

Technical Report 1215

**Effects of Spatial and Non-Spatial Multi-Modal Cues
on Orienting of Visual-Spatial Attention in an
Augmented Environment**

Christian J. Jerome
U. S. Army Research Institute

November 2007



**United States Army Research Institute
for the Behavioral and Social Sciences**

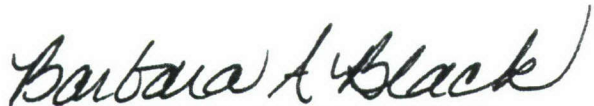
20080107180

Approved for public release; distribution is unlimited.

**U.S. Army Research Institute
for the Behavioral and Social Sciences**

**A Directorate of the Department of the Army
Deputy Chief of Staff, G1**

Authorized and approved for distribution:



**BARBARA A. BLACK, Ph.D.
Research Program Manager
Training and Leader Development**



**MICHELLE SAMS, PhD.
Director**

Technical review by

Dr. John Barnett, U.S. Army Research Institute
Dr. Michael Singer, U.S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Technical Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-MS, 2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926.

FINAL DISPOSITION: This Technical Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) October 2007		2. REPORT TYPE Final		3. DATES COVERED (from... to) July 2003 –November 2007	
4. TITLE AND SUBTITLE Effects Of Spatial and Non-Spatial Multi-Modal Cues on Orienting of Visual-Spatial Attention in an Augmented Environment				5a. CONTRACT OR GRANT NUMBER 	
				5b. PROGRAM ELEMENT NUMBER 622785	
6. AUTHOR(S) Christian J. Jerome (U.S. Army Research Institute)				5c. PROJECT NUMBER A790	
				5d. TASK NUMBER 233	
				5e. WORK UNIT NUMBER H01	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: DAPE-ARI-IF 12350 Research Parkway, Orlando, FL 32826-3276				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 2511 Jefferson Davis Highway Arlington, VA 22202-3926				10. MONITOR ACRONYM ARI	
				11. MONITOR REPORT NUMBER Technical Report 1215	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES Subject Matter POC: Christian Jerome					
14. ABSTRACT (<i>Maximum 200 words</i>): Visual search tasks are known to be cognitive capacity demanding and therefore may be improved by training in an augmented reality (AR) environment. During the experimental task, 64 participants searched for enemies (while cued from visual, auditory, tactile, combinations of two, or all three modality cues) and tried to shoot them while avoiding shooting the civilians (fratricide) for two 2-minute low-workload scenarios, and two 2-minute high-workload scenarios. The results showed significant benefits of attentional cueing on visual search task performance. These benefits were revealed by improved performance in reaction time and accuracy from the haptic cues alone, auditory cues alone, and the combination of the visual and haptic cues together. Fratricide occurrence was shown to be amplified by the presence of the audio cues. The two levels of workload produced differences within individual's task performance for accuracy and reaction time. Accuracy and reaction time were significantly better with the medium cues than all the other cue specificities and the control condition during low workload and marginally better during high workload. Cue specificity generally resulted in better accuracy and reaction time with the medium cues.					
15. SUBJECT TERMS Augmented Environments, Augmented Reality, Dismounted Infantry, Training, Presence, Simulator Sickness					
SECURITY CLASSIFICATION OF			17. LIMITATION OF ABSTRACT Unlimited	18. NUMBER OF PAGES 63	21. RESPONSIBLE PERSON Ellen Kinzer Technical Publication Specialist (703) 602-8047
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified			

Technical Report 1215

**Effects of Spatial and Non-Spatial Multi-Modal Cues on
Orienting of Visual-Spatial Attention in an Augmented
Environment**

Christian J. Jerome
U. S. Army Research Institute

Technology-Based Training Research Unit
Stephen Goldberg, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926

November 2007

Army Project Number
622785A790

Personnel Performance
and Training Technology

Approved for public release; distribution is unlimited.

ACKNOWLEDGMENT

This research could not have been completed without the support of the Consortium of Universities, especially Dr. Robert Ruskin, the University of Central Florida, especially Dr. Mustapha Mouloua, the Institute for Simulation and Training, especially the Media Convergence Laboratory, and Dr. Bob Witmer of the Army Research Institute. I would also like to thank my reviewers, Drs. Michael Singer and John Barnett for their insightful comments and recommendations.

EFFECTS OF SPATIAL AND NON-SPATIAL MULTI-MODAL CUES ON ORIENTING OF VISUAL-SPATIAL ATTENTION IN AN AUGMENTED ENVIRONMENT

EXECUTIVE SUMMARY

Research Requirement:

Advances in simulation technology have brought about many improvements to the way we train tasks, as well as how operational tasks are performed in the field. Augmented reality (AR) is an example of how to enhance the user's experience in the real world with computer generated information and graphics. The purpose of this research was to determine if AR cueing could be used successfully to focus the user's attention on specific locations in the environment.

Procedure:

Visual search tasks are known to be capacity demanding and therefore may be improved by training in an AR environment. The first step is demonstrating that the performance can be improved within the training environment. During the experimental training task, 64 participants searched for enemies (while cued from visual, auditory, tactile, combinations of two, or all three modality cues) and tried to shoot them while avoiding shooting the civilians (fratricide) for two 2-minute low-workload scenarios, and two 2-minute high-workload scenarios. The attention cues were also varied in the amount of spatial information they possessed (i.e., specificity), including a control condition (no specificity), small, medium, and large cue specificities. Measures of performance included accuracy and reaction time.

Findings:

The results showed significant benefits of attentional cueing on visual search task performance as revealed by benefits in reaction time and accuracy from the presence of the haptic cues and auditory cues when displayed alone and the combination of the visual and haptic cues together. Fratricide occurrence was shown to be amplified by the presence of the audio cues. The two levels of workload produced differences within individual's task performance for accuracy and reaction time; counter intuitively, low-workload levels produced lower performance than high-workload levels. Accuracy and reaction time were significantly better with the medium cues than all the others cue specificities and the control condition during low-workload and marginally better during high-workload. Cue specificity generally resulted in better accuracy and reaction time with the medium cues.

Utilization and Dissemination of Findings:

These results are in support of Posner's (1978) theory that, in general, cueing can benefit locating targets in the environment by aligning the attentional system with the visual input pathways. Since attention can be cued using this AR system, increasing accuracy and efficiency performance, tasks can be trained using such systems in order to teach and provide practice in visual search tasks. The cue modality does not have to match the target modality. This research is relevant to potential applications of AR technology. Furthermore, the results identify and describe perceptual and/or cognitive issues with the use of displaying computer generated augmented objects and information overlaid upon the real world. The results also serve as a basis

for providing a variety of training and design recommendations to direct attention during military operations. Such recommendations include cueing the Soldier to the location of hazards, and mitigating the effects of stress and workload.

EFFECTS OF SPATIAL AND NON-SPATIAL MULTI-MODAL CUES ON ORIENTING OF VISUAL-SPATIAL ATTENTION IN AN AUGMENTED ENVIRONMENT

CONTENTS

	Page
INTRODUCTION	1
Attention Theory.....	2
Contemporary Attention Theory.....	3
Detection of Signals.....	6
Orienting Attention and Orienting Reflex	6
Cueing of Attention	7
Auditory.....	7
Haptic.....	8
Visual.....	8
Cue Reliability/Trust.....	9
Workload	9
Multiple Resource Theory	9
Breadth of Attention	10
Augmented Reality	11
Rationale for Present Research	12
Research Hypotheses and Predictions	13
METHOD	15
Participants.....	15
Apparatus	15
Augmented Reality Simulator.....	15
Video Display.	15
Audio Display.	16
Haptic Display.	16
Tracking Device.....	16
Input Device.....	16
Computers.....	16
Research Methodology	17
Cues.....	17
Design.	17
Procedure.	18
Independent Variables	19
Between Subjects.....	19
Within Subjects.....	19
Dependent Variables.....	19
Task Performance.	19
Secondary Task.....	19
Questionnaires.....	19

CONTENTS (continued)

RESULTS	20
Descriptive Statistics.....	20
Tests of Specific Hypotheses.....	20
Other Interesting Findings	30
Interactions.....	30
Other Measures of Interest.....	34
Presence.....	34
Simulator Sickness.....	35
Secondary Task Performance	35
DISCUSSION	37
Cueing Modality Effects.....	37
Workload Analysis	39
Cue Specificity Effects	40
Implications	40
Cueing Guidelines.....	42
Conclusion	42
REFERENCES	43
APPENDIX A: TASK SCENARIO CHARACTERS.....	A-1
APPENDIX B: MR MOUT PORTAL MAPS	B-1
APPENDIX C: BIOGRAPHICAL QUESTIONNAIRE	C-1

LIST OF FIGURES

FIGURE 1. BROADBENT'S (1958) MODEL OF ATTENTION (AFTER BROADBENT, 1958, P. 299).....	4
FIGURE 2. MODIFIED MODEL OF ATTENTION.....	5
FIGURE 3. WICKENS' (1984) MODEL OF INFORMATION PROCESSING (AFTER WICKENS, 1992, P. 17).....	6
FIGURE 4. MODEL REPRESENTING HUMAN SYSTEMS INVOLVED IN THE EXPERIMENTAL TASKS	12
FIGURE 5. DIFFERENCE IN ACCURACY FOR ABSENCE AND PRESENCE OF VISUAL, AUDITORY, AND HAPTIC CUES.....	21

CONTENTS (continued)

FIGURE 6. DIFFERENCE IN REACTION TIME FOR ABSENCE AND PRESENCE OF VISUAL, AUDITORY, AND HAPTIC CUES.	22
FIGURE 7. OBSERVED ACCURACY PERFORMANCE ACROSS CUE MODALITY CONDITIONS.	23
FIGURE 8. OBSERVED REACTION TIME PERFORMANCE ACROSS CUE MODALITY CONDITIONS.	24
FIGURE 9. FRATRICIDE OCCURRENCE FOR THE ABSENCE AND PRESENCE OF VISUAL, AUDITORY, AND HAPTIC CUES.	25
FIGURE 10. FRATRICIDE OCCURRENCE INTERACTION BETWEEN VISUAL AND HAPTIC CUES	25
FIGURE 11. FRATRICIDE OCCURRENCE ACROSS CUE MODALITY CONDITIONS.	26
FIGURE 12. ACCURACY PERFORMANCE BETWEEN CUE MODALITIES FOR LOW AND HIGH WORKLOAD	27
FIGURE 13. ACCURACY OF TARGET ACQUISITION FOR THE FOUR LEVELS OF CUE SPECIFICITY.	28
FIGURE 14. REACTION TIME OF TARGET ACQUISITION FOR THE FOUR LEVELS OF CUE SPECIFICITY.	29
FIGURE 15. INTERACTION BETWEEN CUE SPECIFICITY AND AUDIO CUES FOR ACCURACY (LEFT PANEL) AND REACTION TIME (RIGHT PANEL).	30
FIGURE 16. INTERACTION BETWEEN CUE SPECIFICITY AND HAPTIC CUES FOR ACCURACY (LEFT PANEL) AND REACTION TIME (RIGHT PANEL).	31
FIGURE 17. THREE-WAY INTERACTION BETWEEN VISUAL CUES, AUDIO CUES, AND CUE SPECIFICITY.	31
FIGURE 18. THREE-WAY INTERACTION BETWEEN VISUAL CUES, AUDIO CUES, AND CUE SPECIFICITY FOR REACTION TIME.	32
FIGURE 19. FOUR-WAY INTERACTION FOR ACCURACY FOR WORKLOAD LEVEL X VISUAL CUES X AUDIO CUES X CUE SPECIFICITY.	33
FIGURE 20. FOUR-WAY INTERACTION FOR REACTION TIME FOR WORKLOAD LEVEL X VISUAL CUES X AUDIO CUES X CUE SPECIFICITY.	34

CONTENTS (continued)

FIGURE 21. EFFECTS OF LOW AND HIGH WORKLOAD FOR ACCURACY FOR LIGHT EXTINGUISHING.	36
FIGURE 22. EFFECTS OF LOW AND HIGH WORKLOAD FOR REACTION TIME FOR LIGHT EXTINGUISHING.	36
FIGURE A-1. TASK SCENARIO CHARACTERS.....	A-1
FIGURE B-1. MR MOUT PORTAL MAPS.....	B-1

LIST OF TABLES

TABLE 1. DESIGN OF CUED ATTENTION EXPERIMENT	18
TABLE 2. DESCRIPTIVE STATISTICS FOR DEPENDENT MEASURES	20
TABLE 3. DESCRIPTIVE STATISTICS FOR ACCURACY AND REACTION TIME MEASURES FOR EACH CONDITION	21
TABLE 4. PLANNED COMPARISONS FOR CUE MODALITY CONDITIONS	23
TABLE 5. DESCRIPTIVE STATISTICS FOR FRATRICIDE OCCURRENCE.....	24
TABLE 6. PLANNED COMPARISONS FOR CUE MODALITY CONDITIONS FOR FRATRICIDE	26
TABLE 7. DESCRIPTIVE STATISTICS FOR WORKLOAD.....	27
TABLE 8. CORRELATIONS OF OTHER MEASURES OF INTEREST	35

EFFECTS OF SPATIAL AND NON-SPATIAL MULTI-MODAL CUES ON ORIENTING OF VISUAL-SPATIAL ATTENTION IN AN AUGMENTED ENVIRONMENT

Introduction

Humans are endowed with the ability to take in information from the environment by transforming energy at the sense organs into electro-chemical neural activity. The mechanisms by which each sense modality transforms energy, however, have certain perceptual capabilities and limitations. For example, humans can hear a wide range of sounds but are limited by the frequency and intensity that can cause a sensory neuron to fire. Clearly, if the sensory neuron does not fire, the stimuli will not be sensed, perceived, or attended to. Being aware of these limitations, we have developed sensory aids that can enhance or amplify the environmental signals so that we may be able to better detect and react to them. These aids include such simple devices as sunglasses to help improve vision by reducing glare and eyeglasses that correct abnormalities in the shape of the eye, smoke and carbon monoxide detectors to warn us of the presence of a fire before our senses can, to telescopes and binoculars that make objects visible that would normally be impossible to see due to their small retinal image.

Sensory limitations are not exclusive; we have attention limitations as well. Much research has been done exploring human limitations such as dual-task performance, i.e. the inability to perform two tasks simultaneously when they compete for the same attentional resources, visual search problems, i.e. the target of interest is surrounded by non-target distracters and does not 'pop out' leading to a longer (serial) search time. On the other hand, a distracter that 'pops out' can pull attention away from the true target, and thus undermine target detection performance. These issues have been studied from different applied goals, e.g. aviation (Bronkhorst, Veltman, & Breda, 1996), automotive driving (McKnight & McKnight, 1993), and target detection during military operations (Itti, Koch, & Niebur, 1998). What they all have in common is that without knowledge about attentional limitations, the design of the interface and controls could inadvertently create a dangerous situation where accidents might happen and people could get hurt.

