was completed under four different conditions: standing, walking without obstacles, walking with obstacles on the left, and walking with obstacles on both the left and right. A buzzer sounded when subjects stepped on an obstacle. In all conditions, obstacles were presented with orange tape on the treadmill on both sides of the subject, but the particular experimental condition determined whether certain obstacles could safely be ignored (i.e. stepped on without sounding the buzzer). The reaction time tasks required subjects to press a key on a numeric keypad when given instructions verbally (e.g. north-west), numerically, or spatially. The results showed that the time required for subjects to complete tasks on the HMD was poorer when they needed to monitor for obstacles on the treadmill than when simply walking. The different types of reaction time tasks had no influence on obstacle avoidance. No difference was observed between the need to monitor for obstacles only on the left versus the monitoring of obstacles on both sides; however, this result may be attributed to the fact that the task requiring subjects to avoid obstacles on the left also required them to ignore the obstacles on the right. While the results implicate a cost for dual task performance with overlapping imagery, they cannot inform as to how much of that cost was attributable to the overlap.

Research by Seagull and Gopher (1997) suggests that with training, one can be taught to increase head movement in order to better scan one’s environment. Pilots were asked to fly through a simulated canyon environment displayed monocularly using a one-eye see-through HMD or binocularly on a display screen. The experiment consisted of a pre-training phase, in which pilots flew through the canyon simulation, a training phase, in which pilots flew through the canyon while performing a secondary task designed to encourage head movement, and a post-training phase, in which pilots flew through the canyon as in the pre-training phase. Note that no secondary tasks were presented in the pre- and post- training phases. In one training condition, pilots were asked to capture a target, in which a diamond-shaped object (target) was presented at different distances from the center of the canyon, and subjects needed to turn their head in order to position a square reticle, appearing in their forward field of view, over the target while flying through the canyon. In a second condition, the secondary task required not only target capture but also head re-orientation – once the target was captured, pilots need to move their head back to the center of the direction of travel. Two other conditions were included to compare the effectiveness of the training – one in which pilots were asked to perform an irrelevant secondary task which did not require head movement and another in which no secondary task was presented. A comparison of flight performance in the post-training phase showed better performance from those subjects who had been trained to move their heads, and no difference between the two training conditions. In fact, subjects who had not been given the same training minimized the amount of their head motion. The use of an irrelevant secondary task, which did not train head motion, did not improve performance relative to those conditions in which head movement was trained. Thus, these findings suggest that in addition to using symbology which facilitates focused and divided attention tasks, training subjects to divide their attention can prevent cognitive tunneling and aid dual task performance.

3.3 Summary

Superimposing information from the near domain onto the far domain is motivated by two factors: eliminating scan time and preventing eye accommodation when switching between symbology displayed in the near domain to that on the far domain and vice versa. Unfortunately, the presentation of near domain information on the far domain increases the amount of clutter in the forward field of view, making it difficult for one to focus on any particular object. On the other hand, if one focuses too much on one domain, ignoring events in the other domain, cognitive tunneling results such that attention is not divided effectively between the near and far domains.

These problems are solvable with the use of conformal symbology, in which the two worlds (near and far) are “fused” to better allow the user to treat both domains as one. In the task of land navigation, the HMD user must be able to not only divide attention effectively between the display and outside world but also scan for hazards in the outside world. However, this task is difficult, as was shown by Sampson (1993). Attentional considerations in the design of the HMD must thus take into account the prevention of cognitive tunneling. It should be noted, however, that cases of cognitive tunneling are rare and the overall advantages to presenting information head-up greatly outweigh the costs of missing surprising and, most importantly, infrequent events. However, so far, little work has been done examining the latter issue. This question is especially important for tasks of land navigation, e.g. those of the foot soldier who must walk through hazardous environments while monitoring for friend or foe positions, some of which may be unexpected.

4. One Eye vs. Two Eyes

The issue at hand is whether using two eyes to view an HMD really are better than one. In terms of complexity, the monocular display is the simplest as it requires only one image source and one set of optics and is cheaper and lighter than the biocular or binocular displays. Advantages for the monocular configuration are a wider field of view of the far domain due to lack of obstruction of the free, or uncovered, eye, which is hypothesized to allow for better target detection performance in the periphery (Cuqlock-Knopp, et al., 1996) and for greater safety when operating under low illuminations (Kooi, 1993; Lippert, 1990). The latter advantage is due to dark adaptation, in which the uncovered eye remains adapted to the dark illumination of the outside world. The use of a high contrast HMD image against a dark low contrast background as is common in two-eyed HMDs increases the amount of visual effort needed to use the HMD and as a result, will impair the long term comfort of the device (Lippert, 1990).

Disadvantages to the monocular display, such as the lack of depth information, the potential for binocular rivalry, and the small amount of space for information display, offset its benefits (National Research Council, 1997). Both biocular and binocular transparent HMD configurations are believed to give the user greater visual comfort, improve detection and recognition of obstacles and targets as a result of increased range resolution performance in the forward field of view, and require less training than the monocular display (Blake and Fox, 1981; Lippert, 1990). The biocular display is slightly more complex than the monocular HMD as it requires a second set of optics, thus making the display slightly heavier than the monocular one, but eliminates the problem of binocular rivalry through two-eyed presentation of data. An advantage to this configuration is a wider field of view for the display of information relative to the monocular configuration (National Research Council, 1997).

The comparison of one-eyed versus two-eyed viewing would not be complete without taking into consideration the addition of stereopsis. The binocular HMD is the most complex of the three, requiring two sets of optics and two image sources. This display is the only one which allows for stereoscopic viewing and three-dimensional (3D) depth perception (Davis, 1997), and as a result is the heaviest and most difficult to adjust to the viewer (National Research Council, 1997). The benefits of stereo increase as the quality of visibility conditions decrease. Stereoscopic viewing in non-HMD environments has already been shown to aid tasks of exploration (Cole, et al., 1991), teleoperation (Drasic, 1991), manual tracking (Kim et al., 1987; Sollenberger and Milgram, 1993), neurosurgery (Sollenberger and Milgram, 1989), and spatial awareness (McCormick and Wickens, 1995). The motivation for stereo viewing includes improved spatial perception of the HMD depicted artificial scene with better visual filtering of noise, enhanced image quality, better object recognition, less training time, and greater user satisfaction (Davis and Hodges, 1995; Drasic, 1991). In the design of HMDs, Davis (1997) lists scenarios in which binocular viewing is more effective than biocular viewing. These are:

- the presentation of a visual scene in an egocentric, perspective view rather than an exocentric view
- the presence of monocular cues which provide ambiguous information that could be presented more effectively in stereo
- the use of a static display rather than dynamic one
- the presentation of ambiguous objects and complex scenes
- the tasks to be performed require ballistic movement or accurate manipulation of objects within the virtual environment

It is important to mention that not everyone can see in stereo; there are also those who have stereopsis for objects behind the point of fixation but not in front of it or vice versa (Davis, 1997).

The benefits and costs associated with the monocular, biocular, and binocular HMD configurations will be examined in greater detail. We will first describe the occurrence of binocular rivalry, the result of which can be an unstable view. Experiments examining binocular rivalry as well as potential solutions to the problem will be discussed. The section will then turn to comparisons of monocular, biocular, and binocular displays for tasks performed in the air and on land.

4.1 Binocular Rivalry

The presentation of HMD information with a monocular display can range from simple symbology such as a tracking cross (Gopher, et al., 1992) or a single bar used to indicate boundary limits (Williams and Parrish, 1990) to a map display (Marshak, 1997); the second eye, meanwhile, views the world directly. The visual system is presented with two different, functional images, which it attempts to fuse to form a single one. *Binocular rivalry* is the failure of this process. If the difference between the images is large, then the visual system may not be able to fuse the images. Although the visual system tries to repress the visibility of one image through *binocular suppression*, over time, the dominant image may shift from eye to eye, so that the two monocular views will appear as alternating images (Arditi, 1986; Davis, 1997). Thus, the images perceived may be unreliable (Kooi, 1993).

In general, the dominant image will be the one with greater intensity, contour, contrast, and motion. When presenting images using a monocular HMD, the stronger image will often be that on the display since it will be presented at a much stronger intensity than the scene viewed by the unoccluded eye. It may not be ideal to determine beforehand the eye to which the image should be displayed. Rather, the display should be presented to the eye with higher acuity to ensure that the HMD image will be the dominant view (National Research Council, 1997).

The occurrence of binocular rivalry is not limited to the monocular display but may occasionally occur with binocular configurations of synthetic images. The occurrence is rare, as the visual system will try to fuse the images based on any matching features in the input received by each eye. Only when this fails will rivalry result (Blake and Boothroyd, 1985).

Kooi (1993) examined various techniques and perceptual phenomenon to minimize the occurrence of binocular rivalry and enhance the tendency for image fusion. The goal of the study was to discover a way to allow one eye to remain dark adapted while displaying information on binocular displays. Methods examined included:

- *blur suppression*, which is configured by bringing one eye into near focus and the other into focus at infinity. Although the eye focused at infinity remains out of focus, the sharp image received by the eye in near focus will suppress the blur seen by the other eye.
- *motion fusion*, which, as the name implies, is the use of motion to aid fusion; two monocularly presented random dot patterns will fuse more readily if both are moving in similar directions. In addition, since binocular rivalry occurs over time for static images, the use of motion will prevent rivalry by continuously changing the image.
- *natural windows*, a technique based on the assumption that the visual system can tell whether the presentation of two monocular images conforms to real world expectations. For example, introducing a “frame” into the view presented to one eye allows the visual system to interpret the two different images to be one viewed through a window, thus making use of a naturally occurring occlusion cue.
- *brightness averaging*, the presentation of two monocular images at two different levels of brightness, which can then be fused into one image with a brightness equal to the average of the two different images.

Kooi examined some of these factors in a series of experiments, in which he collected physiological and subjective ratings, to determine how best to configure monocular images so that two different images could be fused to form only one. Images presented from two separate monitors were displayed separately to each eye and were evaluated on the basis of ease of binocular fusion, occurrence of binocular rivalry, and subjective visual comfort of the image. Subjects gave favorable reports for the use of blur suppression but found that while the addition of a window frame served as an effective occlusion cue, it was not very comfortable, rating between the "barely acceptable" and "acceptable" levels for visual comfort. When two images varying in luminance were presented, physiological measures showed greater sensitivity when one eye remained dark-adapted, and subjects reported a surprisingly high degree of visual comfort, which the author attributes to the short viewing time, i.e., 3-5 minutes.

Binocular rivalry may also be minimized with the use of conformal symbology such that information is superimposed on the world in such a way that is *ecologically valid*. Shimojo and Nakayama (1990) suggest that displaying information non-conformally on the forward field of view can occlude objects in the far domain in two ways: ecologically valid or invalid. When non-conformal objects in the near domain are superimposed on objects in the far domain, some of the information in the world is occluded as a consequence. Interocularly unpaired regions result such that areas on each side of an object are only visible to one eye or the other and offer no disparity information. If there is more occlusion by the near domain over the far domain towards the right field of view, the left eye will see more of the far domain. Stimuli which are positioned in depth between the two regions visible to only the left eye will be *opto-geometrically valid*, whereas stimuli presented to only the right eye would be *invalid* as this scenario is impossible in the real world. The reverse is true if the object in the near domain occludes the left visual field. Figure 4.1.1 presents an example.

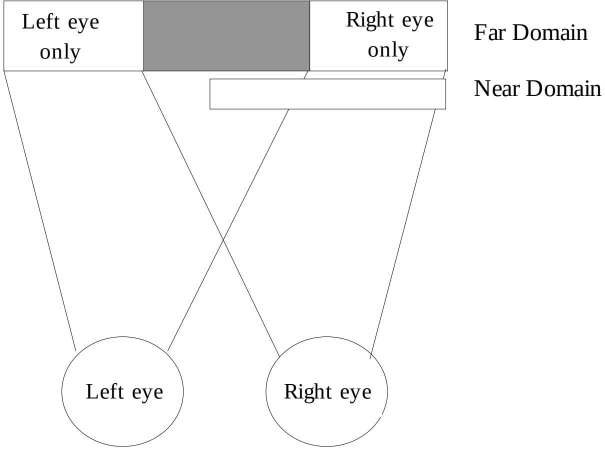


Figure 4.1.1. Occlusion constraints – object in the right visual field.

As Figure 4.1.1 shows, the far domain viewed by the right eye is occluded by an object in the near domain, so that the left eye sees more of the far domain. Objects appearing in depth between the near and far domains should be visible only to the left eye as the view from the right eye is occluded by imagery in the near domain. Objects appearing in depth between the two domains visible to the right eye is therefore impossible. The reverse is true if the imagery in the near domain is displayed to the left eye.

In an experiment conducted by Shimojo and Nakayama, subjects were presented with stimuli consisting of a square shaped region superimposed on rectangular backgrounds with either the left or right region presented monocularly. Subjects viewed the stimuli through a prism haploscope and were asked to indicate whether the monocular region faded in and out of view.

The results showed that subjects perceived fading in and out in the *invalid* conditions but not in the *valid* configurations of the stimuli. Since conformal symbology locates information in the near domain at ecologically valid points relative to the far domain, the occurrences of binocular rivalry may be prevented.

4.2 Monocular vs. Biocular vs. Binocular Displays

The efficiency of information processing has been one criteria used to evaluate performance with monocular or binocular displays. In other words, does each eye constitute a separate information channel such that information processing will be facilitated if tasks are decomposed and divided into the left and right visual fields, as is possible with monocular configurations? So far, results of studies reported below indicate no change in performance if information is presented to one eye or two.

In an experiment by Gopher et al. (1992), subjects were asked to navigate a flight path designated by tunnel imagery. In the binocular condition, a tracking cross and predictor square of the flight tunnel was presented to both eyes; in the monocular condition, the tracking cross and predictor square could be presented to one eye or the tracking cross could be presented to one eye and the predictor square to the other. The results showed no difference in tracking performance due to viewing condition (one eye or two), but more errors were made when the tracking cross and predictor square were presented to different eyes. Additionally, there was no benefit for tracking performance on the flight task with the addition of stereo.

A similar experiment was conducted by Rohaly and Karsh (1997). Stimuli consisted of a Snellen "E" presented face up, down, left, or right in the center of the screen and a silhouette side view of a tank in the periphery. Stimuli could be presented to one eye or divided dichoptically, one to each eye; e.g. the Snellen "E" and the tank was presented to one eye for the monoscopic condition, but in the dichoptic condition, the "E" was presented to one eye and the tank image to the other. Tasks performed varied in location (center or periphery) and workload. In the experiment, subjects were first given a task in the center of the display requiring them to indicate either the presence or absence of the Snellen "E" or to indicate its orientation. Once subjects completed this task, they needed to indicate the direction of a tank appearing in the periphery, which could be presented with low or high levels of clutter. The distance of objects in the periphery from the center of the display was varied in order to determine the useful field of view, the area of the visual field from which information can be acquired with a single glance. The results of the study showed that there was no difference in performance when the tasks were presented to one eye or two. Rather, performance on the target detection task was affected by the level of clutter, indicating that the useful field of view decreased with high levels of clutter from 30° to 10° or less.

HMDs have the ability to utilize human stereo because these displays have optics which can easily present two differing viewpoints. Stereo is one of thirteen different cues that humans use for judging depth. It is compelling (Wickens, Todd, and Seidler, 1989) but not more so than other cues such as motion parallax or occlusion. Its effects are likely to increase when these other cues are missing or impoverished (Pepper et al., 1983; Pepper, Smith, and Cole, 1981) or as the clutter on the display increases (Zenyuh et al., 1988), but not when depth cues are rich. Additionally, the effectiveness of stereo varies with distance; Davis (1997) reports that stereopsis is as effective as texture gradients in providing depth information at a distance of one meter but significantly less effective at a distance of two meters. In HMD environments, both monocular and biocular display configurations lack stereo.

Generally, comparisons of monocular, biocular, and binocular displays have not been examined in the framework of solely one eyed versus two-eyed presentation of information but is tied closely to the issue of non-stereo versus stereo viewing. Studies have examined the potential benefits for the addition of stereo on tasks ranging from divided attention to precision and control of movement to depth judgments in virtual environments. The results of these studies will now be discussed.

Moffitt (1989) collected ocular vergence and visual accommodation data for monocular, binocular, and one-eye occluded HMD imaging to study the effects of switching attention between near and far domain when using HMDs. Two subjects were asked to switch their attention between symbology and background at the sound of a tone and were instructed to scan the background as in target search, although no actual object needed to be found. The visual scene was presented using slides with different combinations of symbology (near domain) and mountainous or cloudy background (far domain). Analysis of eye tracking data showed that binocular viewing in contrast to the two monocular conditions resulted in more accurate vergence and more distant accommodation following the signal to switch.

Performance on path tracing tasks completed using monocular, biocular, and binocular viewing has been compared to determine whether there is an advantage to stereo for precision and control of movements in virtual environments. Hendrix, Brandt, and Barfield (1995) found a benefit for the binocular display in an experiment which compared stereo vs. non stereo – i.e. binocular versus biocular viewing, respectively – performance on a wire tracing task. Subjects were required to move a wire along the path of another wire without actually touching the path. In the stereo condition, the wire was projected out of the monitor towards the viewer. Results showed a performance benefit for the binocular display; subjects using stereoscopic viewing moved along the x, y, and z axes faster than those without stereo and thus completed the task faster with no loss in accuracy. However, Ellis et al. (1997) found no benefit to binocular viewing over monocular or biocular viewing on a similar task. Subjects were asked to trace a path which could be angular or smooth using a cursor that was manipulated with hand or head tracked controls. The path to be traced was presented with a haploscope at the same perceived depth for all three viewing conditions. The results showed that the lack of a viewing effect may have been due to the careful calibration of image position of the stereo stimulus between the monoscopic and stereoscopic conditions.

Comparisons of the three viewing conditions on tasks of depth judgments show a performance advantage for binocular viewing. Ellis and Menges (1995) examined the effects of the monocular, biocular, and binocular displays on depth perception using a see-through display. The stimuli consisted of a rotating tetrahedron placed 58 cm from the subjects' eyes and depicted at the same perceived depth for all three viewing conditions. Subjects viewed the display using a haploscope and adjusted a physical cursor to the estimated location in depth of the tetrahedron. The initial position of the cursor was either 25cm away from the subjects' eyes or 90cm away. The manipulation of viewing condition (monocular, biocular, or binocular) had a significant effect on distance judgment as subjects were more prone to overestimate depth as the depth cues were degraded, i.e. from binocular to biocular to monocular.