In an effort to overcome these attention limitations, system designers have made tasks simpler, reducing the amount of extraneous information that may distract from the stimuli that are of most importance to the task. However, what if the task cannot be simplified any more? Can additional information be added in order to improve the ability to pick out the important information in the environment from the less important/distractions? Augmented reality (AR) is one such tool that may prove to be a valuable attentional aid.

AR provides additional information overlaid upon the real world. The amount of information is thus increased into a small visual area, which can either aid or hinder the user's attentional processes. For example, if a text message is displayed, the individual may focus on the sensory stimuli of the text, higher cognitive processes will interpret the meaning of the words and message, and the individual will likely miss other potentially crucial stimuli coming from the rest of the visual field. However, text messages and other information may help in other ways. Information that helps draw the user's attention toward a target increases target detection efficiency by aligning the attentional system with the visual input pathways (Posner, 1980). As humans, we have a bias towards vision as the primary modality for taking in information. The information that steers attention does not necessarily need to be visual; it may be from any

modality that can provide spatial locations, however with potential for performance costs (Wickens & Liu, 1988).

The purpose of this research effort is to determine if the user can successfully focus attention on specific spatial locations of the visual scene when cued from either visual, auditory, tactile, or a combination of modalities, if there are any differences when the user is cued using similar cues but with no specific spatial location information, and if the user can be cued to focus attention of differing breadths from different cue specificities. The following section of this paper will summarize the literature, first from a basic level describing the theories and models of attention, then next from a more specific level describing the research findings from more recent work on orienting attention.

For years the limits of human attention have intrigued philosophers, theorists, and researchers. Exploring and explaining the processes underlying the causes of these limitations provides invaluable information that helps determine what types of situations are safe and, more importantly, unsafe for human operators. One such example is the maximum number of airplanes an air-traffic controller should have to coordinate (Hopkin, 1995). Other situations may be unsafe not due to the number of items required to monitor but because of the combination of one situation task with another (Wickens, 1984). UPS and FedEx employees drive trucks to deliver packages to addresses they have never visited before. Therefore, navigation is an important aspect of their job. However, navigation aids like global positioning systems (GPS) and other computer aided visual maps may contribute to distraction to the main task of driving the delivery truck (Jerome, Helmick, Mouloua, & Hancock, 2002). A general principle human factors engineers follow is to 'design out' errors. So in the case of the delivery truck, the navigation aid could be designed to be inoperable while the vehicle is in motion. The general idea of attention, the selecting of one thing to concentrate on at the expense of other things, how much an individual can concentrate on at one time, and what combination of tasks a person can safely handle at one time have been investigated theoretically and experimentally and will each be reviewed further as they apply.

Attention Theory

Over the years, attention has been defined many different ways, and the way it is defined is partially influenced by the zeitgeist at the time. Aristotle described it as a narrowing of the senses (Taylor, 2004). Others defined attention based on its particular properties, such as attention as active directing (Lucretius, 1st century B.C., as cited in Hatfield, 1998), attention as involuntary shifts (Augustine of Hippo, 400 A.D., as cited in Hatfield, 1998), attention as clarity (Buridan, 14th century, as cited in Hatfield, 1998), attention as fixation (Descartes, 17th century, as cited in Hatfield, 1998), attention as effector sensitivity (Descartes, but perhaps Lucretius, 1st century B.C., as cited in Hatfield, 1998). Others defined attention as a component or stage of consciousness. Wundt believed that there were two stages to consciousness. First, there was a working memory called the *Blickfeld* where ideas can be temporarily stored and manipulated, and second, there was an *Apperception*, which is under voluntary control, moving about the *blickfeld*, and can be thought of as selective attention (Hatfield, 1998). Others hypothesized about specific attention mechanisms, for example the ability to attend to specific spatial locations (Gibson, Von Helmholtz, Tichenor, as described in Van der Heijden, 1992). These early attempts to define and describe attention have paved the way for later more scientific definitions.

During the early days of psychological investigation, introspection was the main method of (so called) scientific psychological research. Procedurally, introspection simply consists of the individual reporting what is present in his or her conscious awareness during specific tasks or mental processes. Since there was very little science to this method, innovative ideas and findings were scarce even though the scientific question sought was valid and well thought out. For example, one major research question during this period was whether it was possible to divide attention between two or more things at once (Comte & James, as cited in Rosenthal, 1998). Without true experimental methods, it was impossible to provide strong evidence to support the ability or not.

Contemporary Attention Theory

Recent research interest has turned toward the focus of attention. Many studies have revealed the nature of auditory attention, specifically how different 'channels' of information cannot be attended to simultaneously. Cherry (1953) showed that people while listening and repeating back what they hear in one channel, or ear (a task called shadowing) could not report what the unattended channel message was. Many other researchers have investigated the same phenomenon and discovered other interesting characteristics of auditory attention including how characteristics of the attended message affect its detectability, and what characteristics of the message more frequently and easily seem to enter conscious recollection, e.g. if the speaker was male or female, high voice or low voice, etc. (Cherry, 1953). This line of research provides much description of various characteristics of how the auditory attentional system works, however a general model describing how the attentional system works was needed in order to explain how the general idea of attention has limitations and how the different senses bring that information into awareness.

Broadbent's (1958) "Filter" model of attention (see Figure 1) maps the flow of information from the senses through a number of processing stages. The four arrows at the left of Figure 1 represent the multiple simultaneous sensory inputs that all compete for attention controlled by the central selective filter. The selective filter chooses only one of the competing inputs for further processing, which is then passed on to the limited capacity channel, and then to two more advanced subsystems.

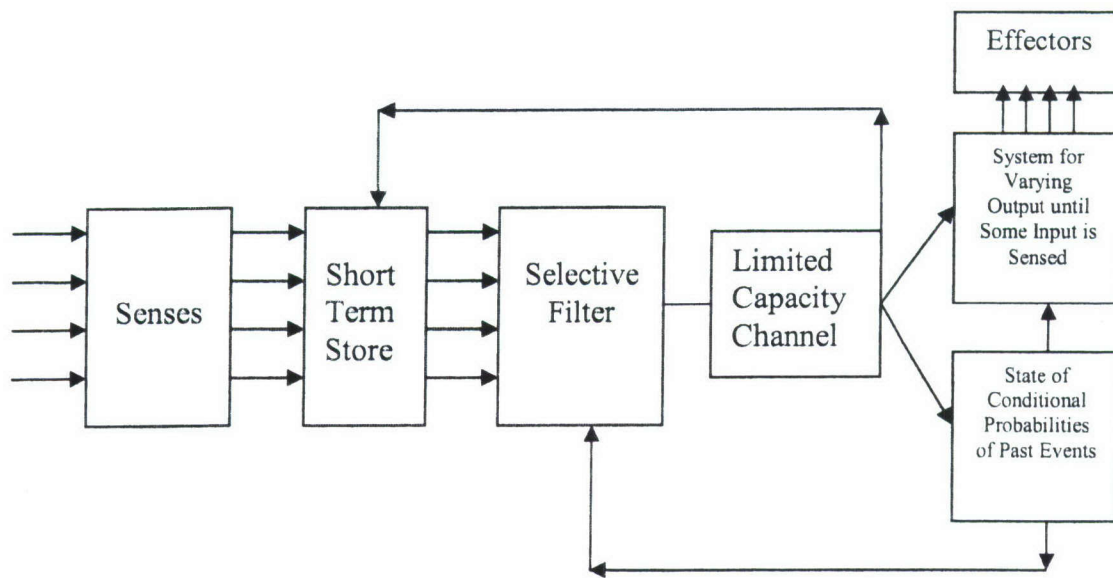
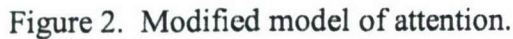


Figure 1. Broadbent's (1958) model of attention (after Broadbent, 1958, p. 299).

What this model basically tells us is that information comes in through the senses (visual, auditory, tactile, etc.) is temporarily stored, then some information is filtered and other information is focused based on importance or saliency. Thus, the filter exists near the actual stimulus itself along the information pathway and this type of filtering became known as *early selection* filtering. Attention shuts down the processing of the channel before the information can be analyzed semantically. There are those that believe that the filtering process takes place at much higher cognitive areas of the brain, and thus happens much later in the information pathway (closer to response selection) and appropriately is called *late selection* filtering (Deutsch and Deutsch, 1963). All information is analyzed semantically and is filtered much later just before reaching consciousness. The general benefits of an attentional cueing paradigm can be explained within these models. For example, the selective filter is augmented by the information; the important stimuli are cued and thus the filter's selection process is aided by the cue. The cue, in a way, acts as an intelligent external selective filter, since the cue is directed by information gathered from reconnaissance or electronic sensors capable of detecting and analyzing information that is beyond the sensory and cognitive capabilities of the individual, or can simply detect and analyze that data much quicker (see Figure 2).



Further work by Wickens (1984) has created an amalgamation of previous models in an effort to more clearly predict behavior for an applied utilization of information processing. Wickens' model (see Figure 3) is a hybrid mostly of Broadbent's model of attention (see Figure 1) and Atkinson and Shiffrin's (1968) Box Model. The most salient similarity to Broadbent's model is that there are limits to the amount of information that can enter into consciousness by entering through the senses and into the short-term store. Broadbent's model represents this with the limited number of arrows representing incoming stimuli, the selective filter that chooses only one sensory channel, and the 'limited capacity filter', while Wickens' model represents this with a limited amount of 'attentional resources' available to perception, decision making, decision execution, and working memory in general (see also Kahneman, 1973). The Box model is similar to Wickens' model in that they both show the flow of information from the environment, through the senses, into consciousness in a short-term storage system, and then either into long-term storage or action, or both. (The Box model is only mentioned here due to its influence on Wickens' model and is not described in detail because it says little about the limits of attention as they apply to the current work).

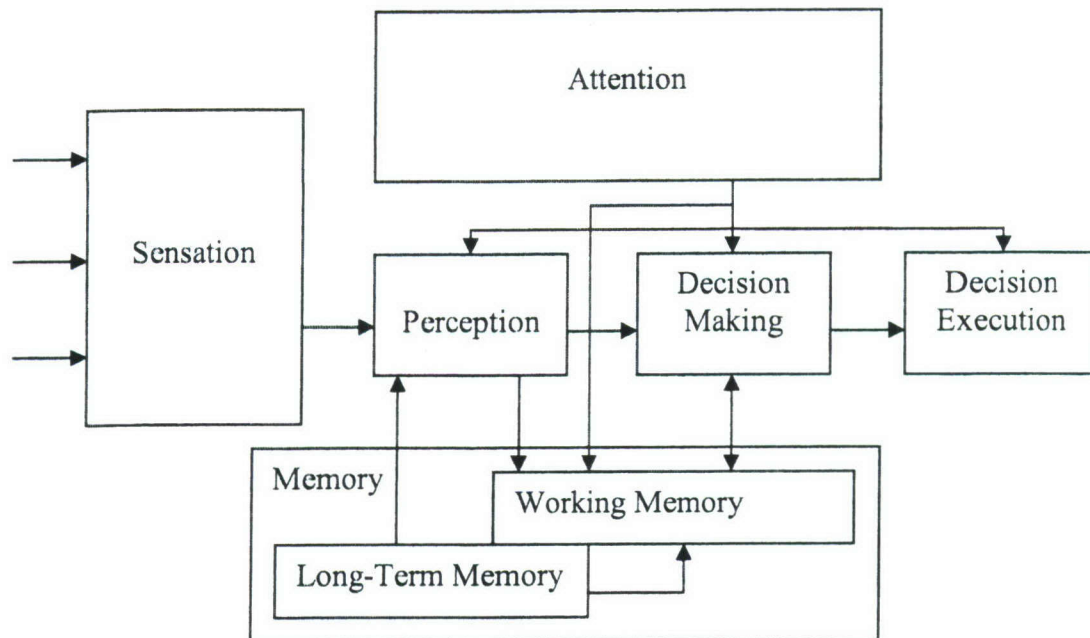


Figure 3. Wickens' (1984) model of information processing (after Wickens, 1992, p. 17).

The model of information processing is important to attention cueing in that it describes how there is a general resource for attention that feeds not only the perception of the sensory information, but also the process of decision making, decision execution, and working memory where the current thought processes take place. The attentional cue frees some of the attentional resources that are used to perceive the target, and those resources may be used to detect other cues or targets or may be used for other tasks like decision making.

Detection of Signals

Research has shown that the detection of signals can be made more efficient by providing information concerning the location of the target (Posner, Snyder, & Davidson, 1980). Apparently, this is caused by an aligning of the attentional system with the pathways involved with the visual location of the target (Posner, 1978), and it reduces the bandwidth or ranges of orientations to which these channels are sensitive (Blanco & Soto, 2002). The question then becomes: How does one provide the location information without distracting from that, or other tasks? The literature provides some insight to this question, and the current work attempts to add to the existing knowledge base.

Orienting Attention and Orienting Reflex

As explained previously, the environment delivers an abundance of stimuli from all modalities. Human limitations reduce our ability to pay attention to multiple channels, so we must choose which stimuli to devote attention to, or an important stimulus draws our attention to it by virtue of its characteristics (Treisman & Gormican, 1988; Broadbent, 1958; Wickens, 1984). Stimuli are categorized into two types based on the way they draw our attention: exogenous cueing and endogenous cueing. Exogenous cueing of attention is stimulus driven,

i.e., the stimulus draws attention to it based on its physical properties in a bottom up fashion. Little higher cognitive processing is involved in the decision to attend to it. Endogenous cueing of attention, on the other hand, is cognitively purposeful and goal driven, where a decision is made to attend to stimuli in a top-down fashion. Endogenous cues are usually displayed at the center of the visual field instead of at the target location.

Some of the earliest evidence suggesting differences in endogenous and exogenous covert orienting of attention was provided by Jonides (1980). His results showed that endogenous cueing was slower at covert orienting of attention than exogenous cueing. He also showed that endogenous cueing was affected by workload, while exogenous cueing was not. However, workload may still have an effect on exogenous cueing since, based upon the previously described models of attention, attentional resources are used to detect any sensory stimuli.

Many studies have shown an advantage when endogenous cues correspond to the target location, i.e., the arrow points to the target direction correctly (Bahri, 1989; Yeh & Wickens, 2000; Hillyard, Luck, Mouloua, Downing, & Woodward, 1990). Even when subjects were specifically told that the cues do not necessarily correspond to the location of the targets, there seems to be an innate reflex to orient attention to where the cue directs. For example, Friesen (2001) told subjects that the gaze direction of a schematic face was not predictive of subsequent target location, response time to target locations correctly cued by the gaze direction were faster than when the gaze direction did not correspond to the target location. Tipples (2002) showed that this was not unique to a schematic face/eye direction. Subjects were told that the arrow cues were not predictive of subsequent target locations, but response times to target locations correctly cued by the arrows were faster than when the arrows were not predictive of the target location. These results suggest that orienting attention to a cue can be highly reflexive and accurate cues are very important for a successful cueing system. Of equal importance are the characteristics of the cue which add to its orienting properties, including what sensory modality the cue is displayed from.

Cueing of Attention

Auditory. Providing additional information in order to improve the ability to pick out the important information in the environment from the less important/distractions is thought to be possible using the auditory modality. Exactly how one can determine the spatial location of a target based on its sound is accomplished by the inter-aural time difference (ITD), i.e., the minute time difference the sound reaches the two separate auditory sensory receivers (the ears).

Auditory cues to auditory targets show a performance improvement for targets on the expected side of the head, supporting the notion that ITDs can be used as a basis for orienting attention (Sach, Hill, & Bailey, 2000). Sach et al. (2000) also showed that a centrally located visual cue was successful in orienting attention for subsequent auditory targets (known as endogenous cueing), which supports top-down attention control, i.e., cognitive resources are required to process the information since the cue is located in a different location than the target, and that the cue is of a different modality than the target. Endogenous spatial orienting in response to predictive cues has been shown to influence localization responses, but also that spatial orienting elicited by uninformative spatial auditory cues can produce validity effects on localization responses, i.e., when the cue was presented at a different location as the target,

detection performance declines (Spence & Driver, 1994). Therefore, the validity of the cue is very important.

Since the targets are visual, also at issue is whether auditory cues can successfully orient attention to a visual target. Spence and Driver (1997) reported that whereas visual exogenous attention tends to follow auditory exogenous attention around, the reverse dependency apparently does not apply, or applies negligibly (note that vision has a limited field of view while audition is omnidirectional). These findings add to the growing list of qualitative differences between exogenous and endogenous covert orienting (Spence & Driver, 1996, 1997). Further, Ferlazzo, Couyoumdjian, Padovani, and Belardinelli (2002) showed that auditory and visual spatial attention systems are separate, as far as endogenous orienting is concerned. Schmitt, Postma, and DeHaan (2000) found that it is important whether the attention system is activated directly (within a modality) or indirectly (between modalities). Others have found that visual cues affect both visual and auditory localization, but auditory cues only affect auditory localization (Ward, 1994; Ward, McDonald, & Lin, 2000). However, Spence and Driver (1997) found that auditory cues affected both visual and auditory target localization, whereas no sign of auditory orienting was found when visual cues were used. There still remains some doubt regarding the efficacy of auditory cueing, especially during various applied situations.

Haptic. Cueing attention has also been shown to be successful with the sense of touch. When vision is first oriented to the body site receiving the tactile stimulation, tactile localization is facilitated (Lloyd, Bolanowski, Howard, & McGlone, 1999). This research describes improvements in tactile target acquisition by visual cues. However, targets from other modalities can be improved by tactile cues (Kennet, Eimer, Spence, & Driver, 2001). Specifically, links in spatial attention from touch to vision can affect early stages of visual processing (Eimer & Driver, 2000).

Visual. The most obvious and logical method to cue attention to a spatial location is through the visual modality. It is generally accepted that there are two main visual pathways that provide distinct information to humans and primates; the ventral “what” pathway and the dorsal “where” pathway (Niebur & Koch, 1996). Since vision is the primary method of determining the identity and location of a target, then cueing using this modality is highly ecologically valid. The purpose of a visual cue is to reduce the amount of parallel information and make the important stimuli salient. The so-called ‘feature integration theory’ explains how vision is broken down into a set of topographic feature maps (Treisman & Gelade, 1980). Within each map, different spatial locations compete for attention, which then feed into a master “saliency map”, which codes for conspicuity over the entire visual field (Itti, Koch, & Niebur, 1998). Visual cues that are similar to the targets based on color and location have been shown to improve localization performance (Ansorge & Heumann, 2003). Pratt and McAuliffe (2002) described this as an inclusive rule as opposed to an exclusive rule. In other words, the attention system orients attention to stimuli that shares similar features to the targets, as opposed to a system that does not orient attention to stimuli that does not share similar features. This means that the attention system actively seeks out saliency; specifically, a stimulus attracts attention first from a bottom-up perspective and does not actively ignore from a bottom-up perspective. This is important from a design standpoint since a warning or alarm should not give many false positives since

ignoring is a higher cognitive process and thus requires more resources. If this were to take place, missing a valid warning would more likely occur.

Cue Reliability/Trust

The location cueing system could be limited by the current technology, i.e., limits of the processors, sensors, and software may create a less than perfect warning system that fails to detect a target or misinterprets a non-target as a target. Imperfect and unreliable information can create costs with the use of such a system. For example, if a non-target is incorrectly detected as a target and presents the cue to the user, the user would incorrectly focus attention on that location at the cost of other locations where the true target is located (Mosier & Skitka, 1996; Yeh & Wickens, 2000). Yeh and Wickens (2000) studied 'attention bias' in which operator focuses attention to an area highlighted by the automation at the expense of other areas of the visual scene, and 'trust bias' in which unwarranted attention is given to the guidance information. Differences in accuracy between valid and invalidly cued targets could be attributed to differences in allocation of attentional resources (Luck et al., 1994). Alternatively, these results could be explained by a reduction in uncertainty about target location (Luck et al., 1996). Acknowledgement is made that the reliability of the warning information is important, however, in order to explore the phenomena of interest in the current work, accurate and reliable warnings will be assumed possible and simulated in order to partial out any attention biases and trust biases that may exist.

Workload

Workload is an important factor, especially with the applied aspects of the proposed attentional cueing aid. The cueing system may be used in various types of situations that vary in the levels of workload experienced. For example, the system may be used in an automobile driving through a rural road with very few road hazards to warn, or may be used in an urban setting with a number of other vehicles and potential hazards. Similarly, the Soldier using such a system may be in a relatively benign environment with few combatants, or may be under heavy attack. Different levels of workload may affect the way the display characteristics aid performance on target detection. Thus, different levels of workload must be included with the current research exploring exogenous spatial cues. Although Jonides (1976) found only endogenous cueing was affected by workload, exogenous cueing still may be affected by workload in the setting employed in the current work, i.e., even though the cue is data driven and reflexive, it still might require cognitive resources to remap the various sensory locations (the location on the body or the sound in the ear) to the visual-spatial location.

Multiple Resource Theory

Vision is generally accepted as the primary modality for taking in information. Many real-world tasks are inherently dual- and multi-tasks involving multiple sensory modalities, but mostly are a combination of visual and auditory. In order to predict and explain differences in performance under high workload, multi-task environments, Wickens (1980; 2002) developed the multiple resource theory. This theory explains that there is greater interference in task performance when tasks share stages, codes, channels of visual information, and sensory modalities. For example, when providing navigation information to a driver of an automobile, the modality that the information is delivered is very important, especially when the driver's

workload is taxed (e.g., when driving fast or during times of high traffic). Since most of the important driving information is visual, providing more visual information for the navigation task could cause a decline in driving performance since both tasks would be competing for the same attentional resources. In-car navigation systems with visual, moving map displays are one example of how navigation information could interfere with the driving task. The multiple resource theory could be used to predict operationally meaningful interference between the driving task and the navigation task, which could not easily be explained by simpler models of human information processing such as 'bottleneck' or 'filter' theory (Broadbent 1958). Although the multiple resource model only includes visual and auditory modalities, this model still has applicability to multi-tasks incorporating other modalities.

Breadth of Attention

The searchlight metaphor has been used to describe perceptual attention (Wachtel, 1967 as described in Wickens, 1992). The searchlight beam shows the current, momentary direction of attention, and the focus of the beam falls upon that which is in consciousness. This metaphor explains some attention limits well. It explains why and how the brain controls the beam and moves it around, and also that there is a limited number of objects humans can illuminate at once (i.e., have in consciousness and process).

The limited resource metaphor is another way that attention is viewed (Wickens, 1992). Different tasks require many different mental operations, and the performance of each of these mental operations depends on the amount of limited resources the individual has to spend. This view explains the problems with time sharing (attempting to do two activities at once) because two activities demand more resources than one. It also explains why some combinations of activities can be performed well together since they are drawing from different resource pools (visual attention vs. auditory attention). Taken together, these metaphors can be used to explain the problems of attention. Limited resources and the breadth of the searchlight show that it would be impossible to process information from sources that are physically far apart without missing incoming information and/or causing primary task performance errors. The eye has a limited field of view (about 60°) within which it can take in information and an even more restricted region of foveal vision where fine details can be seen and processed (e.g., text and icons) (Wickens, 1992). Attention is also characteristic of mental effort, which is synonymous with mental workload.

Posner (1980) also used the metaphor of the searchlight to describe how attention works. The searchlight scans the environment and those things that fall within the searchlight are the things that are aligned within the attentional system and are more likely to be noticed. Other researchers have proposed that the searchlight can be adjusted so that the 'beam' is narrow or wide, including less or more of the environment within it respectively (Jonides, 1980; Crick, 1984; Johnston & Heinz, 1978). The current work will explore this idea of an adjustable searchlight to see if different breadths of attention can be cued using each of the modality conditions. This will be done by cueing various amounts of the visual search scene with each modality type, as described in the Methods section.

Augmented Reality

Augmented reality (AR) is very much like virtual reality (VR) in that they both use new computer driven technologies that deliver sensory information to the user that is intended to replace the real-world sensations. The main difference, however, is that VR attempts to completely replace the real-world and shield any real-world information from the user, while AR attempts to blend the two into a real-virtual mixture. The computer generated sensory stimuli is predominantly visual, however these AR systems can also augment audition, somatosensory, and even olfaction.

Visual displays in AR are of two types: optical-based and video-based (Barfield & Caudell, 2001). Optical-based displays use lenses that allow light to pass through to the individual's eyes so the actual object is seen as it normally would as if viewed through glasses. However, the lens also has the ability to 'superimpose' computer generated images over the real-world images via reflection onto the lens from a small display. These images thus appear as more of a watermark, since they cannot fully occlude the real world image. Also the real-world images are typically reduced in brightness so the computer generated images can be seen more clearly. Video-based displays, on the other hand, completely regenerate the real-world images and do not deliver the light from the objects to the eyes of the user. The images are captured via cameras and are recreated on small displays directly in front of the user's eyes. This allows the computer to seamlessly blend the computer generated images onto the real-world. The benefit of these video-based displays over optical-based displays is that the images can be seen clearly and fully occlude the real-world images; the draw back is that the display must be refreshed, just as a computer monitor or television does, and can cause negative side effects like eye-strain, headache, or other simulator sickness symptoms.

Advanced technologies such as AR provide many benefits to automated systems intended to provide information to the user that might normally be problematic. For example, AR interacts with human abilities to benefit manufacturing and maintenance tasks, reducing the potential for errors, enhancing motivation, and providing concurrent training (Neumann & Majoros, 1998). Automobile drivers may benefit from an AR product called INSTAR which enables the driver to see a transparent floating arrow that informs the driver where to turn en route to the desired destination. It provides information subtly without impairing the driver's view (Rheingold, 2004). From a military standpoint, this type of system could be a great benefit, especially to the dismounted Soldier. Soldiers could be cued to the positions of enemy snipers who had been spotted by unmanned reconnaissance planes (Feiner, 2004). Logically, this would give the Soldier an advantage; providing advance notice to the location of potential targets may reduce the time required to ascertain the threat and the reaction time required to take action. Technologies are being added into what the US Army calls the Future Combat System (FCS), and AR might prove to be a valuable addition to such a system. The interested reader should refer to Weiner (2004) and Barfield and Caudell (2001) for nice reviews of AR technology, what it is, what it can do, and examples of uses.

Performing pilot research to the current work, Jerome, Witmer, and Mouloua (2005) set out to see how well people can locate a visual or auditory cue in a 360-deg mocked-up urban setting using AR cues. The speed and accuracy of finding targets were compared for three cueing conditions: Audio cues only, Visual cues only, Audio and Visual Cues combined. The cues were superimposed on the real world urban setting mockup. Each participant judged the spatial

location of audio and/or visual cues located to the participant's front, side, or back of the Mixed Reality – Military Operations in Urban Terrain (MR MOUT) simulator. Participants pulled the trigger on the simulated weapon when they located the target cue, causing the cue to disappear. The time of the trigger pull was recorded as a measure of the speed of acquiring that target. Immediately after pulling the trigger, participants indicated the cue location by calling out the correct lettered location from among 24 potential locations within 10 seconds of the start of a trial. There were 12 lettered cue locations and 12 lettered distracter locations. If participants were unable to precisely locate the target, they were required to provide their best guess about the target's location. Accuracy was assessed by determining the number of cues correctly located and by measuring the amount of error for incorrectly identified cues.

Results indicated that simple visual and audio cues presented separately or in combination help in target localization. Visual cues pinpointed the target location better than audio cues but did not significantly improve target acquisition speed beyond that provided by audio cues. Combining audio and visual cues improved both the speed and accuracy of locating targets. While the availability of audio cues helped in locating targets in both high and low positions, it helped more for low position targets. The availability of audio cues helped most when targets were not in the immediate line of sight (to the side or behind the participant). In summary, this research demonstrated the power of combining the unique advantages of visual and auditory cues for aiding target acquisition in an augmented reality environment.

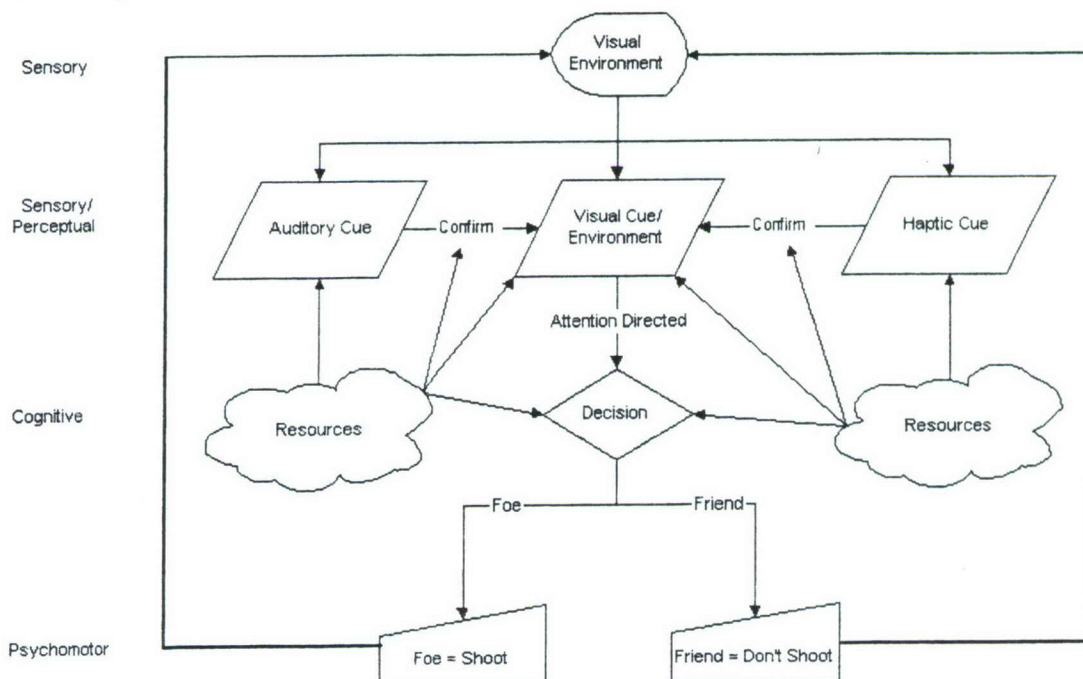


Figure 4. Model representing human systems involved in the experimental tasks.