In the aviation environment, Williams and Parrish (1990) examined the presentation of monocular, biocular, and binocular images for a 3D tracking task performed in conjunction with a secondary monitoring task using a monitor with a field of view constrained to simulate that of an HMD. The primary task required subjects to fly above a specified path, which was presented binocularly in stereo or biocularly without stereo, while monitoring the status of three bars presented monocularly, biocularly, or binocularly in the periphery of the near domain. The addition of stereopsis in the primary task allowed objects (e.g., trees) in the far domain to modulate in depth in front of, behind, or at the same distance as the display. The addition of binocular viewing in the secondary task was used to examine the effects of stereo as a cueing factor such that the bar about to exceed the boundary limits would modulate out in depth toward the subject. The results showed a benefit to stereo in the primary flight path tracking task, which the authors hypothesized to be the result of the availability of more depth information with stereoscopic viewing, thus giving pilots a better idea about their present situation relative to the flight path and future position. In fact, subjects' comments supported this idea. Performance in detecting boundary excursions on the secondary task improved by 9-10% when information was displayed to both eyes as opposed to only one.

Andre and Johnson (1992) attributed performance benefits achieved by the use of stereo to the lack of monoscopic cues and conducted an experiment in which monoscopic depth cues were added in an attempt to minimize performance differences between biocular and binocular displays. Subjects were asked to perform hover maneuvers based on information presented on an opaque HMD, viewed biocularly or binocularly. The level of detail of the scene – that is, the degree to which objects in the simulation aided judgments of vertical, lateral, and longitudinal distance – and level of detail in the ground, the amount of information in the presentation of the surface, which could be either amorphous (without sharp edges) or a patterned grid, was varied. The results were mixed: stereoscopic viewing was a detriment to performance for low hover over detailed ground textures and high hover over less detailed ground, but there were benefits to performance with stereo viewing when the task required low hover maneuvers and salient monocular cues were not present.

However, Eggleston, Janson, and Aldrich (1996) found that the use of stereoscopic imagery altered perceptual judgments. Subjects performed tasks with biocular and binocular HMDs to determine whether there were display effects for judgments of size and distance; that is, whether depth information could be ignored when judging size. The task to be performed required subjects to sit at the intersection of two virtual corridors. A cylinder presented in one corridor at a fixed distance served as the target object, and subjects were required to change the size of the target relative to an object (stimulus) presented in the second hallway at three different distances. In natural viewing, people show *size constancy*. The perceptual system automatically judges the distance of objects and uses this information to perceive an object as being the same true size, independent of their distance from the viewer (which changes the retinal image). However, when depth cues are degraded, this compensation for viewing distance is less effective and the perceived size of an object is more likely to be influenced by the retinal image as the object is viewed at different distances. In a puzzling result, the authors found that *providing* stereo actually led to less size constancy; this difference may have been due to variations in illumination or texture in the virtual world which was dissimilar to the real world.

Ellis and Bucher (1994) found that the addition of stereo into a visual scene altered depth perception for a retinal scene in the background. The task presented to subjects required estimates of distance perception for virtual objects presented with physical objects. Subjects were presented with a virtual stereoscopic image of a rotating tetrahedron (the target object) on a haploscope and asked to move a pointer to the position in space reflecting the perceived distance of the target. For some trials, a physical object (a checkerboard) was presented with the image either at the same location of the tetrahedron's perceived position – and therefore occluding the target – or in front of the perceived target position. The checkerboard was presented statically or rotating with the tetrahedron. Subjects perceived the tetrahedron as moving closer towards them when the checkerboard was not occluding it. Additionally, the perceived distance to virtual objects was five times greater when the checkerboard was rotating versus when the checkerboard was static. The change in position of the virtual image was hypothesized to be the result of changes in ocular convergence required to bring the object into focus on the retina rather than interposition cues, e.g., occlusion.

The task of obstacle detection and avoidance is important for ground navigation and relies upon accurate perception of the far domain through the HMD. Thus, it is important to examine HMD use in such environments. CuQlock-Knopp et al. (1995) compared monocular, biocular, and binocular HMD configurations for off-road terrain navigation tasks. To determine whether the two-eyed presentation disrupted object perception as it prevented dark adaptation, the tasks were performed in varying degrees of light – 3/4 moon or no-moon conditions. Subjects wore either of the three different types of night vision goggles which presented an aided view of the terrain, and navigated through three different courses. Dependent variables included:

- total time to complete the navigation course
- errors resulting from contact with eye level hazards, ground level hazards, or terrain contour hazards; marked decrease in walking pace; request for assistance; stop; or stumble
- subjective ratings of each goggle

In the 3/4 moon condition, subjects navigating the environment with binocular goggles completed the task faster and with greater accuracy than those wearing the monocular and biocular goggles. There were no performance differences between the monocular and biocular goggles. In the low illumination condition (no moon), subjects using binocular goggles made fewer errors than with the other two goggles. Additionally, there was a marginal benefit in response time for the binocular goggles, but no difference in either accuracy or response time between the monocular or biocular goggles. Subjective rankings confirmed the performance benefit for the binocular display; under both moon and no-moon conditions, the binocular configuration was preferred over the biocular and monocular configurations for its utility in depth perception, comfort level, target detection, and environmental awareness.

However, one criticism of the experiment conducted by CuQlock-Knopp, et al. (1995) was the fact that they did not take advantage of the wide field of view of the unobstructed eye, i.e., the navigation tasks performed by the subjects required little scanning of the environment. It is possible, then, that the proposed benefit of the monocular display – greater peripheral vision – was not taken advantage of in the first study. As a result, a second study was conducted which was identical to the first except that a target detection task requiring subjects to scan the environment around them was added. This task consisted of detecting moving human targets or static inanimate targets in the far domain. The findings show that when no light was present (the no-moon condition), subjects performed faster and more accurately with binocular goggles, although there was no difference between the goggles for target detection. In moonlight, there were no differences between the goggles in number of errors or time to complete the task, but contrasts showed that more human targets were detected in the monocular condition than in the biocular condition. Subjective rankings favored the binocular display, which rated highest for the perception of ground level hazards, perception of eye level hazards, perception of terrain contour, target detection, sense of confidence, visual comfort, and timeliness of hazard perception (CuQlock-Knopp, et al., 1996).

The results of the CuQlock-Knopp, et al. (1996) are mixed; that is, the benefit of each type of display comes with some costs to performance. While monocular viewing allowed for better target detection of objects in the periphery, binocular viewing allowed for a better sense of hazard awareness. It is possible that the benefits of stereopsis may be only to supplement depth information, missing due to the lack of monocular cues (Ware, 1995). It is also possible that the poorer performance associated with monocular displays may simply be the result of presentation to only one eye versus two. If each eye is considered independent of the other, and each has the same probability of detecting an object, then the chances of finding the object will be greater when using two eyes versus only one via probability summation (National Research Council, 1997).

As a side note, the benefits of stereoscopic viewing are unclear when it is used in conjunction with a world-referenced display. Before describing the studies, it should be noted that none of these experiments examining stereoscopic viewing and head-tracked displays were conducted in an HMD environment. Arthur, Booth, and Ware (1993) found an additive effect of stereo and head-tracking when they measured performance on a tree tracing task using fish tank virtual reality, in which extra hardware supporting stereoscopic viewing and head tracking is added to a workstation monitor allowing objects to be presented in front of or behind the screen. The results showed lower response times and higher accuracy on the task with a combination of head tracking and stereo than with either head tracking or stereoscopic viewing alone. However, Rekimoto (1995) found that the combination of the two increased the time required to perform the task relative to the use of one factor alone on a similar task to that used by Arthur, Booth, and Ware, although the head-tracked display did improve accuracy.

Although Arthur, Booth, and Ware (1993) and Rekimoto (1995) found accuracy benefits for a head-tracked stereoscopic displays, Barfield, Hendrix, and Bystrom (1997) found improvements in accuracy with a head tracked display or a stereoscopic display, but not both, on a task which required subjects to compare a 3D wire image displayed monoscopically or stereoscopically on a monitor with or without head tracking to a 2D representation of the wire. The accuracy data confirmed that visualizing the 3D structure of the wire benefited from the use of stereopsis or head-tracking, but when one factor was present, the presence or absence of the second factor had little effect.

4.3 Summary

The literature review showed no clear advantage for any of the three display configurations. Table 4.3.1 summarizes the results of the literature in terms of tasks performed.

Each entry within the table represents a study whose identity is shown at the bottom of the table. The study is placed within the column showing the best display. Note that only a few experiments compared all three configurations – monocular, biocular, and binocular. In fact, most of the results described in the literature review were from experiments comparing biocular and binocular configurations, i.e. stereo vs. non-stereo.

Benefits of the addition of stereo for HMDs were present in tasks which required *maneuvering* along a path, whether it be a flight path or wire frame. Comparisons of *depth perception* and *monitoring* tasks showed no advantage to the addition of stereo but a benefit for presentation to two eyes over only one eye. In *target detection*, there was no difference between the monocular, biocular, and binocular configurations in dark illumination, but an advantage for monocular and binocular configurations over the biocular display in lit environments. There was no advantage for any of the three displays on the *wayfinding* task. Subjective preference results were mixed but seemed to favor the display of information to two eyes over only one. The results suggest that each configuration seems best suited for a specific situation, but no one display is optimal for all situations. Additionally, the effectiveness of the display may be influenced by the task environment itself, e.g. amount of light or amount of contrast between the world and the display.

Task	Monocular	Biocular	Binocular	No Difference
Attention Switching				
biocular - binocular			(8)	
Judgments of Depth				
binocular - biocular		(4)	(7)	
Maneuvering (Tracking, Tracing, Peg-in-hole task)				
monocular - binocular				(9)
biocular - binocular			(1)	
monocular - biocular - binocular			(10)	(5)
Monitoring				
monocular - biocular - binocular		(10)	(10)	
Target Detection				
dark				
monocular - biocular - binocular				(2) (3)
light				
monocular - biocular - binocular			(2)	(3)
monocular - biocular	(3)			
Wayfinding				
monocular - biocular - binocular			(2)	(3)
Subjective Preferences				
monocular - biocular – binocular			(2) (3)	

(1) Andre and Johnson (1992)
(2) CuQlock-Knopp, et al (1995)
(3) CuQlock-Knopp, et al (1996)
(4) Eggleston, Janson, and Aldrich (1996)
(5) Ellis, et al. (1997)
(6) Ellis and Bucher (1992)
(7) Ellis and Menges (1995)
(8) Gopher, et al (1992)
(9) Moffitt (1989)
(10) Williams and Parrish (1992)

Table 4.3.1. Summary of results comparing monocular, biocular, and binocular presentation of information.

5. Summary

The configuration of HMDs must be evaluated with respect to the display taxonomy set forth in Figure 1.3.1. Melzer and Moffitt (1997) note that the selection of an HMD may differ based on the requirements of the task to be performed. For example, the HMD selected will be different if the task is a nap-of-the-earth reconnaissance mission versus a preliminary virtual architectural walk-through of a design. In the first scenario, the HMD should be see-through and world-referenced to provide the pilot with information in the forward field of view as to what objects are in the far domain (possibly enemy targets) and where they are and to allow for unconstrained head movement for target detection. Additionally, a monocular view may be beneficial for nighttime missions, allowing one eye to remain dark-adapted. However, when selecting an HMD suitable for the latter case, the presentation of objects in the far domain would contribute little and might actually impair the task of the architect viewing his design. In this case, the HMD should be binocular and opaque, immersing the architect into the virtual environment to allow him to interact within the space to determine whether building specifications have been met.

The results of the literature review show a clear benefit in *display* for the presentation of information using an HMD or HUD over that of a head down display due to time savings from the reduced scanning distance. The presentation of information conformally to the world (i.e., world-referenced) also increases the benefit for head-up/helmet-mounted presentation relative to the head-down condition; its advantages are not only reduced clutter in the forward field of view due to the use of conformal symbology but also the prevention of binocular rivalry. Technology limitations of the system underlie the choice and implementations of the two different frames of reference, however. In developing a world-referenced display, the technology needs to provide for highly accurate real time tracking of head position and updating of images as information on the display is slaved to the user's current head orientation. With screen-referenced displays, however, the presentation of symbology is independent of head position and orientation. The use of world-referenced symbology may aid tasks of focused and divided attention. By linking imagery in the near domain to imagery and events in the far domain as in the world-referenced configuration, the user will be able to search the display and find the information needed faster relative to the screen-referenced configuration. However, questions persist as to the feasibility of a helmet mounted sight for the use of land navigation.

Evaluations of HMDs have cursorily examined dual-tasking scenarios such as whether the user will be able to walk or navigate with the HMD given that he may be monitoring the terrain while monitoring information on an HMD. The results of Sampson's experiment (1993) indicated that monitoring for obstacles in the world and performing tasks with an HMD simultaneously may lead to missed obstacles – the results of which could be dangerous, e.g. stepping on a land mine, although this experiment did not compare HMD performance to that with a head-down display. In a comparison of monocular, biocular, and binocular HMDs, CuQlock-Knopp, et al. (1996) found that errors in navigation were reduced with the use of binocular displays when no light was available but that there was no performance difference between the three displays in the moonlight conditions. It is possible that monocular cues present in the environment were visible only in the 3/4 moon conditions but not in the no-moon condition, thus preventing the subjects from effectively distinguishing objects and perception of edges within the environment.

Relatively few experiments have compared monocular to biocular viewing, i.e. the advantages for one eye relative to two eye information displays. Williams and Parrish (1992) found a benefit to two eyed viewing on a monitoring task and hypothesized that the findings were the result of probability summation. However, CuQlock-Knopp, et al. (1996) showed an advantage for the monocular display over a biocular display for target detection in land navigation, but this benefit may have been a result of the hardware used. That is, the experiment was conducted using night vision goggles so that the monocular configuration allowed one eye to remain unoccluded whereas the biocular goggles presented displays to both eyes. It is possible that the unconstrained field of view for one eye allowed for better scanning of targets in the periphery, aiding the target detection task.

The literature is lacking in experiments examining the effects of one eye versus two eyed viewing on the presentation of symbology – i.e., conformal or non-conformal. Additionally, information regarding how the use of conformal or world-referenced imagery facilitates performance on tasks of focused and divided attention while available for HUDs is sparse for HMDs. That is, will the user will be able to focus attention on information in the near domain (the display), the far domain (the world), as well as switch or divide attention between both domains; or alternatively, will he rely only on the symbology presented on the display and ignore obstacles in the world, or vice versa? Finally, while information for HUDs has revealed that event expectancy (surprise) can substantially influence the detectability, the issue of expectancy has not been addressed in the HMD domain.

In order to address these issues, a study was conducted in which subjects were asked to perform a target search and detection task using a simulated HMD, which could be either world-referenced or screen-referenced. Target search was manipulated by expectancy and priority, such that subjects searched for one of three expected targets in each trial with the potential for an unexpected (i.e., rare) target to be present in any of the trials. Subjects were instructed to prioritize search for the unexpected target over that of the expected targets. The target detection task was simulated as an Army scouting mission in unfamiliar territory. In order to determine whether symbology manipulations in the study would affect performance differently between experts (i.e. those with prior experience on a similar task) and novices, both Army personnel (experts) and civilians (novices) participated as subjects. Note that Army personnel had the advantage of actual field training so that they had in a sense been taught a strategy for scanning the environment for targets.

One potential advantageous aspect of world-referenced imagery is the ability to present target cueing in a position that directly overlaps the inferred location of a far domain target. Hence searching for the target, subjects were sometimes aided by the presentation of cueing symbology, displayed in either a world-referenced or screen-referenced manner. The unexpected target was presented in conjunction with both cued and uncued expected targets in order to examine whether the presence of cueing would direct subjects' attention to specific areas of the display, benefiting the detection task for the expected target but at the cost of missing the unexpected target. Since the unexpected targets were uncued, we hypothesized that detectability for the unexpected target when paired with a cued target would be lower than detectability for the unexpected target with an uncued target.

Symbology was presented monocularly or biocularly to determine whether an advantage for viewing condition was present. The ability to attend to information on the HMD as well as divide attention between near and far domains was examined by the inclusion of a near domain secondary monitoring task performed in conjunction with the visual search task. Note that the search task could be viewed as a focused attention task on the far domain when the target was uncued but a divided attention task between domains when the target was cued, such that symbology presented in the near domain aided the subject in finding the target in the far domain.

Data on how subjects scanned the environment (e.g,. the amount of time subjects needed to detect the target based on its location and the number of times the target passed through the field of view) was collected to determine whether search through the simulated environment was similar to that in the real world and to discover how display referencing and target cueing influenced the subject's strategies in target search. Once all the targets had been found, a posttest was administered in which subjects were asked to report the configuration of targets in the environment, in order to understand whether differences in display formatting affect overall terrain understanding.

6. Method

6.1 Subjects

Sixteen subjects (12 male, 4 female) participated in the experiment. Eight were civilian graduate students or staff at the University of Illinois; eight were Army personnel (6 worked with the Reserve Officers Training Corps at the University of Illinois, and 2 were part of the Army National Guard Reserve).

6.2 Task Overview

The task performed by subjects consisted of three stages: (a) target detection, (b) target identification, and (c) target heading. The target *detection* task (the primary task) required subjects to scan the display looking for any one of four target objects: three of the targets were presented on a total of 90% of the trials (30% each) and were therefore expected; the fourth was presented only 10% of the time and was unexpected. Subjects were not told which target to search for. While searching for the target, subjects were asked to perform a secondary monitoring task, displayed on the simulated HMD, which stopped once the target was found.

Cueing of the target's location was presented for half the expected targets to aid the detection task; the unexpected target was never cued. Targets were presented serially; only one target was displayed at any time, except in the case of the unexpected target, which was always presented in conjunction with an expected target. However, only one object was detected per trial. Subjects were instructed that detecting the unexpected target – a nuclear weapon – took precedence over standard target detection.

Once the target was found, the subject was required to *identify* the target as either friend or foe and give the target's *heading*, the current compass direction of the target with respect to his current location.

Subjects viewed symbology on half the trials monocularly and the other half biocularly. The far domain was visible to both eyes.

6.3 Apparatus

The terrain was displayed on the walls of the Cave Automatic Virtual Environment (CAVE), a 10x10x9-foot room sized video environment. The subject was seated in the center of the CAVE, as shown in Figure 6.3.1.