Rationale for Present Research

AR shows great promise as a potential hazard detection aid in transportation and the military. Baby carriages and children could wear transponders which send a signal to a receiver equipped in a vehicle. This receiver could then warn the driver when approaching the children if they are in their immediate pathway. The U.S. military can benefit from such a device by

providing information about enemy combatant locations collected from unmanned reconnaissance or remote sensors. The information given the automobile driver or the Soldier can increase the likelihood of detecting a hazard or providing advance notice that may provide more time to the individual to react to the situation, increasing the chances of successful countermeasures in time.

Along with the benefits, however, the costs must also be considered. A cueing system that provides hazard location information using different modalities or combinations of modalities may aid or hinder performance. The methods of information delivery thus should be explored and determined prior to system use in order to avoid a poor design, causing more problems than benefits. Although driven by the technology, engaging in user testing in order to make the design more user centered is desirable. The increased safety alone warrants effort into these issues; however, increasing user satisfaction during tasks is important as well. The current work is similar to previous studies, but goes beyond what they have done by using AR to deliver the attentional cues in a 360-degree view of the world, as well as looking at an applied military-like task. The results of such testing could potentially impact the development of an AR system used to warn against potential hazards in the 360-degree environment around the individual.

Research Hypotheses and Predictions

The goal of this research is to show how to best provide hazard location information to a Soldier in the most efficient way, with the presentation of the information, the perception of that information, the decision to act or not act, and the accuracy and speed of such actions. Based on the Army's interests for applications of the cueing system, we will focus only on overt orienting of attention, i.e., with eye, head, and body movements towards the cue and target since overt shifts of attention are how one generally orients attention. Covert orienting of attention is not of interest since it almost always precedes an overt orientation. Also, we are only interested in visual targets and the effects of different cueing modalities since a target in this case is an enemy or hazard that must be confirmed visually before a decision to act is made. Per Broadbent's (1958) and Wickens' (1992) models of Information Processing, since the attentional cue is hypothesized to free some attentional resources, in general target detection performance will benefit from attentional cues. More specifically, since the tactile modality is utilized less than visual or auditory in the experimental task, there are more free resources for this modality. However, since the visual cues are fully exogenous cues, and the haptic and auditory cue are partly endogenous cues, more cognitive resources are required to 'remap' the information from the haptic and auditory cue onto the visual environment. Also, since the visual modality is necessary and sufficient to identify the target and confirm the information provided by the cue, the visual modality should also have advantages over the other modalities (see Figure 4). Due to the characteristics of the experimental task, especially the lack of control of where the participant will be looking when each of the targets will be displayed, it can be assumed with confidence that more targets will appear outside of the field of view than in (especially since the head mounted display (HMD), has a limited field of view and can only display a very limited number of potential target locations at one time). Therefore, the visual cues when used alone will have a great disadvantage. However, combining visual cues with other modality cues will be beneficial. In summary, there are three factors affecting the performance of each of these cues: a) attentional resources, b) endogenous/exogenous cueing, and c) location of targets outside field of view. In general, the addition of any of the cues is expected to improve performance. How will

each modality contribute to the benefits of attentional cueing? The haptic and auditory cue will be sensed immediately upon presentation whereas the visual cue might not (when out of field of view). However, the cost for the cognitive remapping of the spatial information upon the visual scene will also lead to a considerable disadvantage. One last notable issue is that the nature of auditory cues makes them slightly more difficult to localize than the haptic cues. Taking all this into consideration, the specific hypotheses are as follows:

- Hypothesis 1: Cueing modality will affect target detection.
 - Prediction 1.1: Each cue modality will significantly improve performance over the absence of the cue.
 - Tested by MANOVA F-test to determine significance of contribution of each modality to target detection accuracy and reaction time performance.
 - Prediction 1.2: Accuracy and Reaction time performance will benefit from the presence of Haptic and Visual cues together more than Auditory and Visual cues together, this combination would be better than visual cues alone, visual cues would be better than haptic cues, haptic cues would be better than auditory cues, and auditory cues would be better than no cues at all (i.e., $\text{Haptic/Visual} > \text{Auditory/Visual} > \text{Visual} > \text{Haptic} > \text{Auditory} > \text{Control}$). Two groups are intentionally left out of these planned comparisons: a) Haptic/Visual/Auditory, since the combination of all three is difficult to predict to be the most beneficial since nothing is known about how much information is too much, nor what interactions might affect the performance of tasks; and b) Auditory/Haptic, since the combination of two exogenous cues, without an endogenous cue is not expected to improve performance over either cue alone.
 - Tested by planned comparison independent t-tests to determine which groups are significantly better than the others.
 - Prediction 1.3: Each cue modality will significantly reduce fratricide (i.e. shooting non-combatants/civilians) occurrence over the absence of the aid of the cue.
 - Tested by MANOVA F-test to determine significant contribution of each modality to target detection accuracy and reaction time performance.
 - Prediction 1.4: Fratricide occurrence will be reduced with the presence of Haptic and Visual cues together more than Auditory and Visual cues together, this combination will be better than visual cues alone, visual cues will be better than haptic cues, haptic cues will be better than auditory cues, and auditory cues will be better than no cues at all (i.e., $(\text{i.e. Haptic/Visual} < \text{Auditory/Visual} < \text{Visual} < \text{Haptic} < \text{Auditory} < \text{Control})$).
 - Tested by MANOVA F-test to determine significance of contribution of each modality to fratricide occurrence.
 - Prediction 1.5: Workload will interact with cueing modality. Per Treisman's Feature Integration Theory, early selection of attention will occur during times of high workload, and late selection will occur during times of low workload.

However, the degree of early or late selection will also be determined by the cueing modality.

- Prediction 1.6: Workload will affect visual performance more than auditory, and will affect auditory more than haptic.
 - Tested by MANOVA looking to see if there is an interaction between workload and cue modality.
- Hypothesis 2: Cue specificity will affect target detection.
 - Prediction 2.1: The levels of performance across increasing cue specificity will be an inverted U function.
 - Prediction 2.2: Higher workload will lead to narrowing of attention, and thus smaller and larger cue specificity will be hindered over medium cue specificities. The inverted U function will be moderated by workload.

Method

Participants

Prior to recruiting any participants and collecting data, a power analysis was performed to estimate the number of participants needed to obtain sufficient statistical power. Based on the power analysis results, 64 University of Central Florida students were recruited to participate in this research. There were 30 females and 34 males, ranging in age from 18 to 34 years ($M = 20.39$, $SD = 3.49$). All of the participants volunteered to participate and were treated in accordance with the principles of ethical treatment of human research participants (American Psychological Association, 1992). All participants self-reported to have unaided or corrected 20/20 vision and normal color vision.

Apparatus

Augmented Reality Simulator. MR MOUT was a physical mock-up of urban 2-story buildings and chroma-key portals which aid in the display of computer generated environmental objects including walls, rooms, tables, and even mobile entities like civilians and enemies. The building façades enclosed a rectangular area of approximately 30 square meters. The MR MOUT simulator was run by proprietary software that controlled the presentation of the visual, audio, and tactile stimuli. The real world video feed was taken in by the Canon HMD (described below), processed by the computer and software, combined with the computer generated augmented information, and the resultant output was fed back to the HMD imaging elements in real time, and was synchronized with the audio and tactile stimuli to generate a multi-modal environment for the user that seamlessly combined the real and the virtual.

Video Display. A Canon VH-2002 video see-through AR HMD was used to view the MOUT environment. It included a VGA display, 640x480 pixels at 60Hz. The display size was $H51^\circ \times V37^\circ$. The space between the eyes was 63mm, with a convergence position of 2m. The HMD (minus the cables) weighed 325g, and the transmission box weighed 290g. The

transmission box and some of the devices for the tracking system were housed inside a cargo vest that the participants wore fitting loosely outside the haptic vest (described below).

Audio Display. Audio was presented through strategically placed speakers via a 2-tier surround sound system. The computer sound card used was a Creative Labs Audigy 2NX. The program that controlled the audio output was called ASIO4ALL, which allows each output channel to be addressed individual by the mixed reality sound engine. The speakers were Yamaha model MSP3, which are 30W 60Hz. Sixteen speakers were used positioned at heights of 1.37m and 4.57m.

Haptic Display. The haptic vest was constructed from a drysuit retro-fitted with hook-and-loop fasteners so it could be adjusted to fit the size of the participants (see Figure 30). Thirty-two vibro-tactors (manufactured by K'OTL, Model 6DL-05WA, speed 8000 +/- 1500 rpm) were incorporated into the vest at 8 zones. Two zones were located at the upper chest, 2 at the lower abdomen, 2 at the upper back, and 2 at the lower back. Vibrations were applied to the part of the torso corresponding to the external location, relative to the participant, where the target is located. These spatial orienting haptic cues were dynamic, i.e., the user's position and orientation were tracked and the cue followed the participant's movements.

Tracking Device. The Intersense IS-900 VETracker was used to track user position and head movement, as well as aiming direction of a simulated weapon. The IS-900 uses a combination of 2 tracking technologies; inertia position and ultrasound. The inertial position tracking uses gyroscopes and accelerometers to sense the positional changes in the sensors and delivers high update rates (Kindratenko, 2001). The ultrasound component is responsible for keeping the inertial module from drifting. The ultrasound transmitters are housed in 6 tracking bar and send out a 40 kHz pulse that is picked up by receivers on the HMD and the simulated weapon.

Input Device. The input device was a simulated M-4 carbine. The M-4 is a compact version of the M-16 rifle with a collapsible stock. It looked and felt nearly identical to an actual M-4 except that it was retro-fitted with a grenade launcher mounted under the barrel and had the wireless tracking device mounted to the side of the magazine. It also had a wireless transmitter which attached the rifle trigger to the mouse button on one of the computers. This enabled the computer to automatically record the location and time of each trigger pull.

Computers. Three separate, but identical, desktop computers rendered the audio, visual, and tactile stimuli for the MR MOUT environment. Each had 2.8 GHz Xeon processors, 1 GB of RAM, and had NVIDIA Quadro4 900 XGL video cards. Windows XP Professional with service pack 2 was the operating system for one of the computers. This computer was responsible for the story engine, the sound engine, and the physics engines (path planning and ray casting to determine what was shot). It also ran the GUI software used by the experimenter to control the starting and stopping of the scenario, as well as to choose which experimental conditions would be selected. The other two computers used Red Hat Linux as the operating system. The first Linux machine controlled the haptic vest, the sensor server, and the observer views (capture and render). The other Linux computer was responsible for HMD video capture, the graphics engine (render) for the user's view.

Research Methodology

The 64 participants completed the cued attention task in a one hour session. All participants completed an informed consent, a short demographic questionnaire and a simulator sickness questionnaire prior to any experimental tests. Following the tests, participants again completed a simulator sickness questionnaire and a presence questionnaire. Participants were monitored for simulator sickness symptoms during the testing. The research design for the test is described below.

Cues. The visual cue was a yellow and black box surrounding the location of the target. The auditory cue was 6 blasts of pink noise presented from the location of where the target was located via the surround sound system. Tactile cues were vibrations using 8 vibro-tactors whose location corresponded to the environmental location in relation to where the vibrations occurred on the body (e.g., upper right chest vibration corresponded to the right hand side of the second story of the building directly in front of the participant).

Design. The experiment used a $2 \times 2 \times 2 \times 4 \times 2$ mixed factorial MANOVA design to determine the effects of cue modality on attention directing. The between subjects variables were visual modality (present or absent), auditory modality (audio present or absent), and tactile modality (haptics present or absent); the within subjects variables were cue specificity (none, wide, medium, and narrow) and workload (high and low). This design provided the tests of main effects of cue presence or absence for each of the three modalities, workload, and cue size, and their interaction with the dependent variables (Tabachnick & Fidell, 2001). A sample of 64 participants was randomly assigned to one of 8 groups representing the 8 modality combinations. All participants interacted with targets using a simulated rifle and responded by either engaging or not engaging potential targets during judgmental use force (shoot/don't shoot) scenarios. The major dependent variables were measures of combat proficiency, such as: a) reaction time of the target kills (in seconds) and b) accuracy of targets killed (hit/miss ratio). The Secondary Task involved street lights coming on at various times during the scenario and remaining on for 5 seconds or until extinguished. Participants were instructed to extinguish the lights as soon as possible by shooting them out. Lights came on at 8 different times (not equally spaced in time) during both the high workload and low workload parts of the scenario (see Appendix D). The NASA TLX (described below) and the speed and accuracy of extinguishing lights in the secondary task were used to assess the workload experienced in the two scenarios.

Table 1
Design of Cued Attention Experiment

Between Ss Factor						Within Ss Factors					
Cue Modality											
Auditory (2 levels)		Haptic (2 levels)		Visual (2 levels)		Cue Specificity (4 levels)			Workload (2 levels)		
Absent	Present	Absent	Present	Absent	Present	None	Wide	Medium	Narrow	High	Low

Procedure. All data collection occurred at the SFC Paul Ray Smith Simulation and Training Center located in the Central Florida Research Park in Orlando, Florida. After being briefed on the purpose of the research, the tasks for the experiment, and any risks involved, participants read and signed a consent form. They were randomly assigned to an experimental condition, which grouped participants as to what attentional cue type(s) they would receive. Participants also were assigned a participant number so all data would remain anonymous. They then completed demographic and baseline simulator sickness questionnaires. Following these preliminaries, each subject performed a target detection task which required them to identify the target as enemy or civilian, and to engage the enemy (shoot them) while identifying the civilians without harming them (fratricide avoidance). Following these practice trials, each participant searched for enemies and tried to shoot them while avoiding shooting the civilians during each of four 2-minute scenarios. Two scenarios were fast-paced, high intensity scenarios (16 possible targets in 2-minutes) while the other 2 were slow-paced, low intensity scenarios (8 possible targets in 2-minutes). The scenarios were delivered staggered; half the subjects got a slow scenario first while the other half got a fast scenario first (i.e., slow, fast, slow, fast; or fast, slow, fast, slow). The 2 scenario types were expected to differ in the workload that they provide.

Independent Variables

Between Subjects. The between subjects variable was Cue Modality, which had 6 levels: Auditory (absent vs. present) vs. Haptic (absent vs. present) vs. Visual (absent vs. present)

Within Subjects. The within subjects variables were Cue Specificity and Workload. Cue specificity had 4 levels: non-spatial vs. wide vs. medium vs. narrow. Workload had 2 levels: high vs. low

Dependent Variables

Task Performance. Two measures of combat proficiency were used to evaluate task performance including: a) accuracy measured by hit/miss ratio and b) reaction time of the target kills.

Secondary Task. Street lights came on at various times during the scenario and remained on 5 seconds or until extinguished. Participants were instructed to extinguish the lights as soon as possible by shooting them out. Lights came on at 8 different times (not equally spaced in time) during both the High Workload and Low Workload Scenario.

Questionnaires. Immersive Tendencies, Presence, Simulator Sickness, Demographics, NASA-TLX subjective workload. Immersive Tendencies was measured using Witmer & Singer's (1998) Immersive Tendency Questionnaire (ITQ). The ITQ contains three subscales, (a) involvement, (b) focus, and (c) propensity to play and enjoy video games. Presence was measured using the Presence Questionnaire (PQ), which is a 28-item questionnaire using a seven-point scale format based upon the semantic differential principle where the ends of the scale are anchored by opposing descriptors, but has a mid-point descriptor as well (Witmer & Singer, 1994; 1998). It is composed of 4 groups of conceptually similar items, which include (a) involvement (10 items), (b) sensory fidelity (8 items), (c) adaptation/immersion (7 items), and (d) interface control (3 items). Simulator sickness was measured using the Simulator Sickness Questionnaire (SSQ), which is a 16-item questionnaire that is a shortened version of the Motion Sickness Questionnaire (Kellogg, Kennedy, & Graybiel, 1965) where 12 items were deleted that were inappropriate for measuring simulator sickness. It summarizes 3 distinct symptom clusters, including (a) nausea (stomach awareness, increased salivation, burping), (b) oculomotor (eye strain, headache, blurred vision, difficulty focusing), and (c) disorientation (dizziness, vertigo) (Kennedy, Lane, Berbaum, & Lilienthal, 1993). The NASA-TLX is a multidimensional rating scale in which information about the magnitude and sources of six workload-related factors are combined to derive a sensitive and reliable estimate of workload (Hart & Staveland, 1988). These factors include mental demand, physical demand, temporal demand, performance, effort and frustration. The scale is presented in pair combinations of these factors where the subject is asked to select which factor in the combination is the most relevant to the task in hand. This is followed by a subjective scale for each factor in which the subject rates each individual demand. The combination of these two measures is scored and provides a good estimate of mental workload.

Results

Objective and subjective data were collected for each participant as a function of their performance on the tasks in the augmented environment. The data were analyzed using the Statistical Package for the Social Sciences, SPSS® version 12. A MANOVA for mixed-factorial designs looking at the 2 x 2 x 2 x 4 x 2 factors was performed, as well as follow up t-tests for planned comparisons.

Descriptive Statistics

Data screening of all data collected indicated that the distributions of four dependent variables were non-normal. The univariate analysis showed that three were skewed, and there was an outlier for the other. Per the procedure recommended by Tabachnick and Fidell (2001), the next most extreme score to the outlying case was identified, and this score was used in place of the outlier's score, but changed by one unit away from the mean. All variables were then analyzed from a multivariate perspective to see if the outliers are less extreme from within each experimental group. The only non-normal variable was Fratricide occurrence, which was positively skewed ($|z| = 2.31$). Since this is only marginally skewed and may be a true representation of the distribution in the population (i.e., floor effect for fratricide—does not occur very often), nothing further was done to adjust. Table 2 summarizes the overall means and standard deviations for the dependent variables measured.

Table 2
Descriptive Statistics for Dependent Measures

Measure	<i>N</i>	<i>M</i>	<i>SD</i>
Accuracy (% Engaged)	64	.39	.13
Reaction Time (in s, 5 s max.)	64	4.36	.23
Light Accuracy (% Engaged)	64	.09	.08
Light Reaction Time (in s, 5 s max.)	64	4.94	.06
Fratricide Occurrence (% Engaged)	64	.03	.03
NASA TLX Subjective Workload	64	73.49	9.95
Presence	64	119.80	22.07
Simulator Sickness	64	24.84	26.01
Immersive Tendencies	64	66.19	13.32

Tests of Specific Hypotheses

Hypothesis 1 stated that *cueing modality would affect target detection*. The first prediction for this hypothesis states that each cue modality would significantly improve performance over the absence of the aid of the cue, i.e. presence of haptic would be better than absence of haptic; presence of visual would be better than absence of visual; and presence of auditory would be better than absence of auditory.

Table 3 summarizes the descriptive statistics for accuracy and reaction time which both are used to measure different aspects of target detection performance. A 2 x 2 x 2 x 4 x 2 mixed-factorial MANOVA was performed to measure the effects of the presence of the three cue modality types on the dependent variables. Participants benefited from the presence of the haptic cues for reaction time, $F(1, 56) = 28.38, p < .01$, and for accuracy, $F(1, 56) = 40.87, p < .01$. All others were non-significant. Figures 5 and 6 illustrate these results.

Table 3

Descriptive Statistics for Accuracy and Reaction Time Measures for Each Condition

	Visual	Audio	Haptic	<i>M</i>	<i>SD</i>	<i>N</i>
Accuracy	0	0	0	0.26	0.07	8
	0	0	1	0.44	0.13	8
	0	1	0	0.36	0.07	8
	0	1	1	0.50	0.12	8
	1	0	0	0.31	0.10	8
	1	0	1	0.43	0.09	8
	1	1	0	0.32	0.13	8
	1	1	1	0.48	0.10	8
Reaction Time	0	0	0	4.55	0.14	8
	0	0	1	4.26	0.25	8
	0	1	0	4.42	0.14	8
	0	1	1	4.16	0.25	8
	1	0	0	4.47	0.18	8
	1	0	1	4.28	0.19	8
	1	1	0	4.51	0.20	8
	1	1	1	4.23	0.20	8

Note: 0 = absence, 1 = presence of each modality cue

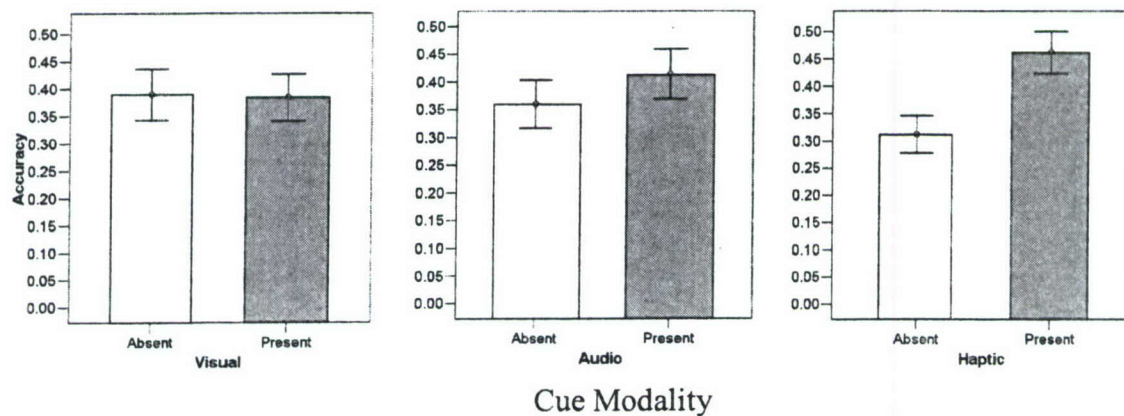


Figure 5. Difference in accuracy for absence and presence of visual, auditory, and haptic cues. Points represent the mean ratio of targets engaged vs. total number of targets; vertical lines depict standard error of the means.

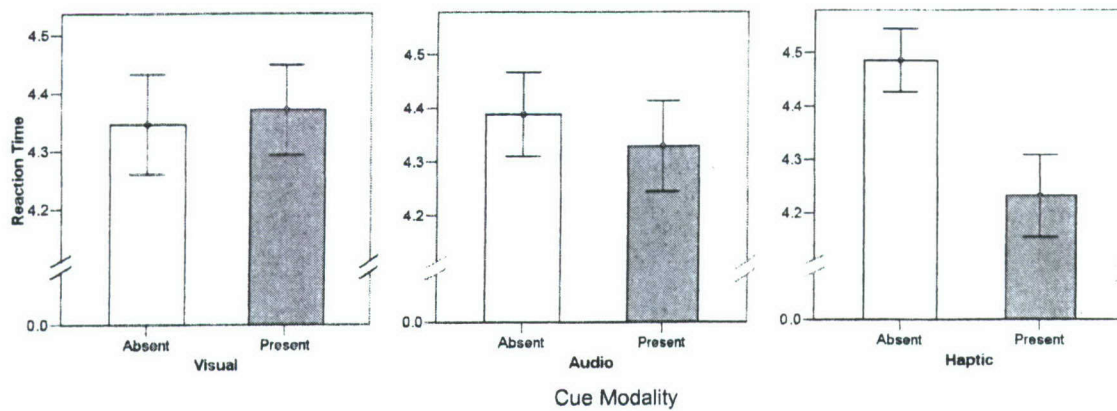


Figure 6. Difference in reaction time for absence and presence of visual, auditory, and haptic cues. Points represent the mean reaction time for targets engaged; vertical lines depict standard error of the means.

The second prediction for this hypothesis states that accuracy and reaction time performance would benefit from the presence of Haptic and Visual cues together more than Auditory and Visual cues together, this combination would be better than visual cues alone, visual cues would be better than haptic cues, haptic cues would be better than auditory cues, and auditory cues would be better than no cues at all (i.e., Haptic/Visual > Auditory/Visual > Visual > Haptic > Auditory > Control). These planned comparisons were carried out testing each one of these predictions with individual t-tests, summarized in Table 4. Figures 7 and 8 summarize the results and shows the expected pattern of performance effects, except for the haptic condition, which showed much higher performance than what was expected. For the Accuracy measure, Haptic/Visual ($M = .43$, $SD = .09$) was significantly better than Audio/Visual ($M = .32$, $SD = .13$), $t(42) = 2.31$, $p < .05$ (predicted); Haptic ($M = .44$, $SD = .13$) was significantly better than Visual ($M = .31$, $SD = .10$), $t(42) = -2.57$, $p < .01$ (not predicted); and Audio ($M = .36$, $SD = .07$) was significantly better than no cues ($M = .26$, $SD = .07$), $t(42) = 1.99$, $p < .05$ (predicted). For the Reaction Time measure, Haptic/Visual ($M = 4.28$, $SD = .19$) was significantly faster than Audio/Visual ($M = 4.51$, $SD = .20$), $t(42) = -2.57$, $p < .05$ (predicted) and Haptic ($M = 4.26$, $SD = .25$) was significantly faster than Visual ($M = 4.47$, $SD = .18$), $t(42) = 2.19$, $p < .05$ (not predicted).

Table 4
Planned Comparisons for Cue Modality Conditions

	Contrast	Value of Contrast	SE	t	df	p
Accuracy	Haptic/Visual vs. Audio/Visual	0.11	0.05	2.31*	42	.03
	Audio/Visual vs. Visual	0.01	0.05	0.16	42	.88
	Visual vs. Haptic	-0.13	0.05	-2.57**	42	.01
	Haptic vs. Audio	0.08	0.05	1.57	42	.12
	Audio vs. Control	0.10	0.05	1.99*	42	.05
Rxn Time	Haptic/Visual vs. Audio/Visual	-0.23	0.09	-2.47*	42	.02
	Audio/Visual vs. Visual	0.04	0.09	0.45	42	.66
	Visual vs. Haptic	0.21	0.09	2.19*	42	.03
	Haptic vs. Audio	-0.16	0.09	-1.66	42	.10
	Audio vs. Control	-0.13	0.09	-1.39	42	.17

* $p < .05$. ** $p < .01$.

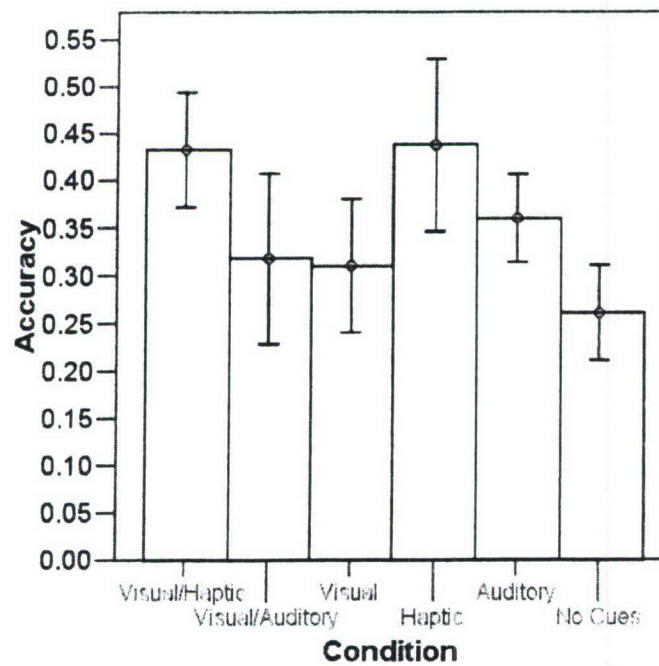


Figure 7. Observed accuracy performance across cue modality conditions. Points represent the mean ratio of targets engaged vs. total number of targets; vertical lines depict standard error of the means.

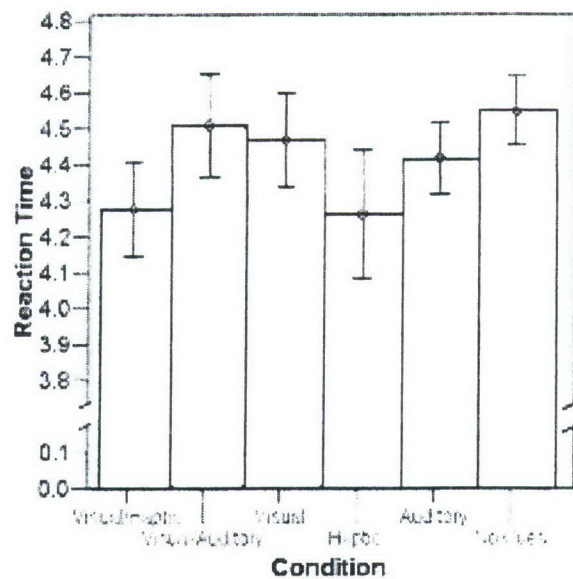


Figure 8. Observed reaction time performance across cue modality conditions. Points represent the mean reaction time for targets engaged; vertical lines depict standard error of the means.