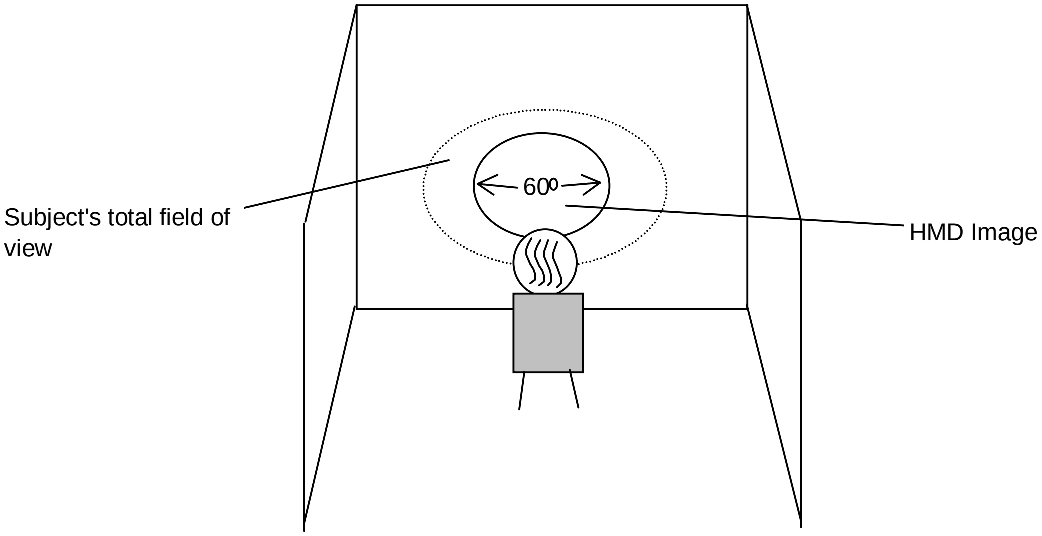


Figure 6.3.1. The subject, seated in the CAVE.

An actual HMD was not used in the experiment. Instead, subjects wore head-tracked shutter glasses, which could be used to display the symbology to one eye or two. HMD imagery was superimposed on the CAVE walls and constrained to a field of view of 60° laterally and vertically. Note that subjects' field of view was not constrained to 60°, that is, subjects could see far domain information in the periphery of the simulated HMD.

6.4 Displays / Tasks

The displays were created from static two-dimensional rendering of three-dimensional images depicting hilly terrain. The terrain was developed using geographical data of Austin, TX, Detroit, MI, and Jordan Valley, UT, downloaded from the U.S. Geological Survey web site. The target stimuli, shown in Figure 6.4.1, were placed in the terrain.

(a) Tank: Friend Foe (b) Soldier: Friend Foe

(c) Land Mine (d) Nuclear Device

Figure 6.4.1. Stimuli: (a) Tanks (cued) (b) Soldiers (uncued) (c) Land Mine (d) Nuclear Device

The tanks, soldiers, and nuclear devices were camouflaged, i.e. colored in shades of brown, green, and black; land mines were presented in black. Since the shading of the terrain varied, the intensity of the targets was adjusted adaptively at each location so that the contrast ratios between the target and the terrain were similar for all targets. The greater salience of the nuclear device was insured by presenting them at a higher contrast ratio with the background than the other three targets. The location of tanks and 50% of the land mines were cued with an arrow pointing in the direction of the target based on the subject's current head position. All soldiers and 50% of the land mines were uncued.

Friend-or-foe identification was dependent on the direction in which the target was pointing. Friendly targets pointed towards the left and enemy targets pointed towards the right. No identification was required if the target was a land mine or nuclear device.

Examples of the displays are presented in Figure 6.4.2, which depicts the field of view of the images, the horizon line, the cueing arrow, and the box containing the secondary task.

In Figure 6.4.2, the pictures show terrain and symbology presented on a CAVE wall. Symbology was presented in green by the simulated HMD and superimposed onto the wall. The visual region of HMD-depicted information was 60° laterally x 60° vertically.

Heading was presented either non-conformally (i.e., screen-referenced, Figure 6.4.2a), or conformally (i.e., world-referenced, Figure 6.4.2b) with respect to the horizon line. The four cardinal directions were marked on each heading tape. Note that the heading tape displayed in Figure 6.4.2a was constantly present on the HMD in a pre-determined location, whereas heading information in Figure 6.4.2b was superimposed on the true horizon line, and as a result, the location of the heading tape on the HMD changed as the subject moved his head vertically in order to examine the environment.

On cued trials, a cue was presented to signal the current lateral and vertical location of a target with respect to the subject's head orientation. For example, if the target was presented to the right of the subject, then a right pointing arrow appeared on the HMD as shown in Figure 6.4.2, indicating the presence and general direction of a target. If the target was above and to the right of the subject – e.g., located on top of a mountain, then an arrow pointing towards the upper right corner of the HMD appeared. Note that targets appearing directly in front of the subject within the forward field of view were not cued by this symbology but rather designated by a target lock-on indicator, or reticle, which will be discussed later.

Cueing information was presented with two levels of conformality – screen-referenced or partially head-referenced. An example of a screen-referenced display is shown in figure 6.4.2a. In this display, the cueing arrow is located at the bottom of the display, three quarters of the way down from the top, indicating symbolically the direction of the target. The left-right direction of the arrow represented the side on which the cued target was located in relation to the subject. Its angle of inclination represented the approximate angle (above or below) the subject's horizontal line-of-sight in the visual field. Thus, in this example, the target is to the right of and above the current orientation. Figure 6.4.2b shows a world-referenced display; in the experiment, a cueing arrow positioned on the perimeter of the screen display, pointed directly toward the 3D location of a target. In this example, the arrow would indicate that the target was to the right and above.

The cueing arrow could be positioned at the edges of the perimeter of a circle whose diameter subtends 40° of visual angle. In contrast to the screen-referenced arrow, this arrow could move continuously as the subject’s head orientation changed.

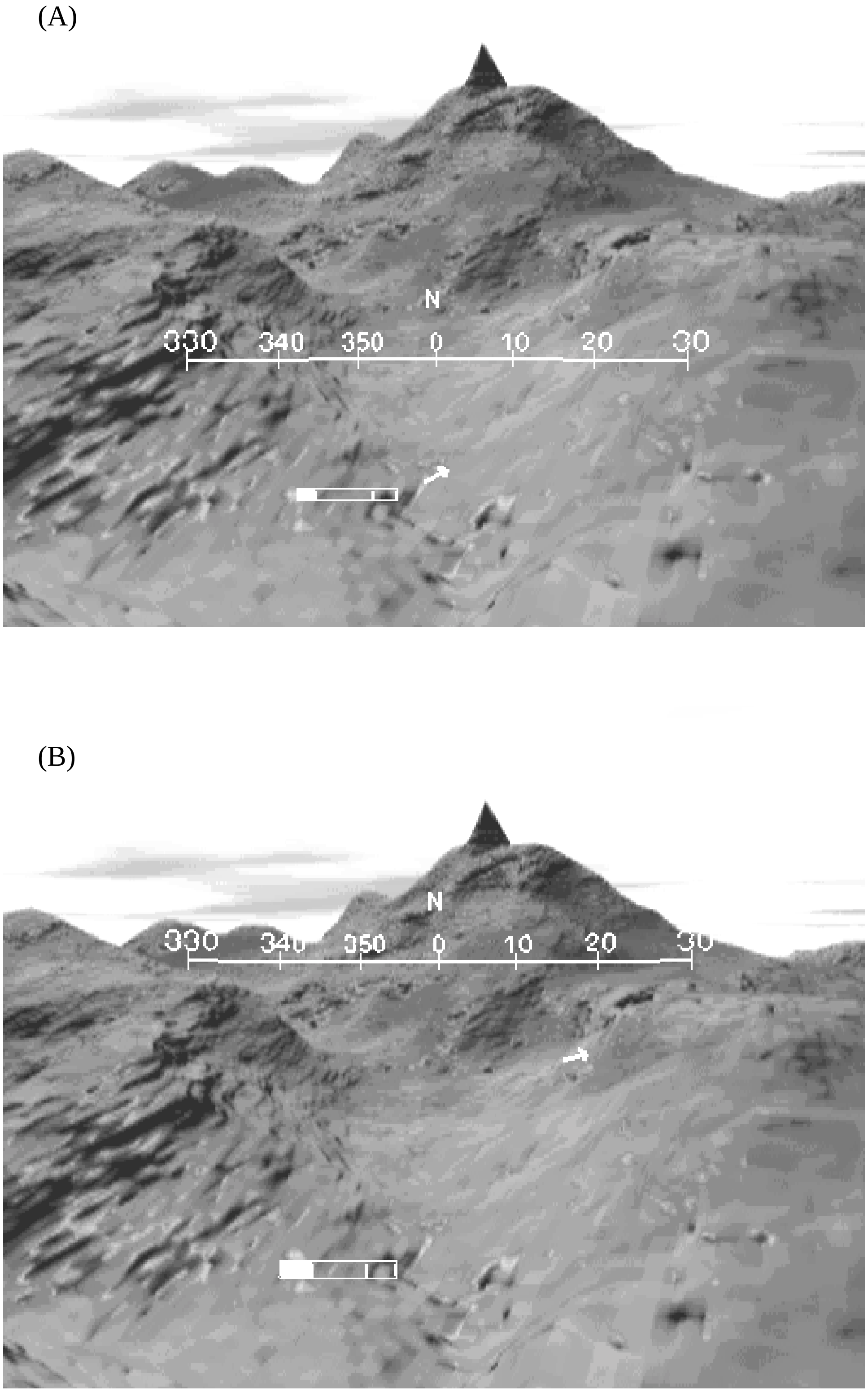
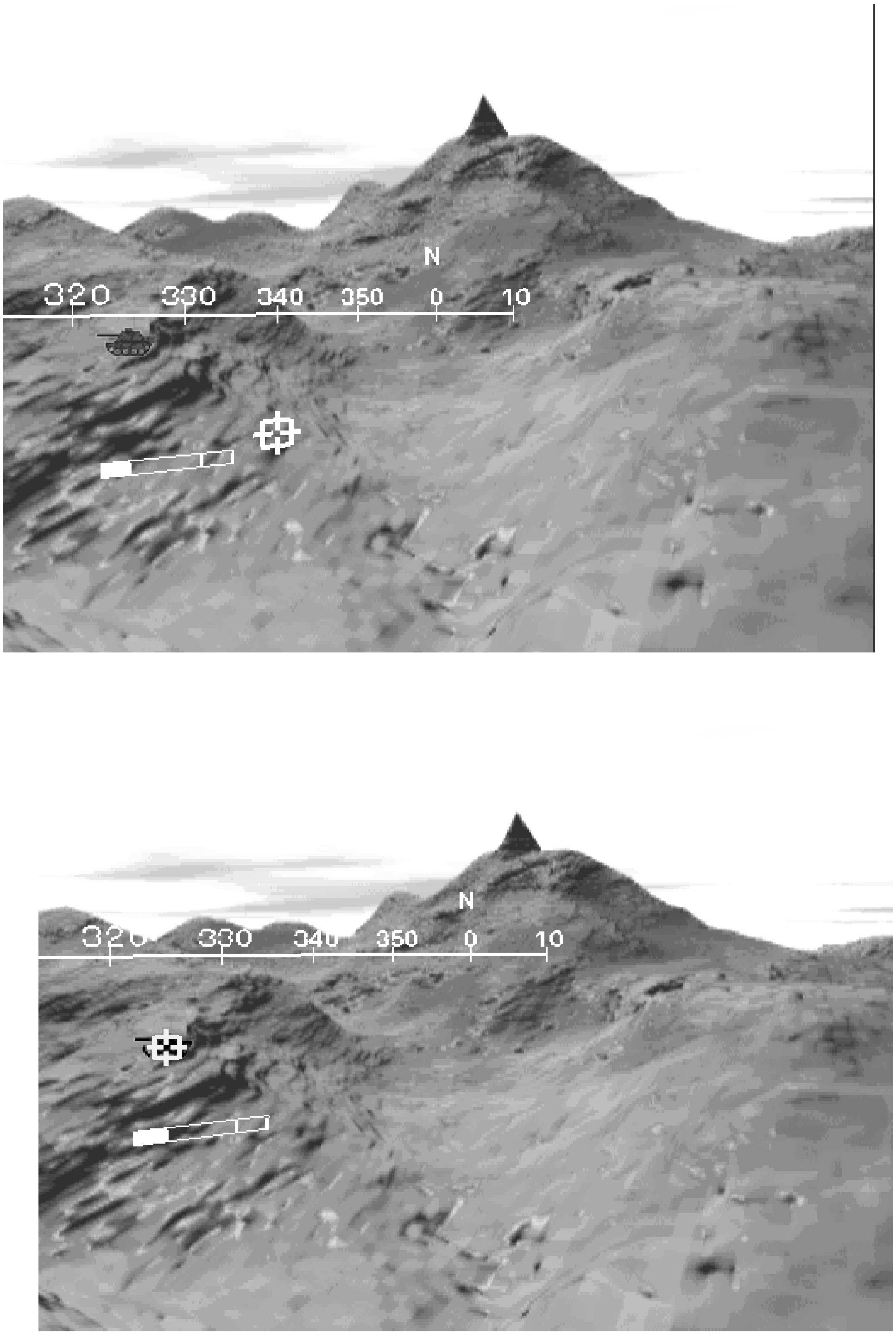


Figure 6.4.2. Displays – (a) screen-referenced symbology, (b) world-referenced symbology

Once the target was in the subject's field of view, i.e. visible through the HMD, a target lock-on reticle appeared on the display either non-conformally or conformally as shown in Figure 6.4.3.



(A)

(B)

Figure 6.4.3. Target lock-on. Note that the tank is in the far domain and is being viewed through the HMD. (a) Non conformal or screen-referenced. (b) conformal or world-referenced.

In the non-conformal condition (Figure 6.4.3a), the lock-on information appeared in the same location as the cueing arrow to indicate to the subject that the target was in his field of view. In the conformal display (Figure 6.4.3b), the same lock-on reticle was displayed over the actual object. The lock on reticle was *not* used to signal the presence of any uncued targets which might appear in the subject's forward field of view. The advantage of the lock-on symbology is that its occurrence signaled the presence of the target somewhere within the subject's field of view for the screen-referenced display and signaled the specific location of the cued target for the world-referenced display. However, the disadvantages for cueing are the increased clutter on the display and for world-referenced cueing, the potential for the symbology to obscure information crucial to the task, e.g., identifying markings such as the direction the object is facing, which the scout must examine to identify the target as friend or foe.

Subjects viewed the symbology (and the secondary task) monocularly or biocularly. In the monocular condition, symbology was presented only to the dominant eye, superimposed on the view of the CAVE wall, while in the biocular view, both eyes could look through the imagery to see the CAVE wall.

Subjects were given a secondary task, displayed only on the HMD, to perform continuously throughout the experiment. They were told that enemy troops were tracking their location by using radio frequency as input. Thus, as the subjects searched for targets, they were also required to jam the enemy's radar frequency so that they remained undetected. To do this, subjects needed to monitor a horizontal bar, presented at the lower left edge of the HMD (as shown in Figure 6.3.1). The solid bar gradually grew longer horizontally, filling in the rectangle from left to right. When it passed the first marker, subjects had 5 seconds to jam the enemy's frequency by responding with a button press. Responding before the solid bar passed the first marker had no effect. The solid bar increased at a variable rate created as the sum of four sine functions. The bar reached the first marker between three and five seconds from the start of the secondary task. Once subjects responded to the task, the bar would reset. The task continued until the target was detected.

6.5 Experiment Design

<table>
<tr><td></td><td></td><td></td><td colspan="2">Within Subjects
Viewing Condition</td></tr>
<tr><td></td><td></td><td></td><td>Monocular</td><td>Biocular</td></tr>
<tr><td rowspan="2">Between Subjects</td><td>Military
Civilian</td><td>World Referenced</td><td>TARGET TYPE
CUED | UNCUED
Tank Mine | Mine Soldier | Nuclear Device
← <i>Expected</i> → | <i>Unexpected</i></td><td>TARGET TYPE</td></tr>
<tr><td>Military
Civilian</td><td>Screen Referenced</td><td>TARGET TYPE</td><td>TARGET TYPE</td></tr>
</table>

The experiment was a mixed design as shown in Table 6.5.1.

Table 6.5.1. Experimental design.

As Table 6.5.1 shows, the presentation of display (world-referenced versus screen-referenced), and subject population (military versus civilian) were examined between subjects. The manipulations of target type (cued versus uncued targets, high versus low expectancy) and viewing condition (i.e., monocular versus biocular) was examined within subjects. The secondary task was present on all trials.

Six different terrains, created from taking static “pictures” at different locations of three cities, were used in the experiment. For each viewing condition, subjects were presented with one practice block, consisting of ten search trials, and ten experimental blocks, each containing a set of twenty search trials. The presentation of target stimuli (i.e. tanks, land mines, and soldiers) was serial – that is, only one target was presented per trial and subjects searched the three walls of the CAVE until it was located. The exception was the presentation of the nuclear device, which would appear concurrently with one of the other targets.

In the practice block, subjects viewed ten targets, presented serially. The targets consisted of three tanks, three soldiers, and three land mines, each, and one nuclear device. Tanks, soldiers, and land mines appeared once on each of the three walls. Each experimental block consisted of a total of 20 targets; 6 each of tanks, soldiers, and land mines and 2 nuclear devices. Half the tanks and half the soldiers were friendly – the other half were enemy. On half the trials, cueing was present. This was the case for all tanks and half of the mines. Thus, the presence of the cueing symbol provided subjects with a partial reduction of uncertainty of target type. Each object appeared twice on each wall, except for the nuclear device which appeared once on the left wall and once on the right wall. Targets were presented serially, except for the nuclear device, which was presented in conjunction with either a cued target (tank) or an uncued target (soldier). As it was an “unexpected” target, the nuclear device was presented within 15° of either a tank or a soldier.

6.6 Procedure

The experiment took approximately 2.5 hours during which subjects were given the instructions for the experiment and then performed the experiment. Subjects were instructed to pretend that they were scouts, sent to search for enemies and allies in unfamiliar territory. Their primary task was to find the targets, identify them as friend or foe, if relevant, and send information back to their troop regarding the objects’ position. Their secondary task, described below, was to monitor a radio frequency display, which provided data as to how close the enemy was in tracking their position.

Subjects interacted with the display using a wand and shutter glasses. A diagram of the wand is presented in Figure 6.6.1.