The third prediction for Hypothesis 1 stated that each cue modality would significantly reduce fratricide occurrence over the absence of the aid of the cue (i.e., presence of haptic would be better (lower occurrence) than absence of haptic; presence of visual would be better (lower occurrence) than absence of visual; and presence of auditory would be better (lower occurrence) than absence of auditory).

Table 5 summarizes the descriptive statistics for accuracy (occurrence) of fratricide. A $2 \times 2 \times 2 \times 4 \times 2$ mixed-factorial ANOVA was performed to measure the effects of the presence of the three cue modality types. Participants were hindered by the presence of the audio cues, $F(1, 56) = 12.96, p < .01$. All others were non-significant. Figure 9 illustrate these results. There was also an interaction for visual and haptic cues on accuracy performance, $F(1, 56) = 10.50, p < .01$, $\eta^2 = .16$ (see Figure 10). This indicates that when each cue is presented alone, fratricide occurrence is diminished; however, when both cues are presented together, fratricide occurrence is amplified, even surpassing the level of occurrence with neither cue.

Table 5
Descriptive Statistics for Fratricide Occurrence

Visual	Audio	Haptic	M	SD	N
0	0	0	0.02	0.02	8
0	0	1	0.01	0.02	8
0	1	0	0.05	0.03	8
0	1	1	0.02	0.02	8
1	0	0	0.01	0.02	8
1	0	1	0.02	0.02	8
1	1	0	0.02	0.02	8
1	1	1	0.07	0.05	8

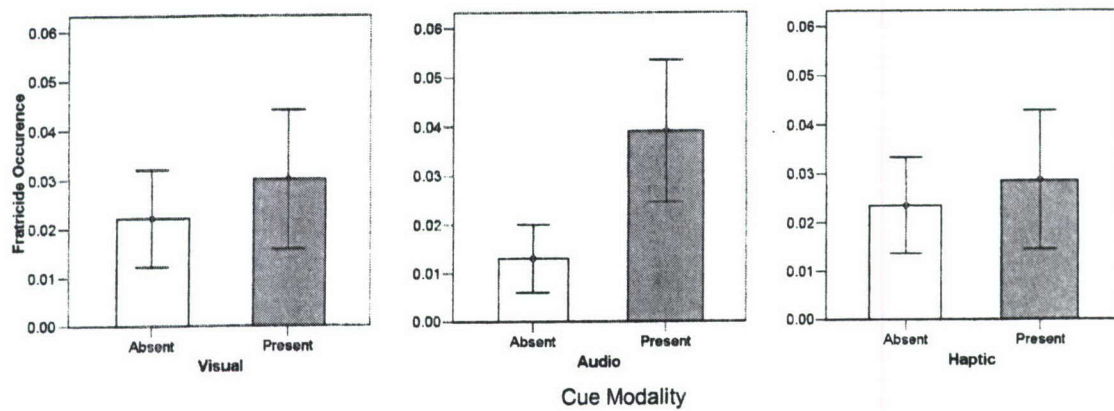


Figure 9. Fratricide occurrence for the absence and presence of visual, auditory, and haptic cues. Points represent the mean ratio of civilian targets engaged vs. total number of civilian targets; vertical lines depict standard error of the means.

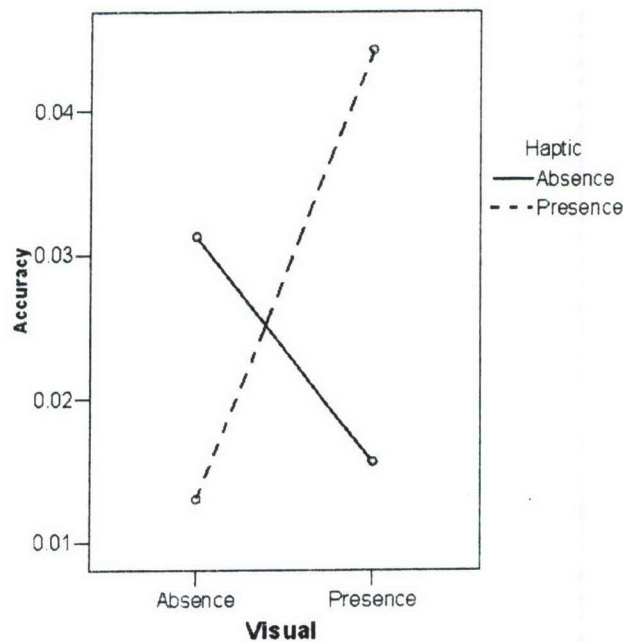


Figure 10. Fratricide occurrence interaction between visual and haptic cues.

The fourth prediction for Hypothesis 1 stated that fratricide occurrence will be reduced with the presence of Haptic and Visual cues together more than Auditory and Visual cues together, this combination will be better than visual cues alone, visual cues will be better than haptic cues, haptic cues will be better than auditory cues, and auditory cues will be better than no cues at all (i.e., Haptic/Visual < Auditory/Visual < Visual < Haptic < Auditory < Control). These planned comparisons were carried out testing each one of these predictions with individual *t*-tests. Figure 11 shows the expected pattern of performance effects, except for the Audio condition, which showed much higher fratricide occurrence than what was expected. Fratricide occurrence was significantly higher for the Audio group ($M = .047$, $SD = .035$) than the Haptic group ($M = .010$, $SD = .019$), $t(42) = -3.07$, $p < 0.01$ (predicted), and significantly higher for the

Audio group ($M = .047$, $SD = .035$) than for the control group ($M = .016$, $SD = .022$), $t(42) = 2.63$, $p < 0.01$ (not predicted). These results indicate that the presence of the audio cue was associated with a greater level of false positive target detections than the haptic cue and the control group which received no cues.

Table 6
Planned Comparisons for Cue Modality Conditions for Fratricide

Contrast	Value of Contrast	SE	<i>t</i>	<i>df</i>	<i>p</i> (2-tailed)
Haptic/Visual vs. Audio/Visual	-0.01	0.01	-0.44	42	0.66
Audio/Visual vs. Visual	0.01	0.01	0.88	42	0.39
Visual vs. Haptic	0.00	0.01	0.00	42	1.00
Visual vs. Audio	0.04	0.01	2.52	42	0.02
Haptic vs. Audio	-0.04	0.01	-3.07**	42	0.00
Audio vs. Control	0.03	0.01	2.63**	42	0.01

** $p < .01$.

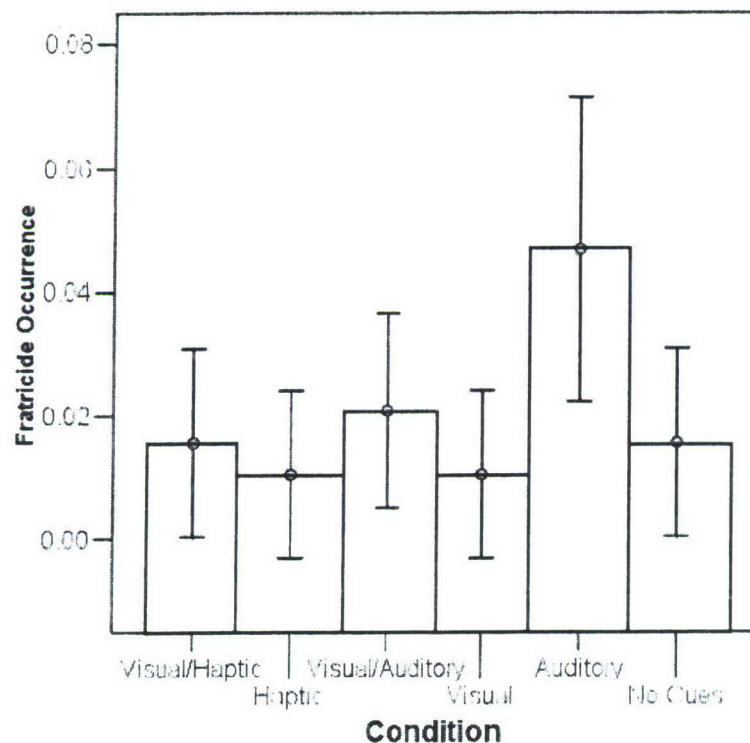


Figure 11. Fratricide occurrence across cue modality conditions. Points represent the mean ratio of civilian targets engaged vs. total number of civilian targets; vertical lines depict standard error of the means.

The fifth and sixth predictions for Hypothesis 1 stated that cueing modality would affect target detection performance; specifically workload will interact with cueing modality, workload affecting visual performance more than auditory, and affecting auditory more than haptic. A 2 x

2 x 2 x 4 x 2 mixed-factorial MANOVA was performed to measure the effects of the presence of the three cue modality types. Participant's target detection performance was significantly affected by the two workload conditions for reaction time, $F(1, 56) = 11.426, p < .001$, and for accuracy, $F(1, 56) = 6.439, p < .05$. All others, including the interactions, were non-significant. Figure 12 illustrates these results. These results indicate that the performance in the low workload condition was worse (lower accuracy and longer reaction time) than the high workload condition (see Table 7).

Table 7
Descriptive Statistics for Workload

Workload	Dependent Variable	<i>M</i>	<i>SD</i>	<i>N</i>
Low	Accuracy	0.35	0.15	64
High	Accuracy	0.40	0.14	64
Low	Reaction Time	4.43	0.28	64
High	Reaction Time	4.32	0.25	64

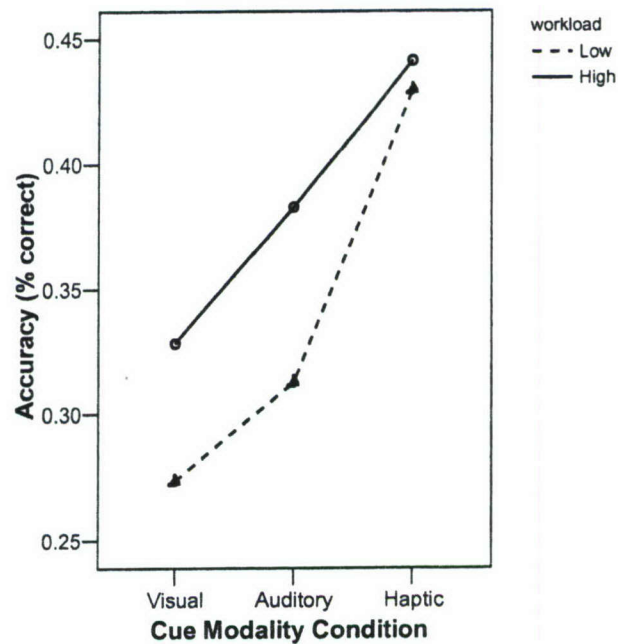


Figure 12. Accuracy performance between cue modalities for low and high workload.

Hypothesis 2 stated that *cue specificity would affect target detection*, specifically, the levels of performance across increasing cue specificity will be an inverted U function. A 2 x 2 x 2 x 4 x 2 mixed-factorial MANOVA was used to measure the effects of the presence of the three cue modality types and the different cue specificities on the performance of the dependent variables. Participant's target detection reaction time performance was significantly affected by the 4 cue sizes, $F(1, 56) = 31.93, p < .01$, as well as their target detection accuracy performance,

$F(1, 56) = 30.25, p < .01$. These results mean that the cue specificities had significant affects on performance; however, planned comparisons must be performed in order to find the cue sizes that provided the significant benefits.

Planned comparisons were carried out comparing each level of cue specificity with each other within each workload group. During low workload, no cues ($M = 4.67, SD = .31$) had greater reaction time than small cues ($M = 4.40, SD = .56$), $F(1, 56) = 12.80, p < .01$, medium cues ($M = 4.20, SD = .58$), $F(1, 56) = 17.67, p < .01$, and large cues ($M = 4.46, SD = .51$), $F(1, 56) = 10.05, p < .01$. Small cues ($M = 4.40, SD = .56$) had greater reaction time than medium cues, ($M = 4.20, SD = .58$), $F(1, 56) = 4.31, p < .05$. Finally, large cues ($M = 4.46, SD = .51$) had greater reaction time than medium cues ($M = 4.20, SD = .58$), $F(1, 56) = 8.56, p < .01$. Overall, these results indicate that the presence of any level of spatial information within the cue benefited target detection reaction time during low workload.

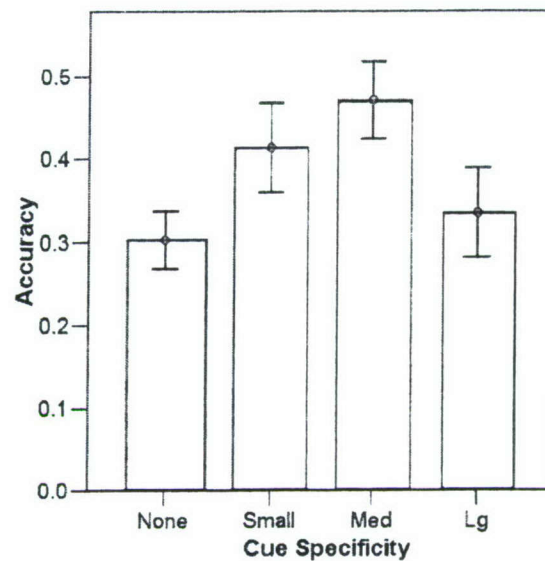


Figure 13. Accuracy of target acquisition for the four levels of cue specificity. Points represent the mean ratio of civilian targets engaged vs. total number of civilian targets; vertical lines depict standard error of the means.

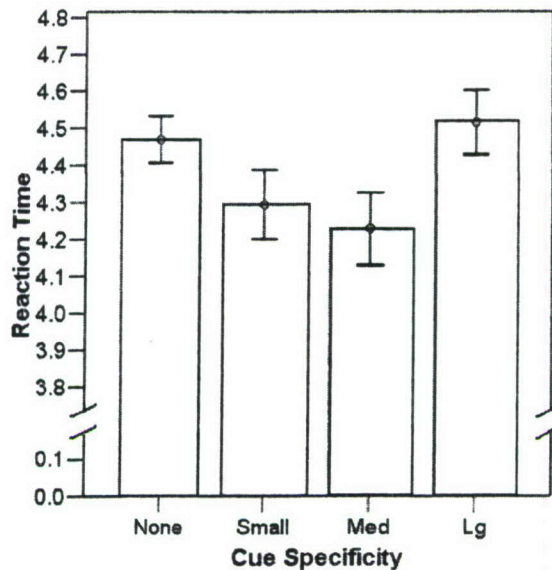


Figure 14. Reaction time of target acquisition for the four levels of cue specificity. Points represent the mean ratio of civilian targets engaged vs. total number of civilian targets; vertical lines depict standard error of the means.

During high workload, large cues ($M = 4.57$, $SD = .38$) had greater reaction time than no cues ($M = 4.27$, $SD = .45$), $F(1, 56) = 16.56$, $p < .01$. Large cues ($M = 4.57$, $SD = .38$) had greater reaction time than small cues ($M = 4.18$, $SD = .40$), $F(1, 56) = 37.03$, $p < .01$. Finally, large cues ($M = 4.57$, $SD = .38$) had greater reaction time than medium cues ($M = 4.26$, $SD = .39$), $F(1, 56) = 27.99$, $p < .01$. Overall, these results indicate that during high workload, the large cue specificity took the longest to detect the target than any other cue specificity.

During low workload, medium cues ($M = .48$, $SD = .25$) had greater accuracy than no cues ($M = .23$, $SD = .20$), $F(1, 56) = 63.62$, $p < .01$, small cues ($M = .39$, $SD = .32$), $F(1, 56) = 13.44$, $p < .01$, and large cues ($M = .33$, $SD = .27$), $F(1, 56) = 17.37$, $p < .01$. Small cues ($M = .39$, $SD = .32$) and large cues ($M = .33$, $SD = .27$) had greater accuracy than no cues ($M = .23$, $SD = .20$), $F(1, 56) = 14.89$, $p < .01$; $F(1, 56) = 8.18$, $p < .01$, respectively. Overall, these results indicate that the presence of any level of spatial information within the cue benefited target detection accuracy during low workload, with the medium sized cue specificity being significantly better than all others.

During high workload, small ($M = .44$, $SD = .20$) and medium cues ($M = .46$, $SD = .20$) had greater accuracy than no cues ($M = .38$, $SD = .19$), $F(1, 56) = 4.53$, $p < .05$; $F(1, 56) = 5.14$, $p < .05$, respectively. Both small cues ($M = .44$, $SD = .20$), $F(1, 56) = 11.73$, $p < .01$ and medium cues ($M = .46$, $SD = .20$) had lower accuracy than large cues ($M = .34$, $SD = .23$), $F(1, 56) = 14.07$, $p < .01$. These results indicate that during high workload, small and medium cues tend to improve target detection accuracy performance over the other cue specificities; however, the interpretation of the interactions below will explain how these findings are contingent upon the levels of other variables.

Other Interesting Findings

Interactions. Reaction time performance was highly correlated with accuracy performance ($r = -.90, p < .01$). Furthermore, the interaction trends for reaction time were identical to that of accuracy, so the interpretation can be used for performance in general, and not specifically to accuracy or reaction time. A significant two-way interaction between cue size and audio cueing was obtained for the measure of reaction time, $F(1, 56) = 4.16, p < .05$, and for accuracy, $F(1, 56) = 6.76, p < .01$. This indicates that a change in performance due to the presence or absence of the audio cue is dependent upon the specificity of that cue (see Figure 15). Performance with the presence and absence of the audio cue is virtually identical, i.e., small and medium cue sizes showed higher performance than no specificity and large cues, but were not affected by the presence or absence of the audio cue. However, the no specificity cues (cue present, but no spatial information) indicated a benefit of the presence of the audio cues over the absence of the audio cues.

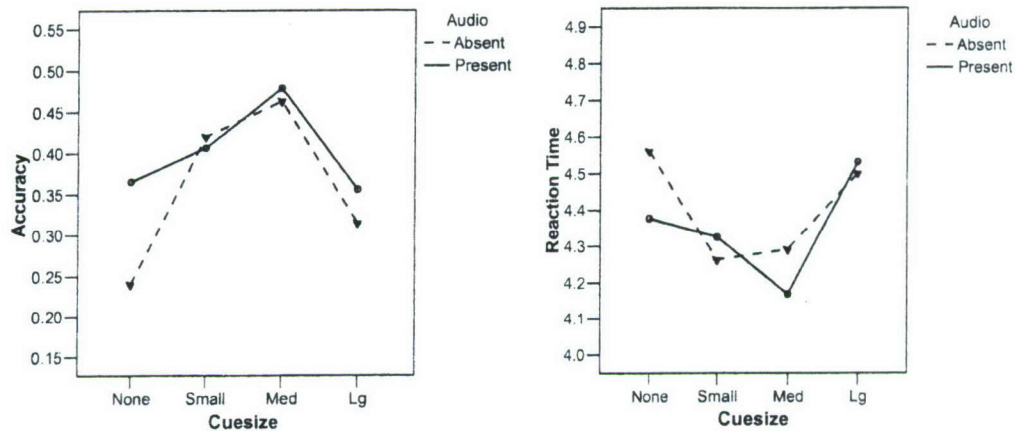


Figure 15. Interaction between cue specificity and audio cues for accuracy (left panel) and reaction time (right panel).

A significant two-way interaction between cue size and haptic cueing was obtained for reaction time, $F(1, 56) = 6.88, p < .01$, and for accuracy, $F(1, 56) = 12.89, p < .01$. The change in performance due to the presence or absence of the haptic cue is dependent upon the specificity of that cue (see Figure 16). When there is no specificity, performance is equivalent between presence and absence of haptic cues. However, when the cue size gets more specific, there is a benefit of the presence of the haptic cues for small, medium, and large cue sizes.

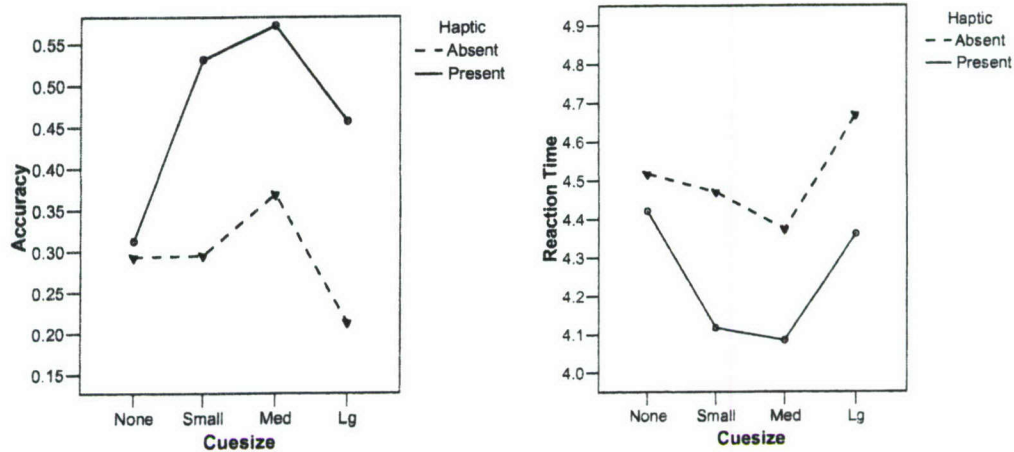


Figure 16. Interaction between cue specificity and haptic cues for accuracy (left panel) and reaction time (right panel).

A significant three-way interaction was obtained for accuracy of cue size X visual X audio, $F(1, 56) = 4.32, p < .05$, and reaction time of cue size X visual X audio, $F(1, 56) = 4.29, p < .05$. This indicates that the change in performance due to the presence or absence of the audio cues is dependent upon the presence or absence of the visual cues and the specificity of those cues (see Figures 17 and 18). When visual cues are absent, the change in performance due to the presence or absence of audio cues is dependent upon the specificity of the cues. Large cues have nearly identical performance for both the presence and absence of the audio cues. However, the no specificity, small and medium specificity cues indicated a benefit with the presence of the audio cues, medium cues showing the highest performance. When visual cues are present, the change in performance is different across cue specificities. The no specificity cues and large cues affected performance nearly the same as when visual was absent, but the small and medium cues caused a reduction in performance when audio is present, and an increase in performance when audio is absent.

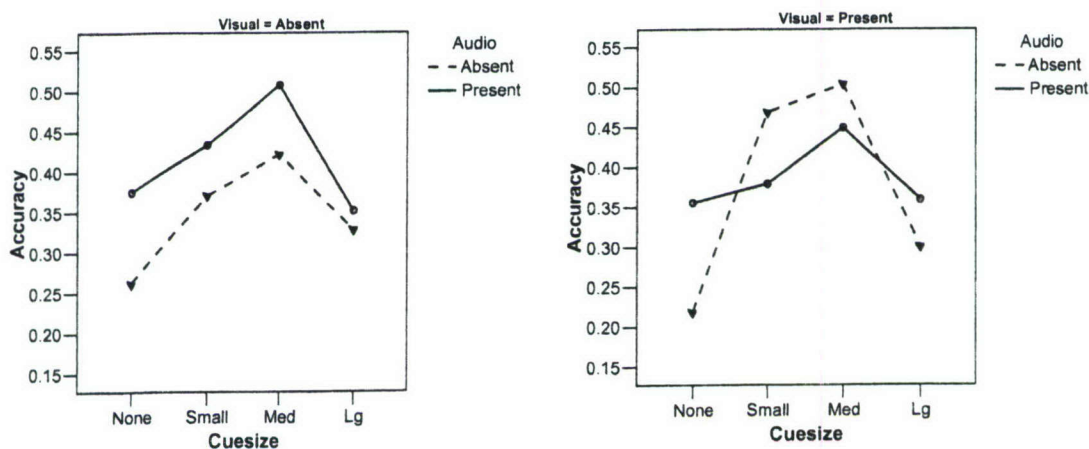


Figure 17. Three-way interaction between visual cues, audio cues, and cue specificity. The left-hand panel shows accuracy performance when visual cues are absent, and the right-hand panel shows performance when visual cues are present.

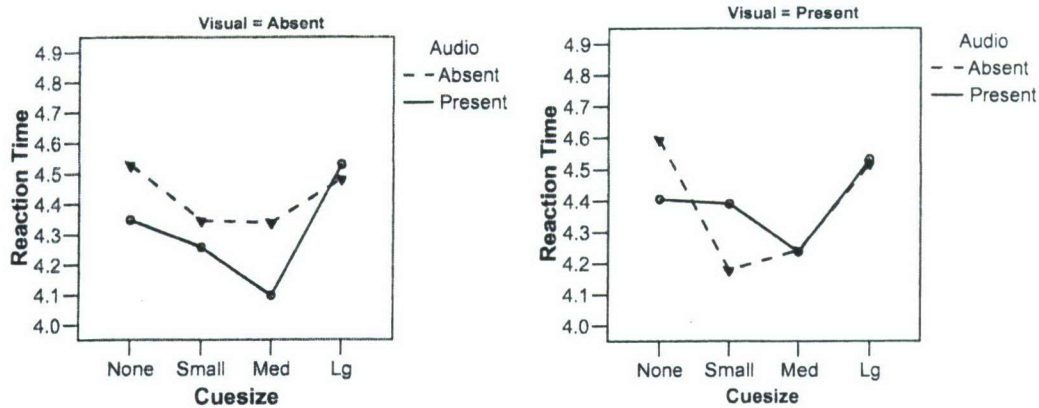


Figure 18. Three-way interaction between visual cues, audio cues, and cue specificity for reaction time. The left-hand panel shows reaction time performance when visual cues are absent, and the right-hand panel shows performance when visual cues are present.

Finally, there was a four-way interaction of workload X cue size X visual X audio for reaction time, $F(1, 56) = 9.37, p < .01$ and for accuracy, $F(1, 56) = 3.89, p < .05$. This indicates that the change in performance due to the workload level is dependent upon the presence or absence of the audio cues, the presence or absence of the visual cues and the specificity of those cues (see Figures 19 and 20). During low workload, the change in performance is identical to the three-way interaction described above, i.e., when visual cues are absent, the change in performance due to the presence or absence of audio cues is dependent upon the specificity of the cues. Large cues have nearly identical performance for both the presence and absence of the audio cues. However, the no specificity, small and medium specificity cues indicated a benefit with the presence of the audio cues, medium cues showing the highest performance. When visual cues are present, the change in performance is different across cue specificities. The no specificity cues and large cues affected performance nearly the same as when visual was absent, but the small and medium cues caused a reduction in performance when audio is present and an increase in performance when audio is absent. However, during high workload the trend is slightly different. When the visual cues are absent, no specificity cues and medium cues benefit more from the presence of the audio cues, and when visual cues are present, the performance is nearly the same for each cue specificity except the no specificity cues, which were improved by the presence of the auditory cues.

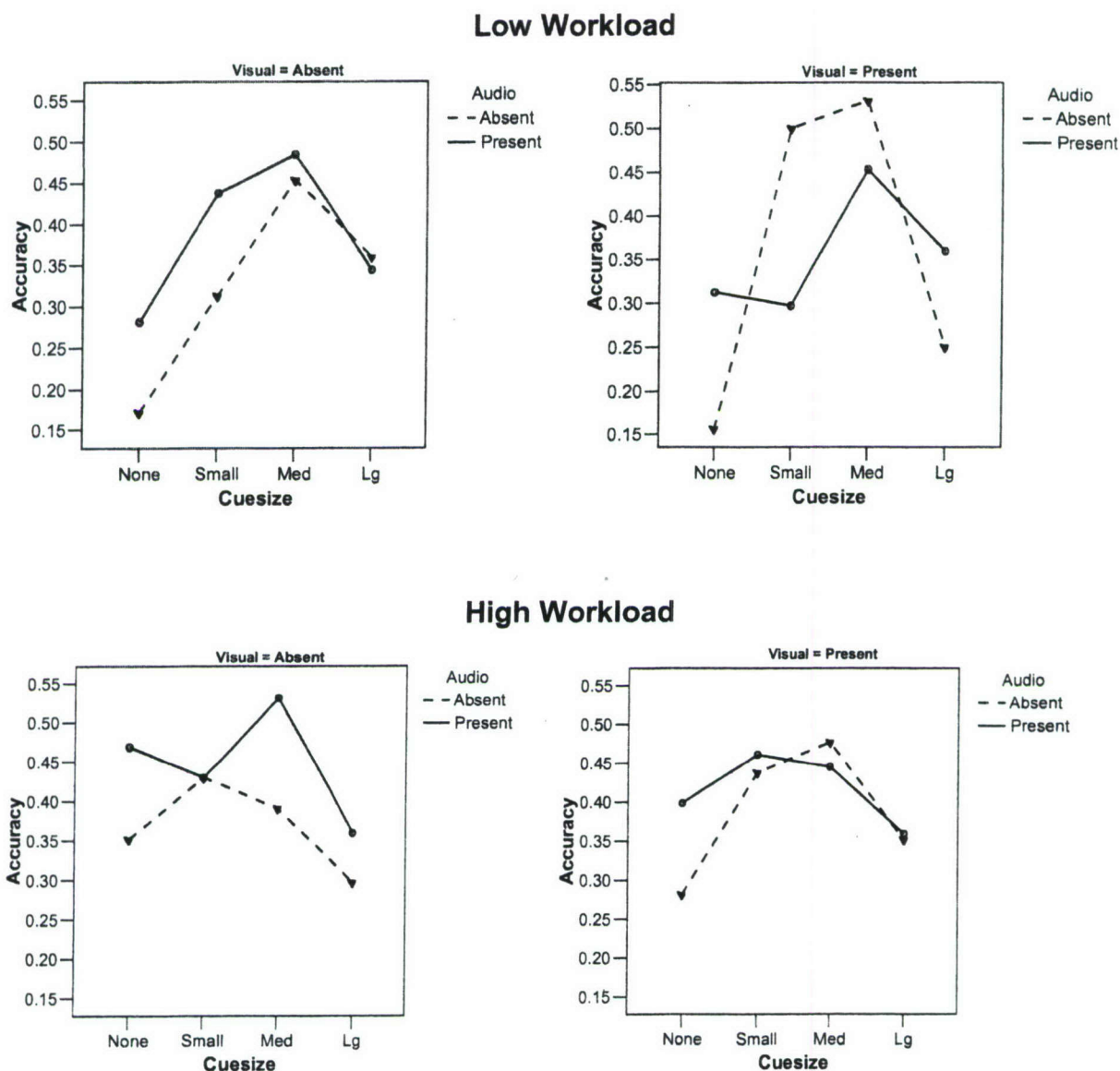


Figure 19. Four-way interaction for accuracy for workload level X visual cues X audio cues X cue specificity. The upper left panel shows performance during low workload when visual cues are absent. The upper right panel shows performance during low workload when visual cues are present. The lower left panel shows performance during high workload when visual cues are absent. And finally, the lower right panel shows performance during high workload when visual cues are present.