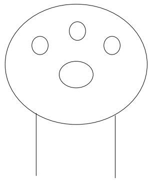


Figure 6.6.1. The wand.

The wand has three buttons and a pressure-sensitive joystick. Only the buttons were used during the experiment to make responses. The joystick was not used at all.

While searching for the target, subjects responded to the secondary task by pressing the right wand button. To indicate that a target was detected, subjects pressed the left button on the wand.

The target identification task required subjects to identify the target as friend or foe. Subjects pressed the left button on the wand again if the target was foe, the center button if the target was friendly, or the right button in the case of a nuclear device. Subjects did not need to identify whether the target was a tank, soldier, or land mine. Note that the button pressed for friend and foe identifications corresponded to the direction the object was pointing, e.g. subjects pressed the left button if the tank or soldier was pointing left. Once the target was detected (land mine) or identified (tank, soldier, or nuclear device), subjects verbally reported its location by stating the target's bearing.

Once the target was detected and reported, the display was darkened. When the subject's head was centered, a subsequent trial, containing a new target, was initiated.

After all of the targets were found within each twenty trial block, subjects were asked to "describe" the location of the targets within the environment to their commanding officer by selecting one of four pictures of the environment, one of which depicted the objects in the same location as in the environment they saw. Of the three incorrect pictures, one showed the tanks placed in different positions, another presented the soldiers in different locations, and the third depicted the land mines in incorrect sites. Not all the targets were presented. That is, targets presented on nuclear device trials were omitted from the pictures since it was not known which target subjects would detect in the nuclear device trials (i.e., would subjects see the missile or would the tank or soldier appearing with the missile capture their attention instead?).

6.7 Performance Measures

The dependent variables collected from the primary target search task were response time and accuracy for target detection, target identification, and target heading. In order to determine whether the symbology influenced the amount of scanning in the environment, data describing the amount of head movement along the x-, y-, and z- axes were collected. Additionally, data concerning the number of times and the amount of time the target was in the view (within 60°, 40°, and 15°) were collected. Note that the center points for the aforementioned view angles are at the center of the shutter glasses, rather than the center of the eyes. Thus, it was possible for a target to pass through the area in the center 15° of the shutter glasses and go unnoticed by the subject, if his eyes were rotated away from the forward axis of the head.

The measures collected from the secondary task were response time and accuracy. Since each subject took a different amount of time in detecting the targets, the number of frequency jamming tasks varied. Thus, accuracy for the task was calculated as a proportion of the number of hits to the number of total frequency jamming tasks viewed.

Finally, measures for the global positioning task were response time, accuracy, and subjects' confidence ratings of their responses.

7. Results

The data were examined in order to determine the effects of expectancy and cueing on target detection and how well attention could be allocated between the near and far domains. Differences in display (world-referenced or screen-referenced), viewing condition (one eye or two), and subject population (military vs. civilian) were factors hypothesized to mediate these effects. Since it was possible for subjects to mistake a terrain feature for an object, trials with heading errors greater than ±20° were scored as incorrect and replaced with the subject's mean response time for like targets in that particular block (i.e. involving the same terrain) displayed on the same wall. This was approximately 5% of the trials. Additionally, outliers which were greater than ±3 standard deviations from the mean were replaced in the same way.

The total data set represented ten dependent variables, consisting of response time and accuracy measures for the primary tasks of target detection, identification, and location, the frequency jamming secondary task, and the global positioning recognition task. These dependent variables were influenced by multiple factors (independent variables):

- Target type, which could be subdivided into comparisons of expected vs. unexpected targets and cued vs. uncued targets
- Display: world-referenced vs. screen-referenced
- Viewing condition: one eye vs. two eyes
- Wall: left, center, or right
- Subject group (military vs. civilian)

Because we do not hypothesize that all dependent variables would plausibly be influenced by all independent variables (or if they were, such influences would not be of theoretical or practical interest), we do not report full ANOVAs on all dependent variables. (These ANOVA tables can be found in Appendix 1). Instead, we parse the presentation of results into seven categories:

1. Effects of target type (expectancy and cueing) and display on the primary and secondary tasks (7.1-7.5)
2. Effects of viewing condition (7.6)
3. Effects of wall (7.7)

Within each of these sections, we present and describe only those effects (and their interactions) that are most relevant to understanding the influence of display augmentations on target detection.

7.1 Expectancy

How effects of the subject's expectation of a target influenced performance was examined by comparing detection performance for tanks and soldiers – both highly expected targets – with detection of nuclear devices – infrequent, low expectation targets. Although expectancy is confounded with physical differences between the stimuli, contrast adjustments were made to ensure that the unexpected targets (nuclear devices) were *more* salient than the expected targets. Subjects were also instructed that the former were of higher priority. The nuclear device trials were separated into two classes based on whether the nuclear device was presented concurrently with a tank or with a soldier. Although mines were also expected targets, the mine trial data were not used for this analysis since mines were cued on half the trials, thus confounding the measure of expectation. Note also that the presentation of unexpected targets never occurred with a mine. No comparisons were made between the two expected targets (tanks and soldiers) as variables affecting performance could not be attributed solely to cueing, i.e. tanks were cued and soldiers were not, but also be attributable to differences in the physical appearance of the stimuli. The direct effects of cueing will be examined in the analysis of mine detection (Section 7.2).

A 2 (display: world-referenced vs. screen-referenced) x 2 (subject population: military vs. civilian) between subjects x 2 (viewing condition: one eye vs. two) x 4 (target type: expected and cued (tank), expected and uncued (soldier), unexpected with cued (nuclear device presented with a tank), unexpected with uncued (nuclear device presented with a soldier)) within subjects ANOVA was conducted on the accuracy and response times for the target detection task. Figure 7.1.1 presents the effects of display and target type on response time (left) and accuracy (right). The bars in the figures show ±1 standard errors from the mean.

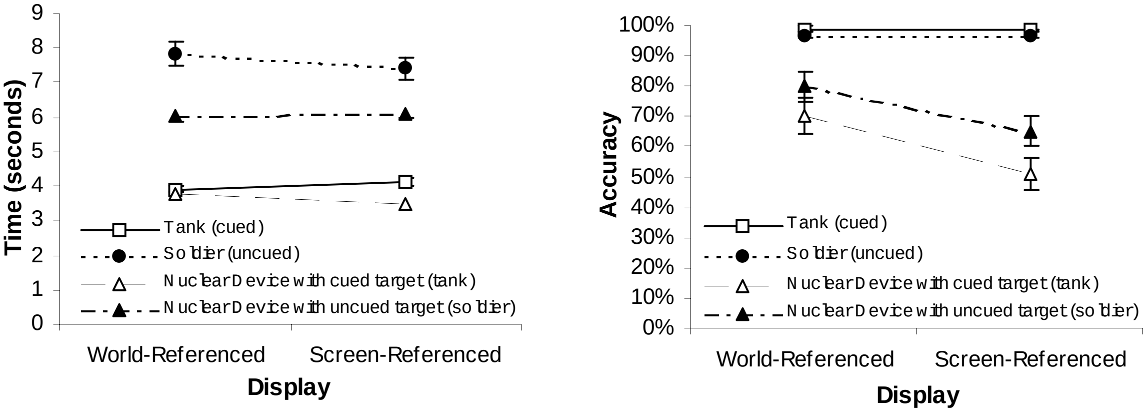


Figure 7.1.1. Response time and accuracy for expected and unexpected targets.

In the graph, the filled symbols are responses when the uncued target (soldier) was present; the open symbols represent responses when the cued target (tank) was present. The analysis for the response time revealed a main effect for target type, F(3, 36) = 59.35, p = .0001. Analysis revealed that the targets associated with the uncued trials (the expected soldier target and the unexpected nuclear target – the filled symbols) were both detected more slowly than targets during the cued trials (the expected tank), F(1, 12) = 26.40, p = .000 and F(1, 12) = 152.51, p = .0001, respectively. Within these uncued trials, the unexpected nuclear weapons were detected more rapidly than the expected soldiers [nuclear device presented with cued object vs. uncued soldier: F(1, 12) = 156.68, p = .0001, nuclear device presented with uncued object vs. uncued soldier: F(1, 12) = 12.92, p = .004]. Within the cued trials (open symbols), when a tank was present, subjects again detected the unexpected nuclear device slightly more rapidly than the expected (and cued) tank target, F(1, 12) = 14.02, p = .003. Comparisons of detection times for the two unexpected targets showed that the nuclear device on a cued trial was detected faster than the nuclear device on an uncued trial, F(1, 12) = 37.86, p = .0001. The data showed that detection of the different target types was not influenced by display referencing, F(2, 24) = .41, p = .66.

Thus, for the expected targets, the presence of cueing (of tanks) greatly reduced response time relative to the uncued soldiers. The unexpected nuclear devices were detected faster than the cued tanks when the two appeared on the same screen and were detected more rapidly than the uncued soldiers when they appeared on the same screen as the soldiers. However, the effects of expectancy (tank and soldier versus nuclear devices) were manifest differently in accuracy than in response time reflecting a speed-accuracy trade-off. The accuracy data shown in Figure 7.1.1 revealed a main effect of target type, F(3, 36) = 113.55, p = 0.0001, showing near perfect detection accuracy for the two expected targets (tanks and soldiers) at a level that was substantially greater than the accuracy for the unexpected targets (nuclear devices). The analysis on the accuracy data also indicated that the presence of cueing marginally reduced the accuracy of detecting the unexpected target (compare the two bottom lines on the curve); that is, detection of the unexpected target was more likely when it was paired with the uncued soldier (72%) than with the cued tank (53%) [F(1, 12) = 15.51, p = 0.002]. The data revealed no overall effect of display, F(1, 12) = 0.78, p = 0.40. However, analyses on the soldier trials showed an interaction which was due to better detection of nuclear devices in the world-referenced condition than the screen-referenced condition, F(1, 12) = 4.61, p = 0.05. Despite the fact that nuclear devices on the tank trials (open triangles) showed the same trend as the nuclear devices on the soldier trials, the effect here was not significant [F(3, 36) = 1.12, p = 0.35], presumably because of the greater variance in this within subject comparison.

7.2 Effects of Cueing

In order to determine the effect of cueing, unconfounded by stimulus type, a comparison of the detection of cued versus uncued targets (land mines) was conducted. The data were analyzed using a 2 (display) x 2 (subject population) between subjects x 2 (cueing: cued vs. uncued) x 2 (eye) x 3 (wall: left, center, and right) within subjects ANOVA. Figure 7.2.1 shows the results for the target detection task.

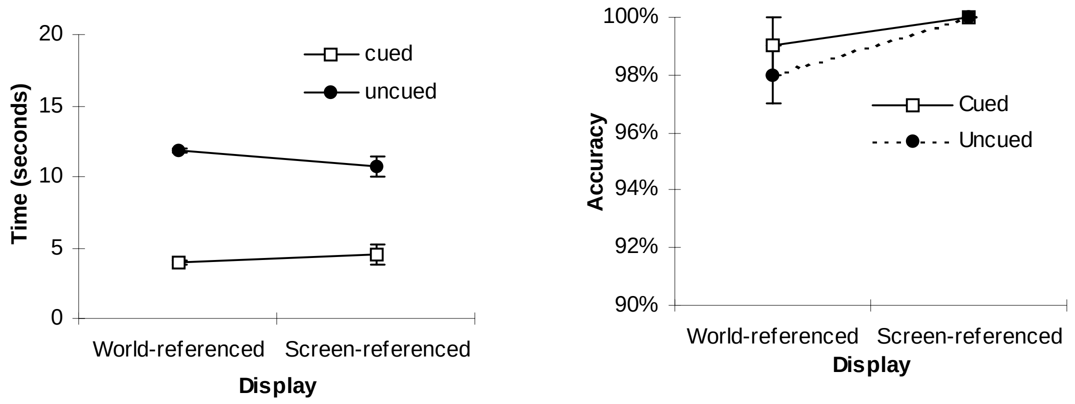


Figure 7.2.1. Effects of cueing: Land mine detection.

Data regarding mine detection showed a large benefit for target cueing, F(1, 12) = 194.27, p = .0001, replicating the cueing response time advantage for tanks seen in Figure 7.1.1. There was no effect of display, F(1, 12) = .15, p = .70, nor was the interaction between target cueing and display significant, F(1, 12) = 2.76, p = .12. Additionally, the data revealed no differences due to subject population, F(1, 12) = .93, p = .35, nor was there an interaction between cueing and subject population, F(1, 12) = .25, p = .63.

Analysis conducted on the accuracy data revealed no differences due to cueing, F(1, 12) = 2.13, p = .17, display, F(1, 12) = .53, p = .48, nor subject population, F(1, 12) = .13, p = .72, nor were there reliable interactions between these variables (display x cueing: F(1, 12) = .53, p = .48; display x subject population, F(1, 12) = .008, p = .29; cueing x subject population: F(1, 12) = .001, p = .72).

7.3 Divided Attention: Results of Secondary Task Performance

In order to determine how well subjects were able to divide their attention between information presented in the display and information in the far domain, ANOVAs were conducted on the response time and accuracy data for the secondary task. A 2 (display) x 2 (subject population) x 2 (eye) x 2 (cueing) ANOVA was conducted on the data for secondary task performance. Because the data for the secondary task was recorded continuously across the block of twenty search trials, it was not possible to examine its effects as a function of target type. The latency and accuracy with which subjects responded to the secondary task are presented in Figure 7.3.1.

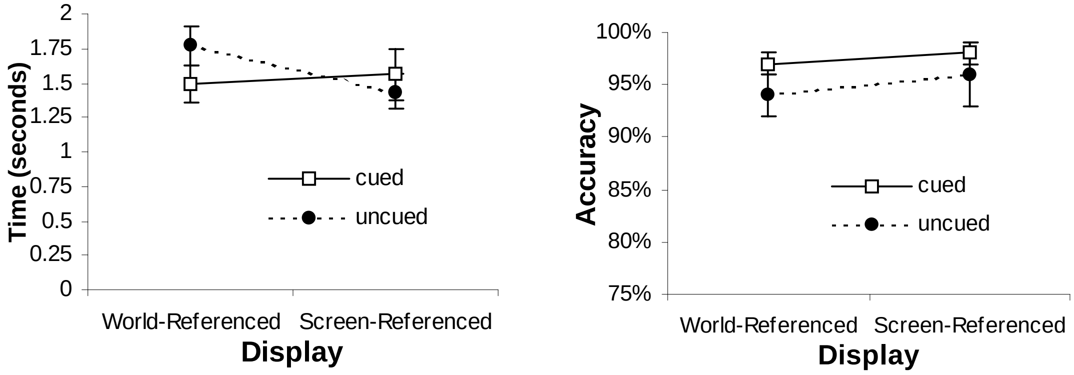


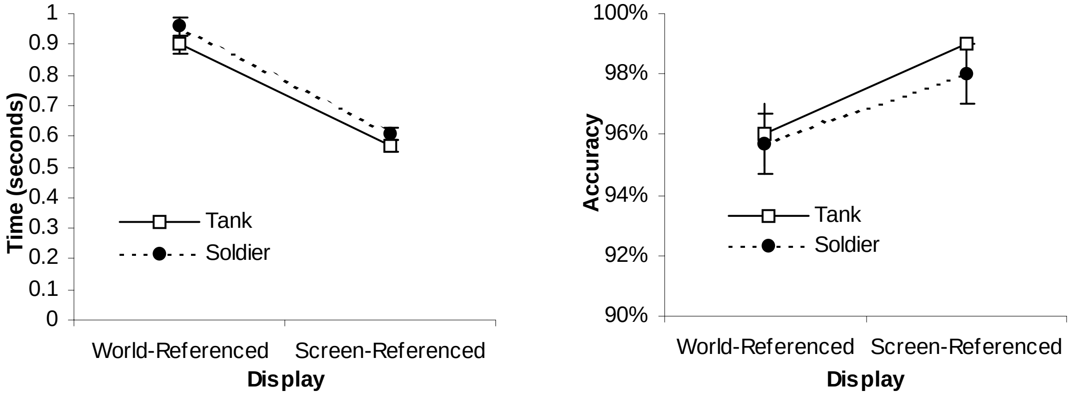
Figure 7.3.1. Response time and accuracy for the secondary task.

The response time data for the secondary task showed no effect of display, F(1, 12) = .46, p = .51, viewing condition, F(1, 12) = .63, p = .44, or subject population, F(1, 12) = .04, p = .84. The interaction between display and cueing was not significant, F(1, 12) = .42, p = .53. The accuracy data also revealed no effect of display, F(1, 11) = .33, p = .58, viewing condition, F(1, 12) = 0.00, p = .97, or subject population, F(1,12) = 1.23, p = .29. There was no interaction between display and cueing, F(1, 11) = .12, p = .73. Thus in general, secondary task performance was uninfluenced by any of the experimental variables.

7.4 Display effects for target identification and heading tasks

The results presented so far have described the data analysis of the effects of cueing and expectancy on allocating attention between the near and far domains. Subjects were also asked to perform two far domain tasks in addition to target detection – target identification and target heading – in order to determine whether the use of conformal or non-conformal imagery could facilitate performance once the target had been detected. This was essentially a single task response, since the secondary task was inactive during this phase. Response time for the target identification task was measured by the time delay between detection and identification. As noted, identification was not required for the mines or nuclear devices, since these were always assumed to be hostile. A 2 (display: world-referenced vs. screen-referenced) x 2 (subject population: military vs. civilian) between subjects x 2 (viewing condition: one eye vs. two) x 2 (target type: tank, soldier) within subjects ANOVA was conducted for the target identification task. Figure 7.4.1 shows the results for the identification task for the tanks and soldiers.

Figure 7.4.1. Response time and accuracy for the identification task.



As Figure 7.4.1 shows, a marginal effect of display was present, F(1, 12) = 3.67, p = .08, such that friend – or foe – identification was faster with the screen-referenced display than the world-referenced display. However, a marginally significant interaction between display and subject population revealed that this effect was only present for the military, F(1, 12) = 3.31, p = .10. A 0.6 second advantage was shown for the screen-referenced display versus the world-referenced display, t(46) = 5.51, p = .0000. Civilian performance on the identification task was essentially no different between the two displays.

Analysis on the accuracy data also revealed a main effect of display, F(1, 12) = .4.64, p = .05. As Figure 7.4.1 shows, subjects were more accurate in their identifications with the screen-referenced than the world-referenced display. Both civilian and military subjects showed the same trend.

Data for the target heading task was analyzed using a 2 (display) x 2 (subject population) between subjects x 2 (viewing condition) x 4 (target type: tank, soldier, mine, and nuclear device) within subjects ANOVA. When determining the accuracy for the target heading task, errors in heading greater than ±10° were considered incorrect. Figure 7.4.2 shows the results.

Figure 7.4.2. Response time and accuracy for the target heading task.

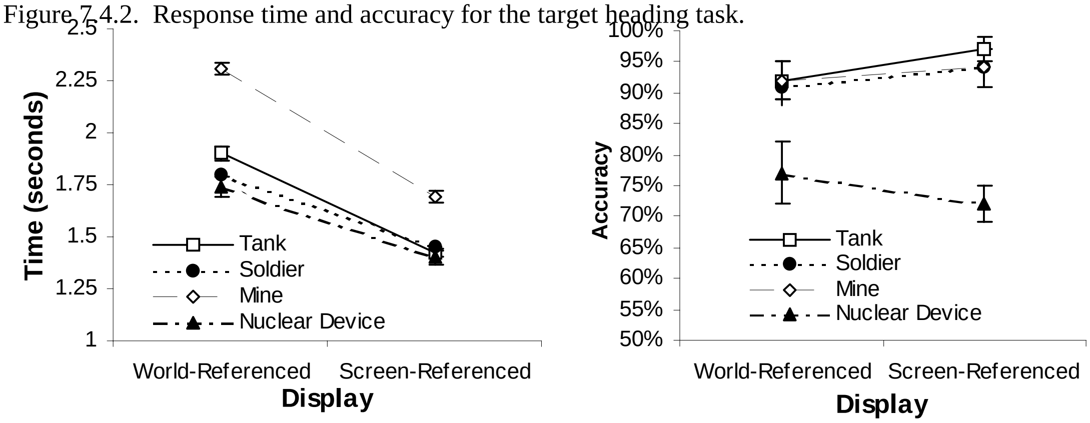


Figure 7.4.2 shows a main effect of display on response times such that heading information was given faster by subjects using a screen-referenced display, F(1, 12) = 6.79, p = .03. Comparisons of display within a target type showed that the screen-referenced display was more time effective than the world-referenced display in providing target heading information for tanks, F(1, 12) = 7.09, p = .02, soldiers, F(1, 12) = 3.49, p = .09, and land mines, F(1, 12) = 11.02, p = .006.

The influence of display on target heading accuracy is also depicted in Figure 7.4.2. The analysis showed no main effect of display, F(1, 12) = .03, p = .87. A significant interaction between target type and display was present, F(3, 36) = 5.39, p = .02, such that heading accuracy for the three expected targets (tanks, soldiers, and land mines) was higher with the screen-referenced display than the world-referenced display, but the opposite was true for the unexpected target (nuclear device).

7.5 Global Positioning Task

A 2 (display) x 2 (subject population) x 2 (viewing condition) x 5 (terrain) ANOVA was conducted on the accuracy for the global positioning task, as presented in Figure 7.5.1.

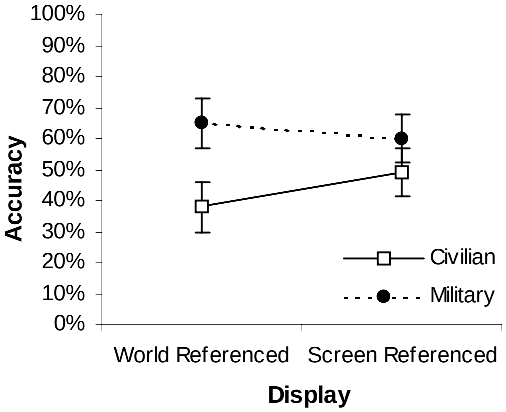


Figure 7.5.1. Accuracy for the global positioning task.

The analysis showed no overall effect of display, F(1, 141) = .17, p = .68, but significant differences in performance due to subject population, F(1, 141) = .6.41, p = .01, such that military subjects were more accurate in their responses than the civilian subjects. The interaction between display and subject population was not significant, F(1, 141) = 1.14, p = .29.

Subjects were also asked to give a confidence rating (1 = not confident, 5 = confident) as to the certainty of their answer. The accuracy data was then converted into performance scores based on subject’s confidence ratings according to the scale presented in Table 7.5.1. Scores decreased with *lower* confidence when the accuracy was correct, and decreased with *higher* confidence when accuracy was incorrect.

Correct	
Confidence	Score
5	10
4	9

Incorrect	
Confidence	Score
1	5
2	4

3	8
2	7
1	6

3	3
4	2
5	1

Table 7.5.1. Confidence score based on response accuracy .

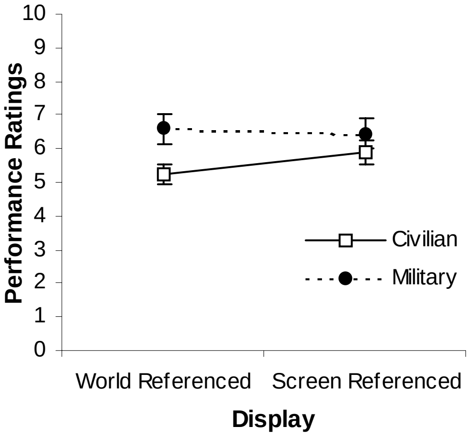


Figure 7.5.2 shows the performance ratings for the global positioning task.

Figure 7.5.2. Performance ratings for the global positioning task.

The data showed no main effect of display, F(1, 141) = .48, p = .49 but again a significant effect due to subject population, F(1, 141) = 6.08, p = .01. The interaction between display and subject population was not significant, F(1, 141) = 1.05, p = .31.

7.6 Viewing Condition

A comparison of one eyed vs. two eyed viewing was conducted on the data for tanks and soldiers using a 2 (display) x 2 (subject population) x 2 (viewing condition) x 3 (wall) ANOVA. A similar analysis was conducted for the land mines but included the added dimension of cueing.

The analysis showed that tanks were detected faster with two eyes than one, F(1, 12)= 7.07, p = .02, but there was no difference in response time for target identification, F(1, 12) = .88, p = .37 or target heading, F(1, 12) = .21, p = .65. Analysis on the accuracy data showed no difference in performance due to viewing condition for the three tasks: detection, F(1, 12) = 1.41, p = .26; identification, F(1, 12) = .08, p = .79; heading, F(1, 12) = 1.86, p = .20.

Analysis on the data for the soldiers showed no difference due to viewing condition in response time for the detection, F(1, 12)= .15, p = .70, identification, F(1, 12) = .01, p = .91, or heading, F(1, 12) = .83, p = .38, tasks. However, a significant interaction between viewing condition and subject population was present for the time required for target heading, F(1, 12) = 6.86, p = .02, such that civilians obtained heading information 0.3 seconds faster with two eyes than one but the reverse was true for the Army subjects, who showed a .07 second advantage for one eyed viewing. The accuracy data showed no difference in the accuracy of target detection, F(1, 12) = 1.83, p = .21, identification, F(1, 12) = 0.00, p = .96, or heading, F(1, 12) = .60, p = .45 , due to viewing condition.

The data for land mines showed no advantage for two eyed viewing over one eyed viewing on either the target detection task, F(1, 12) = 1.49, p = .25 or the target heading task, F(1, 12) = .30, p = .60. However, a significant interaction between viewing condition and cueing was present for target heading, F(1, 12) = 19.57, p = .0008, as shown in Figure 7.6.3.

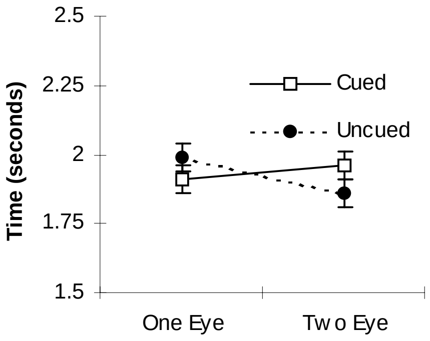


Figure 7.6.3. Effect of cueing on viewing condition for target heading task.

As Figure 7.6.3 shows, heading information for cued targets was obtained faster when information, and particularly the heading tape, was presented to one eye rather than two; however, heading information for uncued targets was faster when information was presented to both eyes rather than only one.

No effect due to viewing condition on accuracy for either the target detection, F(1, 12) = .67, p = .53, or target heading, F(1, 12) = .56, p = .47, tasks was present.

The analysis of data for the global positioning recognition task revealed a statistically marginal effect of viewing condition on subjects' ability to remember the position of targets within the terrain, F(1, 63) = 3.23, p = .08, such that subjects completed the task slightly faster when they had viewed the terrain with both eyes rather than only one. There was no effect of viewing condition on accuracy of completing the task, F(1, 141) = .56, p = .46.

It is important to note that in none of the analyses presented in section 7.6 did the effect of viewing condition interact significantly with the referencing of the display. In other words, the decision on whether information should be displayed to one eye or two can be made independently of display design, i.e. display referencing of the symbology.

7.7 Scanning Strategies

The data for detected expected targets was further examined in order to provide some insight into subject's scanning strategies. Analysis was conducted in order to determine whether the location – or wall – on which the target item was presented played a role in subjects' ability to detect the target. A 2 (display) x 2 (subject population) x 2 (viewing condition) x 4 (target type: tank, soldier, cued mines, and uncued mines) x 3 (wall) repeated measures ANOVA was conducted on the response times and accuracy for target detection. The means are presented in Figure 7.7.1.

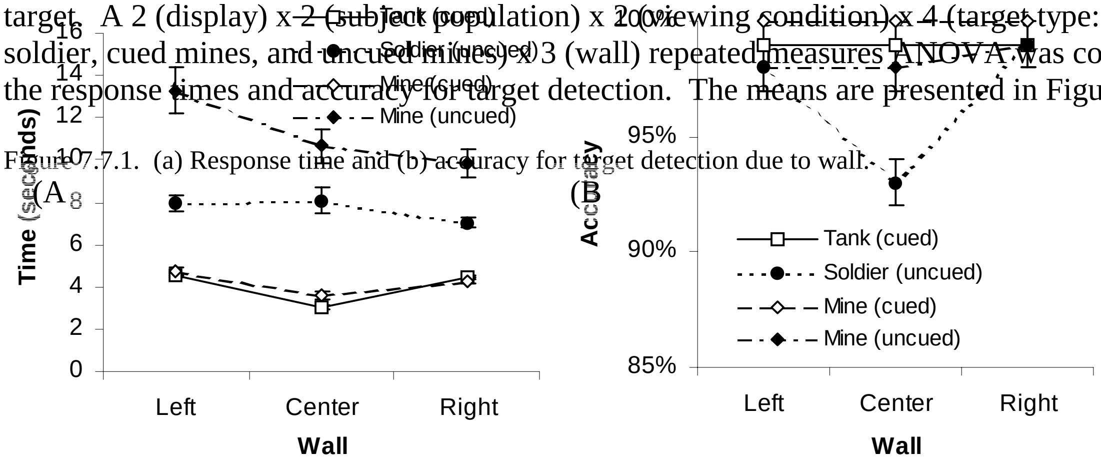


Figure 7.7.1. (a) Response time and (b) accuracy for target detection due to wall.

As Figure 7.7.1 shows, there was an overall effect due to the wall on which the targets were displayed, F(2, 24) = 9.42, p = .001. Targets were detected faster when they were presented on the center or right wall than when they were presented on the left wall (center vs. right: F(1, 12) = .10, p = .76; center vs. left: F(1, 12) = 12.78, p = .004; right vs. left: F(1, 12) = 12.40, p = .004).

The analysis of the accuracy data also revealed a main effect due to wall, F(2, 24) = 5.57, p = .01. However, as the figure shows, interpretation of this main effect depends upon the wall x target type interaction, F(6, 72) = 6.00, p = .001, which reveals that a center wall deficit was only shown for detection of the soldiers (uncued).

Analysis was conducted in order to determine the number of times a target appeared in the 40° field of view (relative to the center axis of the shutter glasses) before the target was detected. This particular angle was chosen because it corresponded to the size of the HMD imagery field around which the cueing arrow (if present) would change into a reticle to indicate the presence of a target within the subject's field of view. A 2 (display) x 2 (subject population) x 2 (viewing condition) x 4 (target type: tank, soldier, cued mines, and uncued mines) x 3 (wall) repeated measures ANOVA was conducted on the data shown in Figure 7.7.2.

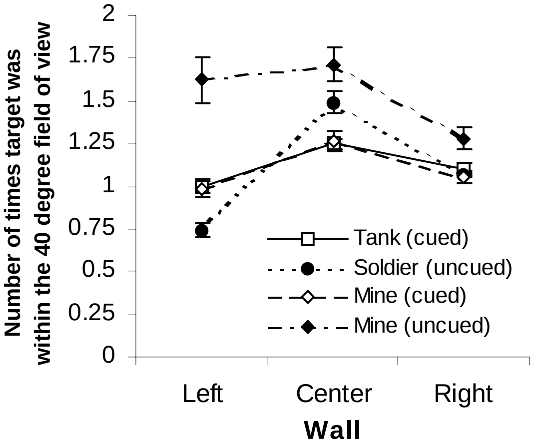


Figure 7.7.2. Number of times a target appeared in the 40° field of view.

The analysis revealed a main effect of wall, F(2, 24) = 22.51, p = .0001. The number of times a target was in the field of view before being detected was lower for targets located on the left and right walls than on the center wall [left vs. center: F91, 12) = 30.77, p = .0001; right vs. center: F(1, 12) = 27.55, p = .0002]. There was no difference between the left and right walls, F(1, 12) = .81, p = .39. The results also showed a significant interaction between wall and target type, F(6, 72) = 8.58, p = .001, such that the effect of wall was reduced for cued targets.

8. Discussion

The current experiment was conducted to determine whether manipulations of helmet mounted display design and viewing condition could aid tasks of focused attention in the near and far domains as well as divided attention between the two. The data suggest that subjects' expectancies of the targets and the presentation of cueing information aided visual search for expected targets in the simulated world but captured attention in a way that resulted in a cost for the detection of unexpected targets in the far domain. The search task was a difficult one in the sense that subjects were searching for multiple targets whose identity was unknown. Their only clue available to reduce uncertainty was that if a cueing arrow was present at the start of a trial, then they needed to search for a tank, land mine, or nuclear device. Otherwise, targets in the far domain could be a soldier, land mine, or nuclear device. The greater the potential for multiple targets in the scene, the more templates the subject needed to activate in order to complete his task. The results will be interpreted in a model of information processing, described in section 8.2.

8.1 Integration of Results

The primary purpose of the current experiment was to examine the effectiveness of a simulated HMD from an attentional perspective. Such a perspective was taken in order to highlight the importance of task analyses that clearly distinguishes display benefits on tasks requiring the focus of attention on the near domain (the secondary task represented on the instrumentation) or the far domain (search for uncued targets) or the division of attention between the domains (e.g., using instrument cueing to help far domain detection, or using instrument heading tape, to assist in far domain target azimuth judgment). The perspective is also useful because previous research with the HUD, a design concept with many features in common with the HMD, has also revealed that HUD benefits are modulated by the nature of the task (focused on a domain, divided between domains; Wickens and Long, 1995; Wickens, 1997a). In particular, such research revealed the important role of both target expectancy, and image referencing (world: conformal vs. screen: non-conformal) in moderating HUD benefits. In the current experiment, both of these variables were manipulated, along with two others; an automation-based target cueing device, and the presentation of the image to one versus both eyes. In the following discussion, we consider the effects of these four display and task variables.

The most prominent finding from this study related to the benefits and cost of target cueing. Cued targets were clearly assisted in their detection, a benefit to the divided attention task, and this benefit was realized no matter whether the cueing was world referenced (enabling a reticle to be placed over the target), or screen referenced (the reticle indicated that the target was within the field of view). The benefit was observed for the always cued tanks over the never cued soldiers, although in making this comparison, we never explicitly compared the detectability of these two targets in uncued format. However, the benefit was also observed for the cued over the uncued mines. In the current study, the cueing was 100% reliable, never cueing a "false target". Hence we had no direct way of examining the extent to which automation based cueing or highlighting could lead the soldier down the "garden path" of following the cueing even when it was in error (Conejo and Wickens, 1997). However we did observe an indirect manifestation of this phenomenon, reflected in a cost to cueing for simultaneously viewable uncued targets...the high priority nuclear device. The data in figure 7.1.1 clearly indicate that such a high priority target was more likely to be overlooked if it appeared on the same trial as a cued target, than as an uncued one. Hence, while divided attention between the near (cueing information) and far (cued target) domain was assisted by cueing, focused attention on the far domain was disrupted. Such a disruption could have serious real world implications, and is reminiscent of the similar disruption of the detection of unexpected events and targets caused by a head up display (Wickens and Long, 1995).

The data indicated however that the costs of unexpected event detection were reduced somewhat by the third variable manipulated in the experiment, the "world referencing" of the display, as shown in the accuracy data of Figure 7.1.1. This effect parallels to some extent the findings of Wickens and Long (1995) that conformal (i.e. world referenced) HUD imagery eliminated the HUD-based cost to unexpected event detection. As implemented in the current experiment, world referenced symbology, driven by real time image updating based on head movement, had two somewhat different manifestations. (1) On the one hand, by superimposing imagery with its far domain counterpart (e.g., the horizon line) when the head is in motion, it can create a sense of "fusion" between the near and far domain, that is continuously evident, hence possibly "scene linking" the two domains (Foyle, et al., 1996) in a way that would benefit attention to both. (2) On the other hand, when cueing was present, the world referenced display provides more precise location of the cued target by overlaying the reticle on that target, rather than merely indicating that the target is in the field of view (cueing in the screen referenced display).

Careful consideration of the data reveals that the benefits of world referencing are more likely related to the first of these explanations – an attentional fusing of the two domains – than the second. First, we note that the slight improvement in the detection of the uncued nuclear device in the world referenced display must be attributed to the first, and not the second explanation, since this improvement was observed whether or not cueing was present, and since cueing was on the whole found to disrupt detection of uncued items. Second, the current data provided little evidence that world referenced cueing offered any benefits above and beyond those of screen referenced cueing for detection of the cued targets.

Although world referencing did provide the important benefit to detection of unexpected targets, it also imposed two noteworthy costs to performance. As Figure 7.4.1 reveals, it imposed a cost on classifying the target as friend or foe (i.e. discriminating left from right facing tanks and soldiers); and as shown in figure 7.4.2, world referencing delayed the reporting of target azimuth. The second of these effects can be readily explained by the fact that world referencing sometimes rendered the azimuth scale off the field of view of the HMD, if the head was oriented downward, hence requiring a short time to "look up" and bring the scale back into the field of view. This result is similar to that reported by Andre and Cashion (1993). The cause of the first effect remains somewhat obscure.

In discussing the attentional effects of the manipulations, it is important to note that none of the variables imposed significant changes on performance of the secondary, near-domain monitoring task. On the one hand, this lack of effect makes interpretation of the results somewhat simpler, since tradeoffs between tasks and task domains do not need to be considered (Fadden and Wickens, 1997). On the other hand, it may possible represent a "ceiling effect" of a very easy secondary task.

Finally, we found that the fourth variable addressed by this study, the one vs. two eyed viewing condition, caused only muted effects on performance. We had anticipated possible costs associated with one eyed viewing of the imagery related to binocular rivalry, and a lower intensity the single image (compared with the additive intensity of the two). Of the few results that did show an effect of viewing condition, the greater proportion did favor the two eyed views, supporting the detection of cued tanks and the azimuth reporting of the uncued soldiers (both divided attention tasks).

The results for the global post-task recognition test indicate that subjects' mental representation of the location of objects in the environment was independent of display manipulations of symbology, i.e. world referencing or screen referencing. In fact, the only effect on performance for the global positioning task were attributable to individual differences – in this case, military training. The means for accuracy and performance ratings, shown in Figures 7.5.1 and 7.5.2, indicated that Army subjects were both more accurate and more confident in their responses regarding the location of objects within the environment. From subjective comments gathered during the task, civilians approached this task by attempting to memorize the position of certain objects in the environment, e.g. a tank in the far left edge of the left wall. However, Army subjects tended to remember object's position based on terrain features, e.g. a tank on a hill. It should be noted though that a corresponding absence of influence of our variables on the global reporting task may reflect either the absence of a true effect, or the insensitivity of this particular measure, to reflect differences in spatial re-creation of the environment. It should also be noted that performance on this task was not terribly good, averaging around 50% accuracy (with chance performance level of 25%).

The paradigm that we used to evaluate the simulated HMD appears to offer promise for further use. The tasks appeared challenging, as witnessed by the substantial time delays involved in target detection, ranging from 5 to 15 seconds. Furthermore, mimicking the search for targets in the real world (Wickens, 1992), it is noteworthy that on a number of occasions, the field of view passed over the target, without its being detected. The soldiers who participated in the simulation performed better on the global post-task recognition test than did their civilian counterparts. At the same time it is important to note that for the vast majority of the results (and all that were reported above), soldiers and civilians both showed the same trends of performance across the manipulated variables.

8.2 Model of information processing

Figure 8.2.1 presents a framework for a model of visual search within which the data can be interpreted.

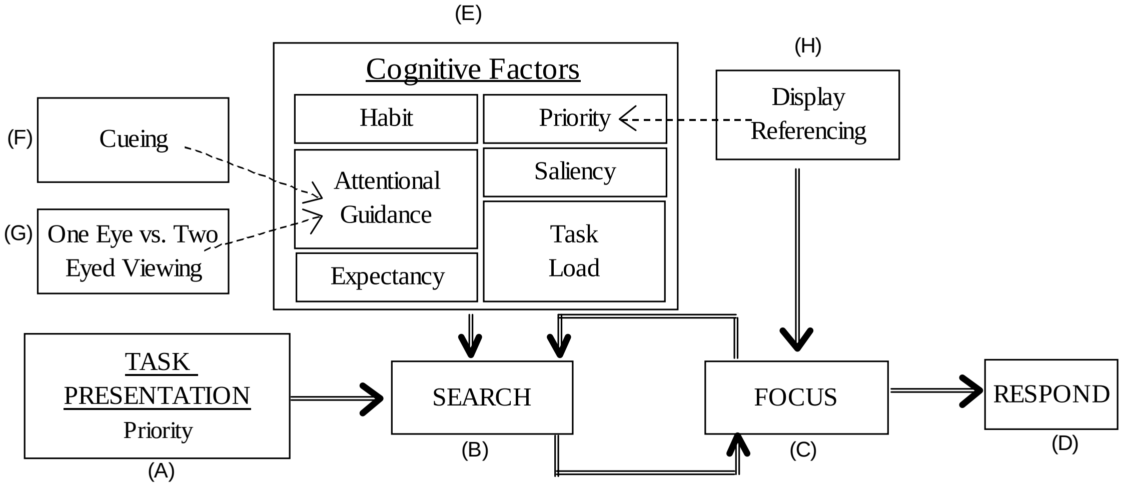


Figure 8.2.1. Proposed model of visual search.

Figure 8.2.1 shows four main stages in the visual search model (boxes [A]-[D]): task presentation [A], search through the environment [B], focus on specific areas of the environment or on the target itself [C], and finally the task response [D]. Boxes [E]-[H] detail cognitive and design factors, which can aid or hinder performance at various stages of the model. The solid arrows describe those stages and factors that have a *direct* effect on the visual search process, the dashed arrows show those factors that *indirectly* influence the process.

The visual search model begins by assuming that subjects are given a task [A] or a set of tasks; in the case of the latter, subjects need to *prioritize* the tasks to be performed. Once subjects have a task in mind, the search stage [B] commences. Processing at this stage is influenced by various *cognitive factors* [E]. Of these cognitive factors, attentional guidance may be affected by *cueing* [F] and *viewing condition* [G] (one eye vs. two eyed viewing), and priority may be influenced by *display referencing* [H] (i.e. the use of world- or screen-referencing). Once search has identified a specific area of the environment which is likely to contain a target, the search process is modulated in the focus [C] stage by the use of a world-referenced or screen-referenced *display* [H]. Note that the search and focus stages of the model may interact as the subject looks for the target, as shown by the feedback loop. Once the target is found, subjects can respond [D] to the various tasks they were asked to perform.

The operations shown in Figure 8.2.1 take time. To improve the efficiency of the search process, displays can be designed to capitalize on known mechanisms to aid visual search, e.g. cueing. The discussion of the model will now turn to examine how different factors examined in the experiment influence processing at the various stages of the model.

Task. In the current study, subjects were asked to perform tasks consisting of target detection, identification, and location. Performance on the tasks can be hindered by the addition of a secondary task, at which point, the subject must prioritize the tasks to be performed. Subjects in the experiment were instructed that the target detection task was more important than the secondary monitoring task.

Search Process. The <u>search</u> stage is guided by the following six factors (shown in [E] of Figure 8.2.1): *habit* in terms of where to look for information, *attentional guidance* which directs the subject to the target's location, *expectancy* based on prior knowledge of where to look as well as what objects to look for, *priority* of the targets to be searched as well as where to search, *saliency* of the target, and *task load*.

The search process begins as the subject divides his attention between the primary target detection task and the monitoring secondary task. The subject is unaware of what the actual target is – rather search commences for one of three targets: if a cueing arrow is present at the start of a trial, the subject searches for a tank, land mine, or nuclear device; otherwise, the subject searches for a soldier, land mine, or nuclear device. Subjects were instructed to *prioritize* their search, such that detecting the nuclear device was given priority over the detection of the other three targets.

What the subjects first search for in the environment is guided by *expectancy*, determined by the frequency of a target's appearance. As shown in Figure 7.1.1, subjects detected expected targets with near-perfect accuracy (98%) but detected unexpected targets only 67% of the time. Where the subject initiates his search in the environment is directed by two factors: *expectancy* of the target's location and *habit*. As an example of the former, military subjects were guided by strategies taught during training. Priority of search was given to features in the environment that had a greater likelihood of containing a target. For example, areas in front of hills were searched before valleys since the height of the hill hid the silhouette of the object whereas being caught in a valley (or lower ground) is normally a military disadvantage in battle. However, since the target objects in the current scenario were not placed in the environment according to military battlefield scenarios, this approach was not the most optimal, and in fact, civilians were faster to detect cued targets than military subjects.

As an example of habit, the data describing scanning strategies show that subjects moved their head clockwise, similar to the pattern used in reading text from left to right. Detection data uncued targets showed faster target detection times on the right and center walls than on the left wall (Figure 7.7.1). When these data are taken into consideration with those showing that subjects were more likely to miss the target on the center wall than on the right wall when it appeared within their field of view (Figure 7.7.2), the combined results suggest that subjects sometimes turned their head from the center wall, where the head was positioned at the start of each trial, to the right immediately after the trial began. That is, when the target was within the subject's field of view on the center wall, they occasionally did not notice the target nor attempt to search the center wall but instead turned their head clockwise, preferring to search the right wall first. If the target was not found on the right wall, subjects then moved their head back to the center and searched that wall for objects before examining the left wall.

Habit no longer directs search when cueing symbology is present to guide the subject's attention to a specific area of the display. The effect of target azimuth for the detection of land mines, i.e. the wall on which the stimulus was presented, was eliminated when the target was cued, suggesting that the valid cueing used here creates an attentional *expectancy* for the target in the cued region. This hypothesis accounts for the target detection data shown in Figure 7.1.1. When the valid cueing was present, unexpected targets were detected faster than expected targets (for those trials in which the subject did in fact detect the unexpected target). The advantage resulted, in part because the locations of the unexpected targets (nuclear devices) were selected by the experimenter to maximize the likelihood that the unexpected target would appear in the subject's field of view as the subject searched for the target (e.g. if the target was a tank located in the center of the left wall, the nuclear device would have been positioned to the right of the tank so that the subject's field of view would pass over the unexpected target as he moved his head from the center wall to the left wall). Thus once attention was cued to the general region of the display during the search phase, the *salience* of the nuclear device resulting from differences in size and shape may have aided subjects in finding the unexpected target – that is, when the nuclear device was detected.

However, the presence of cueing reduced the accuracy for detecting unexpected targets. In other words, the use of cueing proved to be a more dominating factor in guiding search than was target *priority*. Subjects were less likely to detect the unexpected target (the nuclear device) when it was presented with a cued stimulus (tank) than when it was presented with an uncued stimulus (the soldier). That is, the presence of cueing raised the subjects' expectancy of where the target was located, which prevented subjects from more carefully examining the environment for a higher priority target.

The dashed arrow in Figure 8.2.1 shows that viewing condition impacts the benefits of attentional guidance. The lower intensity of the image when the symbology was presented to one eye rather than two was a cost in the performance on the detection of cued tanks and the reporting of target heading for uncued soldiers.

By fusing information between near and far domains through *display referencing, priority* once again becomes an important factor in directing search. Unexpected events were detected more with the world-referenced than with the screen-referenced display whether cueing was present or not. Although we hypothesized that *display referencing* could mediate some of the effects due to *attentional guidance*, the relationship between the two is not represented in the information processing model since data showed that when a display advantage was present (e.g. unexpected targets were detected more accurately with a world-referenced rather than a screen-referenced display), this advantage existed whether cueing was available or not.

Task load was reduced once subjects detected the target, as the secondary task was temporarily stopped for the identification and location tasks.

Focus Stage. Once the subject has narrowed the search field to a subset of the domain, the focus stage commences. Figure 8.2.1 shows that processing at this stage of the model can be aided by display – that is, the use of a world-referenced or screen-referenced display. Evidence to support this hypothesis was the benefit for the world-referenced display in detecting higher priority – but low frequency – nuclear targets over that with the screen-referenced display. The conformal display guided attention outwards to the far domain so that nuclear devices were detected 17% more frequently than with the use of a non-conformal display. This relationship between display design and the ability to focus on information in the far domain was also given support by the results of Fadden and Wickens (1997) and Wickens and Long (1995). In the latter experiment, scene linking symbology with the far domain (i.e. runway) drew attention more effectively to the outside world than the use of non-conformal symbology; as a result, scene-linking counteracted the negative effects of expectancy on detection of low frequency targets.

Although superimposing information from the near domain (the cueing reticle) directly over objects in the far domain was shown to be beneficial for target detection (the mines), a disadvantage for the world-referenced cued display configuration was present for the target identification and location tasks. The differences in performance due to display referencing for the identification task are as yet unexplained; since both cued and uncued targets showed the same trend, the differences in performance between the world and screen referenced displays must be attributable to a factor other than the superimposition of the lock-on reticle). For the target heading task, in the world-referenced display, the benefit of reduced information in the forward field of view by presenting heading information conformal to the horizon line (thereby permitting subjects to "look away" from the heading information) was offset by increased visual scanning time as the subject moved his head between the target and the heading tape in order to obtain target heading information.

Feedback. The feedback loop between the search and focus stages in Figure 8.2.1 indicates that the two stages interact. For example, it was possible that in performing the target detection task, subjects' search led them to focus on a terrain feature. Upon closer examination of what they assumed to be the target, subjects realized their mistake and commenced search again. This feedback loop may describe search for some of the land mines. This target type was the smallest of the four target types and could be mistaken for a terrain feature, and vice versa.

Response. Once the target is detected, the subject can make a response. The subject can then proceed to the next task (e.g. begin search for the next target).

The model presented in Figure 8.2.1 presents a framework for how data in the near and far domains were processed as described by the data collected. The current results point cautiously to the benefits of two of the three design features examined here: world referencing of imagery and biocular viewing. At the same time, the possible benefits of the third feature examined, cueing, remain more ambivalent, and clearly depend heavily both upon the reliability of the automation which imposes the cueing (here the "best case" perfect reliability), as well as the costs of failure to detect targets which cannot benefit from such cueing.

9. Conclusions

The current experiment showed that detection of a target was dependent upon the expectancy of a target as well as the frequency of the target. The more frequent targets (i.e. tanks and soldiers, present 60% of the time) were more highly expected and were detected more accurately than the nuclear devices, presented only 10% of the time. The presence of a cueing information facilitated target detection whether it was presented conformally or non-conformally in target search but hindered the detection of unexpected targets; attention was captured by the presence of the cue and drawn directly to the expected target rather than the unexpected target. However, this effect was mediated by display, such that the use of a world-referenced (or conformal) display allowed subjects to approach the target search task better using priority of targets as a factor in their search.

The results showed a slight benefit for biocular viewing over monocular viewing in the detection of cued tanks and reporting the heading of soldiers and land mines. A global recognition test administered to subjects after each block of trials showed no effect of display or viewing condition; i.e. the representation the subject formed of his environment was independent of the symbology manipulations.

Further research to be performed will examine the effects of expectancy and cueing in the context of a head-up (HMD) vs. head-down (hand-held) display. Additionally, the examination of cueing could involve manipulation of cueing validity in order to determine whether subjects are more likely to scan the environment (and detect unexpected targets) if the cueing were not 100% reliable. The results suggest potential benefits for the use of world referencing and biocular viewing in the design of HMDs but warns of the use of cueing, which aided the detection of targets which were expected but hindered the detection of targets which were not.

Acknowledgements

Many thanks to Ron Carbonari for his time and effort in the development of the program and Jacob Seagull for his help in the development of the experiment and the running of subjects. Research funding for this project was provided by a grant from the Army Research Lab.

References

Andre, A.D. and Cashion, P.A. (1993). Visual scanning in the functional visual field. *Society for Information Display International Symposium Digest of Technical Papers* (pp. 114-117). Playa Del Rey, CA: SID.

Andre, A.D. and Johnson, W.W. (1992). Stereo effectiveness evaluation for precision hover tasks in a helmet-mounted display simulator. *1992 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 1136-1140). IEEE.

Arditi, A. (1986). Binocular vision. In K.R. Boff, L. Kaufman, and J.P. Thomas (Eds.), *Handbook of Perception and Human Performance*. Vol 1, New York: John Wiley.

Arthur, E.J., Hancock, P.A., and Chrysler, S.T. (1993). Spatial orientation in real and virtual worlds. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (pp. 328-332). Santa Monica, CA: Human Factors and Ergonomics Society.

Arthur, E., Hancock, P., and Telke, S. (1996). Navigation in virtual environments. In *Proceedings from the 10th Annual international AeroSense Symposium*. Bellingham, WA: SPIE.

Arthur, K.W., Booth, K.S., and Ware, C. (1993). Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems*. 11(3), 239-265.

Atchley, P., Kramer, A.F., and Theeuwes, J. (1997). Object-based attention in depth. Presented at the 38th Annual meeting of the Psychonomic Society, Philadephia, PA.

Bailey, R.E. (1994). HUD lessons-learned for HMD development. In R.J. Lewandowski, W. Stephens, and L.A. Haworth (Eds.), Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-mounted Displays and Symbology Design Requirements I (pp. 226-237). 2218. Bellingham, WA: SPIE.

Bajura, M., Fuchs, H., and Ohbuchi, R. (1992). Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *Computer Graphics*. 26 (2), 203-210.

Barfield, W., Hendrix, C., and Bystrom, K. (1997). Visualizing the structure of virtual objects using head tracked stereoscopic displays. *1997 IEEE Virtual Reality International Symposium* (pp. 114-119). Los Alamitos, CA: IEEE.

Barfield, W., Rosenberg, C., and Furness, T.A. (1995). Situation awareness as a function of frame of reference, computer graphics eyepoint elevation, and geometric field of view. *International Journal of Aviation Psychology*. 5(3), 233-256.

Barfield, W., Rosenberg, C., and Lotens, W.A. (1995). Augmented-reality displays. In W. Barfield and T.A. Furness (Eds). *Virtual Environments and Advanced Interface Design*. New York: Oxford University Press.

Becklen, R. and Cervone, D. (1984). Selective looking and the noticing of unexpected events. *Memory and Cognition*. 11(6), 601-608.

Bellenkes, A.H. (1988). Helmet-based imaging systems. *Crew Station Technology Report no. 88-1.*

Blake, R. and Boothroyd, K. (1985). The precedence of binocular fusion over binocular rivalry. *Perception and Psychophysics*. 37(2), 114-124.

Blake, R. and Fox, R. (1981). Further developments in binocular summation. *Perception and Psychophysics*. 30, 266-276.

Cole, R.E., Merritt, J.O., Coleman, R., and Ikehara, C. (1991). Teleoperator performance with virtual window display. In J.O. Merritt and S.E. Fisher (Eds.), *Proceedings of the International Society for Optical Engineers (SPIE): Stereoscopic displays and applications II* (pp. 111-119). 1457. Bellingham, WA: SPIE.

Conejo, R., and Wickens, C.D. (1997). The effects of highlighting validity and feature type on air-to-ground target acquisition performance. *University of Illinois Institute of Aviation Technical Report (ARL-97-11/NAWC-ONR-97-1).* Savoy, IL: Aviation Research Lab.

CuQlock-Knopp, V.G., Sipes, D.E., Torgerson, W., Bender, E., and Merritt, J.O. (1996). Human off-road mobility, preference, and target detection performance with monocular, biocular, and binocular night vision goggles. *Army Research Laboratory Technical Report: ARL-TR-1170*.

CuQlock-Knopp, V.G., Torgerson, W., Sipes, D.E., Bender, E., and Merritt, J.O. (1995). A comparison of monocular, biocular, and binocular night vision goggles for traversing off-road terrain on foot. *Army Research Laboratory Technical Report: ARL-TR-747*.

Davis, E. T. (1997). Visual requirements in HMDs: What can we see and what do we need to see? In J.E. Melzer and K. Moffitt (Eds.), *Head Mounted Displays: Designing for the User*. New York, NY: McGraw-Hill.

Davis, E. T. and Hodges, L.F. (1995) Human stereopsis, fusion, and stereoscopic virtual environments. In W. Barfield and T.A. Furness (Eds). *Virtual Environments and Advanced Interface Design* (pp. 145-174). New York: Oxford University Press.

Dede, C., Salzman, M.C., and Loftin, R.B. (1996). ScienceSpace: Virtual realities for learning complex and abstract scientific concepts. In *Proceedings of the Virtual Reality Annual International Symposium* (pp. 246-252). Los Alamitos, CA: IEEE.

Drasic, D. (1991). Skill acquisition and task performance in teleoperation using monoscopic and stereoscopic video remote viewing. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1367 - 1371). Santa Monica, CA: Human Factors Society.

Dudfield, H.J., Hardiman, T.D., and Selcon, S.J. (1995). Human factors issues in the design of helmet-mounted displays. In R.J. Lewandowski, W. Stephens, and L.A. Haworth (Eds.), *Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-mounted Displays and Symbology Design Requirements II* (pp. 214-225). 2218. Bellingham, WA: SPIE.

Duncan, J. (1984). Selective attention and the organization of vital information. *Journal of Experimental Psychology: General*. 113, 501-517.

Ellis, S.R. (1994). What are virtual environments? *IEEE Computer Graphics and Applications*. 14 (1), 17-22.

Ellis, S.R., Bréant, F., Menges, B., Jacoby, R., and Adelstein, B.D. (1997). Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: Viewing conditions and rendering latency. *1997 IEEE Virtual Reality International Symposium* (pp. 138-145). Los Alamitos, CA: IEEE.

Ellis, S.R. and Bucher, U.J. (1992). Distance perception of stereoscopically presented virtual objects optically superimposed on physical objects by a head-mounted see-through display. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Ellis, S.R. and Menges, B.M. (1995). Judged distance to virtual objects in the near visual field. *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (pp. 1400-1405). Santa Monica, CA: Human Factors and Ergonomics Society.

Ellis, S.R., Tharp, G.K., Grunwald, A.J., and Smith, S. (1991). Exocentric judgments in real environments and stereoscopic displays. In *Proceedings of the 35th Annual Meeting of the Human Factors Society* (pp. 1442-1446). Santa Monica, CA: Human Factors Society.

Eggleston, R.G., Janson, W.P., and Aldrich, K.A. (1996). Virtual reality system effects on size-distance judgments in a virtual environment. In *Proceedings of the Virtual Reality Annual International Symposium* (pp. 139-146). Los Alamitos, CA: IEEE.

Eriksen, B.A. and Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Perception and Psychophysics*. 16, 143-149.

Fadden, S. and Wickens, C.D. (1997). Maximizing traffic awareness with a head-up flightpath highway display. *University of Illinois Institute of Aviation Technical Report (ARL-97-1/FAA-97-1).* Savoy, IL: Aviation Research Laboratory.

Feiner, S., MacIntyre, B., and Seligmann, D. (1993). Knowledge-based augmented reality. *Communications of the ACM*. 36(7), 53-62.

Fischer, E., Haines, R.F., and Price, T.A. (1980). Cognitive issues in head-up displays (NASA Technical Paper 1711). Moffet Field, CA: NASA Ames Research Center.

Fisher, S. (1982). Viewpoint dependent imaging: An interactive stereoscopic display. *SPIE: Processing and Display of Three Dimensional Data*. 367, 41-45.

Foyle, D.C., Andre, A.D., McCann, R.S., Wenzel, E.M., Begault, D.R., and Battiste, V. (1996). Taxiway navigation and situation awareness (T-NASA) system: Problem, design philosophy, and description of an integrated display suite for low-visibility airport surface operations. *Proceedings of the SAE/AIAA World Aviation Congress, Paper 965551*. Washington, DC: American Institute of Aeronautics and Astronautics.

Foyle, D.C., McCann, R.S., and Shelden, S.G. (1995). Attentional issues with superimposed symbology: formats for scene-linked displays. *Proceedings of the Eighth International Symposium on Aviation Psychology*. R.S. Jensen (Ed.). Columbus, OH: Ohio State University.

Foyle, D.C., Sanford, B.D., and McCann, R.S. (1991). Attentional issues in superimposed flight symbology. *Proceedings of the Sixth International Symposium on Aviation Psychology*. R.S. Jensen (Ed.). Columbus, OH: Ohio State University.

Geiselman, E.E. and Osgood, R.K. (1995). Head vs. aircraft oriented air-to-air target location symbology using a helmet-mounted display. In R.J. Lewandowski, W. Stephens, and L.A. Haworth (Eds.), *Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-mounted Displays and Symbology Design Requirements II* (pp. 214-225). 2218. Bellingham, WA: SPIE.

Geiselman, E.E. and Osgood, R.K. (1994). Utility of off-boresight helmet-mounted symbology during a high angle airborne target acquisition task. In R.J. Lewandowski, W. Stephens, and L.A. Haworth (Eds.), *Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-mounted Displays and Symbology Design Requirements I* (pp. 328-338). 2218. Bellingham, WA: SPIE.

Gopher, D., Kimchi, R., Seagull, F.J., Catz, I., and Trainin, O. (1992). Flying with dichoptic displays: The interplay between display characteristics and attention control. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 1469-1473). Santa Monica, CA: Human Factors Society.

Haworth, L.A., Sharkey, T., and Lee, A. (1995). TRISTAR III: Helmet-mounted display symbology. In R.J. Lewandowski, W. Stephens, and L.A. Haworth (Eds.), *Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-mounted Displays and Symbology Design Requirements II* (pp. 235-245). 2465. Bellingham, WA: SPIE.

Hendrix, C. and Barfield, W. (1995). Presence within virtual environments as a function of visual and auditory cues. In *Proceedings of the Virtual Reality Annual International Symposium* (pp. 74-82). Los Alamitos, CA: IEEE.

Hendrix, C., Brandt, J., and Barfield, W. (1995). Presence within virtual environments as a function of headtracking and stereopsis. *SAE International Conference Proceedings*. Salem, MA: SAE.

Henry, D. and Furness, T. (1993). Spatial perception in virtual environments: Evaluating an architectural application. *1993 IEEE Virtual Reality International Symposium* (pp. 33-40). Los Alamitos, CA: IEEE.

Hettinger, L.J., Nelson, W.T., and Haas, M.W. (1994). Comparison of target detection performance using a helmet-mounted display and a conventional flight simulator dome display. In *Proceedings of '94 Symposium on Human Interaction with Complex Systems* (pp. 299 - 306). Los Alamitos, CA: IEEE.

Hodges, L.F., Rothbaum, B.O., Watson, B., Kessler, G., and Opdyke, D. (1996). A virtual airplane for fear of flying therapy. In *Proceedings of the Virtual Reality Annual International Symposium* (pp. 86-92). Los Alamitos, CA: IEEE.

Hofer, E.F., Palen, L.A., and Possolo, A. (1993). Flight deck information management: An experimental study of functional integration of approach data. *Proceedings of the 7th International Symposium on Aviation Psychology*. Columbus, OH: Ohio State University.

Jagacinski, R.J. (1977). A qualitative look at feedback control theory as a style of describing behavior, *Human Factors*, 19, 331-347.

Kaplan, D.A. and Brown, C. (March 10, 1997). Goodbye, Pac-Man. *Newsweek*, p. 72-73.

Kim, W.S., Ellis, S.R., Tyler, M.E., Hannaford, B., and Stark, L.W. (1987). Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks. *IEEE Transactions on Systems, Man, and Cybernetics*. 17(1), 61 - 72.

Klymento, V. and Rash, C.E. (1995). Human performance with new helmet-mounted display designs. *CSERIAC Gateway*. 4(4), 1-4.

Kooi, F.L. (1993). Binocular configurations of a night-flight head-mounted display. *Displays*. 14(1), 11-20.

Kramer, A.F. and Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception and Psychophysics*, 50, 267-284.

LaBerge, D. and Brown, V. (1989). Theory of operations in shape identification. *Psychological Review*, 96, 101-124.

Lasswell, J.W. and Wickens, C.D. (1995). The effects of display location and dimensionality on taxi-way navigation. *University of Illinois Institute of Aviation Technical Report (ARL-95-5/NASA-95-2)*. Savoy, IL: Aviation Res. Lab.

Larry, C. and Elworth, C.L. (1972). The effects of pilot age, lighting, and headdown time on visual accommodation. Cited in D.J. Weintraub and M. Ensign (Eds.), *Human Factors Issues in Head-Up Display Design: The Book of HUD*. State-of-the-Art Report, Crew System Ergonomics Information Analysis Center, Wright-Patterson AFB, Dayton, OH.

Larish, I. and Wickens, C.D. (1991). Divided attention with superimposed and separated imagery: Implications for head-up displays. *University of Illinois Institute of Aviation Technical Report (ARL-91-4/NASA HUD-91-1)*. Savoy, IL: Aviation Research Laboratory.

Lippert, T.M. (1990). Fundamental monocular/binocular HMD human factors. In R.J. Lewandowski (Ed.), *Proceedings of the International Society for Optical Engineers (SPIE): Helmet-mounted Displays II* (pp. 185-191). 1290. Bellingham, WA: SPIE.

Marshak, W.P. (1997). Identifying research areas for the digitization of the battlefield. In *Proceedings Advanced Displays and Interactive Displays: Federated Laboratory First Annual Symposium* (1-1 - 1-14). Army Research Laboratory: Adelphi, Maryland.

Martens, M.H. and Wickens, C.D. (1995). Intensity coding Solutions to Display Clutter. *University of Illinois Institute of Aviation Technical Report (ARL 95-10/NASA 95-4)*. Savoy, IL: Aviation Research Laboratory.

Martin-Emerson, R. and Wickens, C.D. (1997). Superimposition, symbology, visual attention, and the head-up display. *Human Factors*. In press.

Martin-Emerson, R. and Wickens, C.D. (1992). The vertical visual field and implications for the head-up display. *Proceedings of the Human Factors Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

May, P. and Wickens, C.D. (1995). The role of visual attention in head-up displays: Design implications for varying symbology intensity. *University of Illinois Institute of Aviation Technical Report (ARL-95-3/NASA-HUD-95-1)*. Savoy, IL: Aviation Research Laboratory.

McCann, R.S., and Foyle, D.C. (1995). Scene-linked symbology to improve situation awareness. *1995 NATO - AGARD Symposium on Situation Awareness* (pp. 16-1 - 16-11). Neuilly Sur Seine, France, NATO Advisory Group for Aerospace Research and Development (NTIS No. AGARDCP 575).

McCann, R.S., Foyle, D.C., and Johnston, J.C. (1992). Attentional limitations with head-up displays. *SID International Symposium Digest of Technical Papers*, XXII.

McCann, R.S., Lynch, J., Foyle, D.C., Johnston, J.C. (1993). Modeling attentional effects with head-up displays. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

McCormick, E.P. and Wickens, C.D. (1995) Virtual reality features of frame of reference and display dimensionality with stereopsis: Their effects on scientific visualization. *University of Illinois Institute of Aviation Technical Report (ARL-95-6/PNL-95-1)*. Savoy, IL: Aviation Research Laboratory.

McGovern, D.E. (1993). Experience and results in teleoperation of land vehicles. In S.R. Ellis, M.K. Kaiser, and A.C. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (pp. 182 - 195). London: Taylor and Francis.

McKenna, M. (1992). Interactive viewpoint control and three-dimensional operations. *ACM*, 53-56.

Melzer, J.E. and Moffitt, K. (1997). HMD design – Putting the user first. In J.E. Melzer and K. Moffitt (Eds.), *Head Mounted Displays: Designing for the User*. New York, NY: McGraw-Hill.

Miller, D.P. (1988). Evaluation of vision systems for teleoperated land vehicles. *IEEE Control Systems*, 8, 37-41.

Moffitt, K. (1989). Ocular responses to monocular and binocular helmet-mounted display configurations. In R.J. Lewandowski (Ed.), *Proceedings of the International Society for Optical Engineers (SPIE): Display System Optics II* (pp. 142 - 148). 1117. Bellingham, WA: SPIE.

Mykityshyn, M.G., Kuchar, J.K, and Hansman, R.J. (1994). Experimental study of electronically based instrument approach plates. *The International Journal of Aviation Psychology*. 4(2), 141-166.

Naish, J.M. (1964). Combination of information in superimposed visual fields. *Nature*. 202 (4933), 641-646.

National Research Council. (1997). *Tactical Display for Soldiers: Human Factors Considerations*. National Academy Press: Washington D.C.

Neisser, U. and Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, 7, 480-494.

Osgood, R.K., Geiselman, E.E., and Calhoun, C.S. (1991). Attitude maintenance using an off-boresight helmet-mounted virtual display. *AGARD Conference Proceedings 517: Helmet Mounted Display and Night Vision Goggles* (pp. 14.1 - 14.7). Neuilly Sur Seine, France, NATO Advisory Group for Aerospace Research and Development (NTIS No. AGARDCP 517).

Osgood, R.K. and Wells, M.J. (1991). The effect of field-of-view size on performance of a simulated air-to-ground night attack. *AGARD Conference Proceedings 517: Helmet Mounted Display and Night Vision Goggles* (pp. 10-1 - 10-7). Neuilly Sur Seine, France, NATO Advisory Group for Aerospace Research and Development (NTIS No. AGARDCP 517).

Pausch, R. (1991). Virtual reality on five dollars a day. *Proceedings of the CHI ’91 Conference on Human Factors in Computing Systems* (pp. 265-269). New York, NY: ACM.

Pausch, R., Proffitt, D., and Williams, G. (1997). Quantifying immersion in virtual reality. *ACM SIGGRAPH `97 Conference Proceedings, Computer Graphics*, August. (submitted for review)

Pausch, R., Shackelford, M.A., and Proffitt, D. (1993). A user study comparing head-mounted and stationary displays. *IEEE Symposium on Research Frontiers in Virtual Reality*. Los Alamitos, CA: IEEE.

Pepper, R.L, Cole, R.E., Spain, E.H., and Sigurdson, J.E. (1983). Research issues involved in applying stereoscopic television to remotely operated vehicles. In J. Ebbeni and A. Monfils (Eds.), *Proceedings of the International Society for Optical Engineers (SPIE): Three-Dimensional Imaging* (pp. 170-174). 402. Bellingham, WA: SPIE.

Pepper, R.L., Smith, D.C., and Cole, R.E. (1981). Stereo TV improves operator performance under degraded visibility conditions. *Optical Engineering*. 20(4), 579 - 585.

Rekimoto, J. (1995). A vision-based head tracker for fish tank virtual reality – VR without head gear. In *Proceedings of the Virtual Reality Annual International Symposium* (pp. 74-82). Los Alamitos, CA: IEEE.

Rohaly, A.M. and Karsh, R. (1997). The effects of divided attention on peripheral target localization. *ARL-TR-1334*. Aberdeen Proving Ground, MD: Army Research Laboratory.

Rolland, J.P., Biocca, F.A., Barlow, T., and Kancerla, A. (1995). Quantification of adaptation to virtual-eye location in see-through head-mounted displays. *1995 IEEE Virtual Reality International Symposium* (pp. 56 - 66). Los Alamitos, CA: IEEE.

Sampson, J.B. (1993) Cognitive performance by individuals using a head-mounted display while walking. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Sampson, J.B. and Warren, P. H. (1992). Human factors of visual display systems for the individual soldier. In S. Kumar (Ed.), *Advances in Industrial Ergonomics and Safety IV* (pp. 1305 - 1309). London: Taylor and Francis.

Seagull, F.J. and Gopher, D. (1997). Training head movement in visual scanning: An embedded approach to the development of piloting skills with helmet-mounted displays. *Journal of Experimental Psychology: Applied*. In press.

Shimojo, S. and Nakayama, K. (1990). Real world occlusion constraints and binocular rivalry. *Vision Research*. 30(1), 69-80.

Sollenberger, R.L. and Milgram, P. (1993). Effects of stereoscopic and rotational displays ion a three-dimensional path tracing task. *Human Factors*. 35(3), 483-499.

Teichner, W.H. and Mocharnuk, J.B. (1979). Visual search for complex targets. *Human Factors*. 21(3), 259-275.

Ververs, P.M. and Wickens, C.D. (1996). Allocation of attention with head-up displays. *University of Illinois Institute of Aviation Technical Report (ARL-96-1/FAA-96-1).* Savoy, IL: Aviation Research Laboratory.

Ware, C. (1995). Dynamic stereo displays. In *Chi '95 Mosaic of Creativity* (pp. 310 - 316).

Weintraub, D.J., Haines, R.F., and Randle, R.J. (1985). Head-up display (HUD) utility, II: Runway to HUD transitions monitoring eye focus and decision times. *Proceedings of the 29th Annual Meeting of the Human Factors Society*. Santa Monica, CA: Human Factors Society.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance*. New York: Harper-Collins Publishers.

Wickens, C.D. (1996). Augmented reality: Cognitive issues in helmet-mounted displays. Foundations of a computational model. Unpublished manuscript.

Wickens, C.D. (1997a). Attentional issues in head-up displays. In: *Engineering Psychology and Cognitive Ergonomics: Integration of Theory and Application*. London: Avebury Technical Pub. Co.

Wickens, C.D. (1997b). Frame of reference for navigation. In D. Gopher & A. Koriat (Eds.), *Attention and Performance*, Vol. 16. Orlando, FL: Academic Press.

Wickens, C.D. and Baker, P. (1995). Cognitive issues in virtual reality. In W. Barfield and T.A. Furness (Eds). *Virtual Environments and Advanced Interface Design* (pp. 514-541). New York: Oxford University Press.

Wickens, C.D. and Long, J. (1995). Object versus space-based models of visual attention: Implications for the design of head-up displays. *Journal of Experimental Psychology: Applied*. 1(3), 179-193.

Wickens, C.D. and Prevett, T.T. (1995). Exploring the dimensions of egocentricity in aircraft navigation displays. *Journal of Experimental Psychology: Applied*. 1(2), 110-135.

Wickens, C.D., Todd, S., and Seidler, K. (1989). Three-dimensional displays: Perception, implementation, and applications. (CSERIAC SOAR-89-01). Wright-Patterson AFB, OH: Armstrong Aerospace Medical Research Laboratory.

Williams, S.P. and Parrish, R.V. (1990). In-simulator assessment of trade-offs arising from mixture of color cueing and monocular, binoptic, and stereopsis cueing information. *Southeastcon '90 Proceedings* (pp. 583-588). Los Alamitos, CA: IEEE.

Yantis, S. (1992) Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*. 24, 295-340.

Zenyuh, J.P, Reising, J.M., Walchli, S., and Biers, D. (1988). A comparison of a stereographic 3-D display versus a 2-D display using an advanced air-to-air format. In *Proceedings of the Human Factors Society – 32nd Annual Meeting*. Santa Monica, CA: Human Factors Society.

Appendix 1: ANOVA Tables

Expectancy

Response Time

Source of Variation	Df	*SS*	*MS*
Military	1	15.15	15.15
Display	1	.373	.373
Military x Display	1	14.48	14.48
Error	12	71.86	5.99
Target Type	3	336.28	112.09
Target Type x Military	3	2.43	.809
Target Type x Display	3	2.31	.772
Target Type x Military x Display	3	3.29	1.08
Error (Target Type)	36	68.0	1.89
Eye	1	.001	.001
Eye x Military	1	1.43	1.43
Eye x Display	1	.018	.018
Eye x Military x Display	1	1.01	1.01
Error (Eye)	12	38.62	3.21
Target Type x Eye	3	4.11	1.37
Target Type x Eye x Military	3	.648	.216
Target Type x Eye x Display	3	7.78	2.59
Target Type x Eye x Military x Display	3	1.34	.445
Error (Target Type x Eye)	36	49.74	1.38

Accuracy

Source of Variation	Df	*SS*	*MS*
Military	1	.000	.000
Display	1	.032	.032
Military x Display	1	.041	.041
Error	12	.499	.041
Target Type	3	7.54	2.51
Target Type x Military	3	.032	.011
Target Type x Display	3	.075	.025
Target Type x Military x Display	3	.217	.072
Error (Target Type)	36	.797	.022
Eye	1	.013	.013
Eye x Military	1	.000	.000
Eye x Display	1	.001	.001
Eye x Military x Display	1	.006	.006
Error (Eye)	12	.416	.035
Target Type x Eye	3	.004	.001
Target Type x Eye x Military	3	.007	.002
Target Type x Eye x Display	3	.081	.027
Target Type x Eye x Military x Display	3	.107	.036
Error (Target Type x Eye)	36	1.08	.030

Cueing

Response Time

Source of Variation	df	*SS*	*MS*
Military	1	24.19	24.19
Display	1	3.97	3.97
Military x Display	1	68.35	68.35
Error	12	311.58	25.964
Eye	1	29.13	29.13
Eye x Military	1	1.11	1.11
Eye x Display	1	4.38	4.38
Eye x Military x Display	1	20.03	20.03
Error (Eye)	12	235.18	19.60
Wall	2	156.70	78.35
Wall x Military	2	45.24	22.62
Wall x Display	2	21.49	10.74
Wall x Military x Display	2	16.91	8.45
Error (wall)	24	307.52	12.81
Cueing	1	2403.61	2403.61
Cueing x Military	1	3.08	3.08
Cueing x Display	1	34.19	34.19
Cueing x Military x Display	1	7.21	7.21
Error (Cueing)	12	148.47	12.37
Eye x Wall	2	18.59	9.29
Eye x Wall x Military	2	28.18	14.09
Eye x Wall x Display	2	1.04	.52
Eye x Wall x Military x Display	2	14.31	7.16
Error (Eye x Wall)	24	173.48	7.23
Eye x Cueing	1	1.19	1.19
Eye x Cueing x Military	1	.480	.480
Eye x Cueing x Display	1	12.12	12.12
Eye x Cueing x Military x Display	1	6.73	6.73
Error (Eye x Cueing)	12	304.29	25.36
Wall x Cueing	2	73.31	36.66
Wall x Cueing x Military	2	40.47	20.23
Wall x Cueing x Display	2	6.12	3.06
Wall x Cueing x Military x Display	2	19.95	9.98
Error (Wall x Cueing)	24	230.23	9.18
Eye x Wall x Cueing	2	9.99	5.00
Eye x Wall x Cueing x Military	2	29.50	14.75
Eye x Wall x Cueing x Display	2	.906	.453
Eye x Wall x Cueing x Military x Display	2	1.89	.946
Error (Eye x Wall x Cueing)	24	241.91	10.08

Accuracy

Source of Variation	df	*SS*	*MS*
Military	1	.001	.001
Display	1	.003	.003
Military x Display	1	.008	.008
Error	12	.075	.075
Eye	1	.001	.001
Eye x Military	1	.000	.000
Eye x Display	1	.001	.001
Eye x Military x Display	1	.003	.003
Error (Eye)	12	.015	.001
Wall	2	.003	.001
Wall x Military	2	.003	.001
Wall x Display	2	.000	.000
Wall x Military x Display	2	.004	.002
Error (wall)	24	.030	.001
Cueing	1	.013	.013
Cueing x Military	1	.001	.001
Cueing x Display	1	.003	.003
Cueing x Military x Display	1	.008	.008
Error (Cueing)	12	.075	.006
Eye x Wall	2	.000	.000
Eye x Wall x Military	2	.004	.002
Eye x Wall x Display	2	.003	.001
Eye x Wall x Military x Display	2	.003	.001
Error (Eye x Wall)	24	.030	.001
Eye x Cueing	1	.001	.001
Eye x Cueing x Military	1	.000	.000
Eye x Cueing x Display	1	.001	.001
Eye x Cueing x Military x Display	1	.003	.003
Error (Eye x Cueing)	12	.015	.001
Wall x Cueing	2	.003	.001
Wall x Cueing x Military	2	.003	.001
Wall x Cueing x Display	2	.000	.000
Wall x Cueing x Military x Display	2	.004	.002
Error (Wall x Cueing)	24	.030	.001
Eye x Wall x Cueing	2	.000	.000
Eye x Wall x Cueing x Military	2	.004	.002
Eye x Wall x Cueing x Display	2	.003	.001
Eye x Wall x Cueing x Military x Display	2	.003	.001
Error (Eye x Wall x Cueing)	24	.030	.001

Secondary Task Performance

Response Time

Source of Variation	df	*SS*	*MS*
Military	1	.047	.047
Display	1	.498	.498
Military x Display	1	.489	.489
Error	12	12.9	1.08
Eye	1	.165	.165
Eye x Military	1	.318	.318
Eye x Display	1	.005	.005
Eye x Military x Display	1	.232	.232
Error (Eye)	12	3.15	.263
Cueing	1	.002	.002
Cueing x Military	1	.037	.037
Cueing x Display	1	.046	.046
Cueing x Military x Display	1	.253	.253
Error (Cueing)	12	1.32	.110
Eye x Cueing	1	.278	.278
Eye x Cueing x Military	1	.391	.391
Eye x Cueing x Display	1	.211	.211
Eye x Cueing x Military x Display	1	.237	.237
Error (Eye x Cueing)	12	.865	.072

Accuracy

Source of Variation	df	*SS*	*MS*
Military	1	.019	.019
Display	1	.005	.005
Military x Display	1	.000	.000
Error	11	.169	.015
Eye	1	.000	.000
Eye x Military	1	.004	.004
Eye x Display	1	.013	.013
Eye x Military x Display	1	.004	.004
Error (Eye)	11	.072	.007
Cueing	1	.012	.012
Cueing x Military	1	.012	.012
Cueing x Display	1	.001	.001
Cueing x Military x Display	1	.001	.001
Error (Cueing)	11	.061	.006
Eye x Cueing	1	.001	.001
Eye x Cueing x Military	1	.001	.001
Eye x Cueing x Display	1	.005	.005
Eye x Cueing x Military x Display	1	.001	.001
Error (Eye x Cueing)	11	.029	.026

Global Positioning Task

Response Time

Source of Variation	df	*SS*	*MS*
Model	17	20180.54	1187.09
Error	63	71124.96	1128.97
Corrected Total	80	91305.50	
Display	1	2.56	2.56
Eye	1	3642.87	3642.87
Military	1	15.32	15.32
Display x Eye	1	394.69	394.69
Display x Military	1	1146.53	1146.53
Terrain	4	6349.80	1609.95
Eye x Terrain	4	474.37	118.59
Military x Terrain	4	8064.41	2016.10

Accuracy

Source of Variation	df	*SS*	*MS*
Model	17	6.71	.394
Error	141	32.91	.233
Corrected Total	158	39.62	
Display	1	.04	.04
Eye	1	.13	.13
Military	1	1.50	1.50
Display x Eye	1	.26	.26
Display x Military	1	.27	.27
Terrain	4	2.53	2.53
Eye x Terrain	4	.79	.20
Military x Terrain	4	1.20	.30

Performance Ratings

Source of Variation	df	*SS*	*MS*
Model	17	144.87	8.52
Error	141	813.93	5.77
Corrected Total	158	958.69	
Display	1	2.79	2.79
Eye	1	1.86	1.86
Military	1	35.11	35.11
Display x Eye	1	13.78	13.78
Display x Military	1	6.09	6.09
Terrain	4	42.86	10.72
Eye x Terrain	4	26.03	6.51
Military x Terrain	4	16.35	4.09

Target Identification

Response Time

Source of Variation	df	*SS*	*MS*
Military	1	.99	.99
Display	1	1.84	1.84
Military x Display	1	1.65	1.65
Error	12	6.00	.50
Target Type	1	.04	.04
Target Type x Military	1	.000	.000
Target Type x Display	1	.001	.001
Target Type x Military x Display	1	.007	.007
Error (Target Type)	12	.170	.014
Eye	1	.005	.005
Eye x Military	1	.031	.031
Eye x Display	1	.001	.001
Eye x Military x Display	1	.004	.004
Error (Eye)	12	.577	.048
Target Type x Eye	1	.003	.003
Target Type x Eye x Military	1	.028	.028
Target Type x Eye x Display	1	.003	.003
Target Type x Eye x Military x Display	1	.001	.001
Error (Target Type x Eye)	12	.141	.011

Accuracy

Source of Variation	df	*SS*	*MS*
Military	1	.001	.001
Display	1	.010	.010
Military x Display	1	.001	.001
Error	12	.025	.002
Target Type	1	.001	.001
Target Type x Military	1	.000	.000
Target Type x Display	1	.002	.002
Target Type x Military x Display	1	.000	.000
Error (Target Type)	12	.006	.000
Eye	1	.000	.000
Eye x Military	1	.002	.002
Eye x Display	1	.001	.001
Eye x Military x Display	1	.002	.002
Error (Eye)	12	.011	.001
Target Type x Eye	1	.000	.000
Target Type x Eye x Military	1	.000	.000
Target Type x Eye x Display	1	.001	.001
Target Type x Eye x Military x Display	1	.000	.000
Error (Target Type x Eye)	12	.009	.001

Target Heading

Response Time

Source of Variation	df	*SS*	*MS*
Military	1	1.47	1.47
Display	1	6.39	6.39
Military x Display	1	.14	.14
Error	12	11.29	.94
Target Type	3	3.64	1.21
Target Type x Military	3	.310	.103
Target Type x Display	3	.376	.125
Target Type x Military x Display	3	.134	.045
Error (Target Type)	363	2.17	.060
Eye	1	.155	.155
Eye x Military	1	.223	.223
Eye x Display	1	.017	.017
Eye x Military x Display	1	.172	.172
Error (Eye)	12	.635	.053
Target Type x Eye	3	.046	.015
Target Type x Eye x Military	3	.063	.021
Target Type x Eye x Display	3	.036	.012
Target Type x Eye x Military x Display	3	.023	.008
Error (Target Type x Eye)	36	.741	.021

Accuracy

Source of Variation	df	*SS*	*MS*
Military	1	.001	.001
Display	1	.003	.003
Military x Display	1	.253	.253
Error	12	1.37	.114
Target Type	3	.838	.279
Target Type x Military	3	.017	.006
Target Type x Display	3	.045	.015
Target Type x Military x Display	3	.005	.002
Error (Target Type)	36	.101	.003
Eye	1	.002	.002
Eye x Military	1	.001	.001
Eye x Display	1	.006	.006
Eye x Military x Display	1	.007	.007
Error (Eye)	12	.251	.021
Target Type x Eye	3	.008	.003
Target Type x Eye x Military	3	.002	.001
Target Type x Eye x Display	3	.015	.005
Target Type x Eye x Military x Display	3	.002	.001
Error (Target Type x Eye)	36	.125	.003

Scanning Strategies

Number of times the target passed through the subjects' 60° field of view

Source of Variation	df	*SS*	*MS*
Military	1	.175	.175
Display	1	.020	.020
Military x Display	1	.109	.109
Error	12	2.63	.219
Eye	1	1.07	1.07
Eye x Military	1	.002	.002
Eye x Display	1	.223	.223
Eye x Military x Display	1	.074	.074
Error (Eye)	12	2.18	.181
Wall	2	18.0	9.00
Wall x Military	2	.122	.061
Wall x Display	2	.044	.022
Wall x Military x Display	2	.325	.162
Error (wall)	24	2.27	.095
Target Type	3	36.1	12.0
Target Type x Military	3	.278	.092
Target Type x Display	3	.160	.053
Target Type x Military x Display	3	.878	.293
Error (target type)	36	4.36	.121
Eye x Wall	2	.915	.458
Eye x Wall x Military	2	.491	.245
Eye x Wall x Display	2	.304	.152
Eye x Wall x Military x Display	2	.629	.314
Error (Eye x Wall)	24	1.22	.051
Eye x Target Type	3	.563	.188
Eye x Target Type x Military	3	.111	.037
Eye x Target Type x Display	3	.432	.144
Eye x Target Type x Military x Display	3	.056	.019
Error (Eye x Target Type)	36	4.31	.120
Wall x Target Type	6	8.66	1.44
Wall x Target Type x Military	6	.392	.065
Wall x Target Type x Display	6	.511	.085
Wall x Target Type x Military x Display	6	.835	.139
Error (Wall x Target Type)	72	7.27	.101
Eye x Wall x Target Type	6	.703	.117
Eye x Wall x Target Type x Military	6	.715	.119
Eye x Wall x Target Type x Display	6	.288	.048
Eye x Wall x Target Type x Military x Display	6	.928	.155
Error (Eye x Wall x Target Type)	72	5.47	.076

Number of times the target passed through the subjects’ 40° field of view

Source of Variation	df	*SS*	*MS*
Military	1	1.75	1.75
Display	1	.031	.031
Military x Display	1	.284	.284
Error	12	6.41	.542
Eye	1	.014	.014
Eye x Military	1	.161	.161
Eye x Display	1	.089	.089
Eye x Military x Display	1	.075	.075
Error (Eye)	12	2.77	.231
Wall	2	8.91	4.45
Wall x Military	2	.684	.342
Wall x Display	2	.272	.136
Wall x Military x Display	2	.122	.061
Error (wall)	24	4.75	.198
Target Type	3	13.52	4.51
Target Type x Military	3	.370	.123
Target Type x Display	3	.766	.255
Target Type x Military x Display	3	.092	.031
Error (target type)	36	6.67	.186
Eye x Wall	2	.750	.375
Eye x Wall x Military	2	.044	.022
Eye x Wall x Display	2	.627	.314
Eye x Wall x Military x Display	2	.327	.164
Error (Eye x Wall)	24	3.66	.152
Eye x Target Type	3	.293	.098
Eye x Target Type x Military	3	.318	.106
Eye x Target Type x Display	3	.014	.005
Eye x Target Type x Military x Display	3	.068	.023
Error (Eye x Target Type)	36	4.49	.125
Wall x Target Type	6	5.68	.947
Wall x Target Type x Military	6	.780	.130
Wall x Target Type x Display	6	.811	.135
Wall x Target Type x Military x Display	6	.808	.135
Error (Wall x Target Type)	72	7.94	.110
Eye x Wall x Target Type	6	.513	.086
Eye x Wall x Target Type x Military	6	.153	.026
Eye x Wall x Target Type x Display	6	.742	.124
Eye x Wall x Target Type x Military x Display	6	.430	.072
Error (Eye x Wall x Target Type)	72	11.92	.166

Number of times the target passed through the subjects' 15° field of view

Source of Variation	df	*SS*	*MS*
Military	1	.106	.106
Display	1	.720	.720
Military x Display	1	.780	.780
Error	12	4.25	.354
Eye	1	.169	.169
Eye x Military	1	.169	.169
Eye x Display	1	.005	.005
Eye x Military x Display	1	.132	.132
Error (Eye)	12	2.14	.179
Wall	2	2.76	1.38
Wall x Military	2	.250	.125
Wall x Display	2	.141	.071
Wall x Military x Display	2	.399	.200
Error (wall)	24	2.66	.111
Target Type	3	3.22	1.07
Target Type x Military	3	.039	.013
Target Type x Display	3	.087	.029
Target Type x Military x Display	3	.163	.054
Error (target type)	36	2.19	.061
Eye x Wall	2	.094	.045
Eye x Wall x Military	2	.011	.005
Eye x Wall x Display	2	.057	.028
Eye x Wall x Military x Display	2	.464	.232
Error (Eye x Wall)	24	2.27	.095
Eye x Target Type	3	.152	.051
Eye x Target Type x Military	3	.083	.028
Eye x Target Type x Display	3	.235	.078
Eye x Target Type x Military x Display	3	.152	.051
Error (Eye x Target Type)	36	2.80	.078
Wall x Target Type	6	4.04	.673
Wall x Target Type x Military	6	.125	.021
Wall x Target Type x Display	6	.356	.059
Wall x Target Type x Military x Display	6	.183	.030
Error (Wall x Target Type)	72	4.24	.049
Eye x Wall x Target Type	6	.235	.039
Eye x Wall x Target Type x Military	6	.205	.034
Eye x Wall x Target Type x Display	6	.168	.028
Eye x Wall x Target Type x Military x Display	6	.435	.072
Error (Eye x Wall x Target Type)	72	4.27	.059