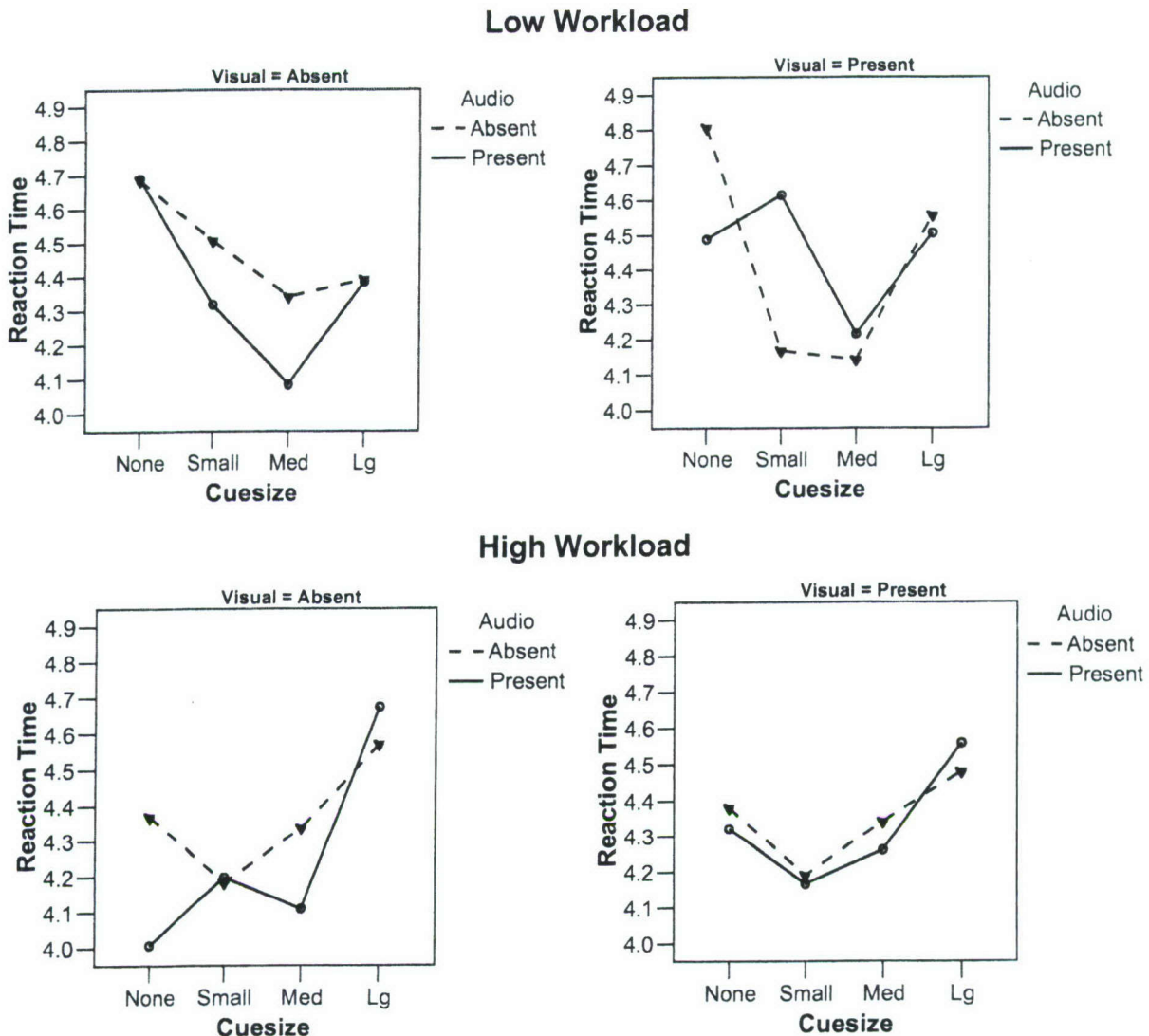


Figure 20. Four-way interaction for reaction time for workload level X visual cues X audio cues X cue specificity. The upper left panel shows performance during low workload when visual cues are absent. The upper right panel shows performance during low workload when visual cue is present.

Other Measures of Interest

Presence. The Presence Questionnaire (PQ) Version 3.0 is a 33 item questionnaire used to measure subjective feeling of the degree of presence which the participants perceived in the AR environment and reported shortly after AR exposure. The PQ score was correlated with both dependent variables representing task performance in the AR environment. Presence was positively correlated with accuracy, $r = .38, p < .01$, which means that as the participant felt more presence in the environment, they engaged more targets; or as they engaged more targets, they felt more presence. Further, presence was negatively correlated with reaction time, $r = -.25$,

$p < .05$, which indicates that as the feeling of presence increased, the participants were quicker at acquiring the targets; or as they acquired targets more quickly, the feeling of presence increased.

Table 8
Correlations of Other Measures of Interest

	Total Severity2	Presence
Total Severity2	--	--
Presence	-0.11	--
Condition	0.19	-0.22
Accuracy	-0.39**	0.38**
Fratricide Occurence	-0.14	0.14
Reaction Time	0.40**	-0.25*
NASA TLX	0.16	0.19
Light Accuracy	-0.08	0.11
Light Reaction Time	0.03	-0.10

* $p < .05$. ** $p < .01$. (2-tailed)

Simulator Sickness. The amount of simulator sickness experienced and reported based on simulator sickness scores after AR exposure was correlated with both dependent variables representing task performance in the AR environment. Simulator sickness was negatively correlated with accuracy, $r = -.39$, $p < .01$, which means that as the participant felt more sickness symptoms in the environment, they engaged fewer targets; or if the participants engaged fewer targets, they felt more sickness symptoms. Further, simulator sickness was positively correlated with reaction time, $r = .40$, $p < .01$, which indicates that as the sickness symptoms increased, the participants were slower at acquiring the targets.

Secondary Task Performance

Reaction time and accuracy of light extinguishing were recorded as measures of secondary task performance. The lights came on at random times, but did not vary across workload conditions (i.e. the lights came on at the same rate and time). Participant's performance were significantly different for each level of workload for reaction time, $F(1, 56) = 4.80$, $p < .05$, and accuracy, $F(1, 56) = 7.27$, $p < .01$, indicating that performance was better during low workload than during high workload (see Figures 21 and 22). None of the cue modalities showed an interaction with workload.

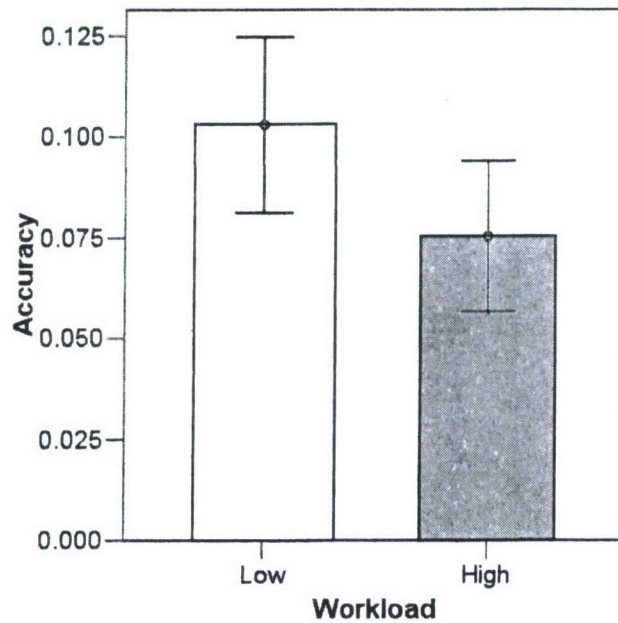


Figure 21. Effects of low and high workload for accuracy for light extinguishing. Points represent the mean reaction time to engage targets; vertical lines depict standard error of the means.

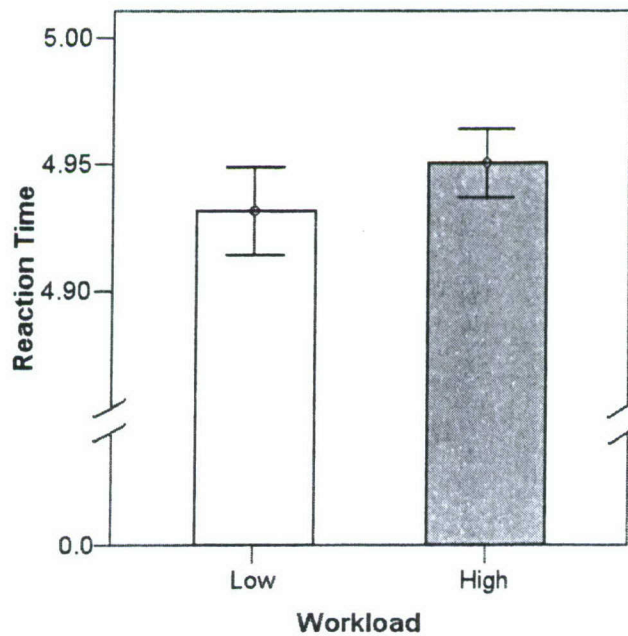


Figure 22. Effects of low and high workload for reaction time for light extinguishing. Points represent the mean reaction time to engage targets; vertical lines depict standard error of the means.

Discussion

The results of the current analysis are consistent with previous findings that performance of search tasks can benefit from a direction cueing system (Sach, Hill, & Bailey, 2000; Kennet, Eimer, Spence, & Driver, 2001; Ansorge & Heumann, 2003). This improved performance is due to a reduction in the amount of parallel information and freeing attentional resources during the perceptual and cognitive stage of information processing and making the target more salient (Kahneman, 1973; Norman & Bobrow, 1975; Wickens, 1992). All previous studies mentioned in this paper, however, have used limited field-of-view displays, even though attention in real world tasks spans much wider around the individual. This research was designed to empirically examine the effects of different modality cues on the performance of locating and interacting with hazards in the environment 360-degrees surrounding the participants in an AR system. The results provided a general view of how well directional auditory, visual, and haptic cues oriented attention to a specific spatial location. Along with the effects of the modality cues, effects of varying levels of workload, and finally the effects of different breadths of attentional cues were investigated. The findings of this research are consistent with previous studies; indicating that cues help orient attention to the location of the visual target, regardless of the type of modality (e.g., visual, auditory, or tactile (Posner, Snyder, & Davidson, 1980; Spence & Driver, 1997; Eimer & Driver, 2000)). In addition, the results of the workload analysis are consistent with previous research showing a workload effect on target detection performance, adding more support that workload affects the performance gained from exogenous cueing.

Furthermore, the results are not consistent with previous studies that found auditory cues only affected auditory localization (Ward, 1994; Ward, McDonald, & Lin, 2000). However, the results are consistent with some previous research that has found significant benefits of auditory cues across modalities (Spence & Driver, 1997). An explanation of these contradictory findings might be related to the equipment used in the current research. Previous studies used devices that deliver the cues and targets using simple PC systems with CRT monitors as the displays. Generally, these monitors span only 30-degrees of the visual field at most, when viewed from 2 feet away, which is not very adequate when testing peripheral exogenous cues. However, MR MOUT is a simulator which is considered a useful tool to empirically examine attentional cueing by displaying targets and cues overlaid upon the environment 360-degrees around the participant (Jerome, Witmer, & Mouloua, 2005). This system was able to display the auditory, visual and tactile stimuli in the desired spatial location. This ensures that the locations of the cues are accurate in relation to the locations of the targets.

Cueing Modality Effects

The results of the present research indicated significant benefits from the presence of the haptic cues but not significant benefits from the auditory or visual cues, when the total unique contribution of each variable is taken into account. The results of the planned comparisons, however, reveal slightly different results. The expected basic performance trend existed, with the exception of haptic performance being much higher than expected. Haptic performance was nearly equivalent to the Visual/Haptic combination performance, however with slightly more variability in scores. The auditory group performed significantly better than the control group, as did the visual/haptic group. These cue modality findings support the previous findings that cueing aids in the detection of visual targets; however, only haptic cues, regardless of the presence of other cues, significantly improved the detection of visual targets in a real/augmented

world consisting of a 360-degree environment. In a more controlled environment when the variance for each group is taken separately, the auditory cues support previous findings. However, the results did not support the benefit of visual cues alone as previously reported by Ansorge and Heumann (2003), Posner (1978), and Posner (1980). These contrasting results might be attributed to format and size of the display used in the current research. Since the field of view was much larger than the one used in previous studies, perhaps different attentional mechanisms underlie the nature of the cues. The 30-degree field of view of the monitors used in previous studies may not be wide enough to truly present stimuli to the participant's visual periphery. Therefore, even visual cues classified as exogenous may in fact be endogenous, in that the mechanisms used to orient attention is more top-down than bottom-up.

In general, these findings support the multiple resource theory proposed by Wickens (1984). Since the target detection task is predominantly a visual task, most of the resources available to the visual modality are used up, thereby disrupting the ability of the visual cue alone to aid in the target detection performance. The auditory cues show moderate performance benefits because this modality is used much less. However, the haptic cues show significant performance benefits because this modality is hardly used.

In addition, the results showed that fratricide occurrence was amplified by the presence of the audio cues but not the visual or the haptic cues. Further, the results of the planned comparisons for the expected trend of fratricide occurrence showed nearly identical performance patterns across modality conditions except for the auditory condition, which was significantly higher than the others. Localization cues (interaural intensity and time difference) can be ambiguous because the intensity and time difference can be accounted for from more than one location, for example front/back confusion (Proctor & Proctor, 1997). Further ambiguity stems from the artificial sound localizations produced when using technology to transform one dimension (mono/Left-Right) to two dimensions. Although the surround sound technique uses multiple speakers, the locations where there are no speakers rely on the combination of sounds originating from multiple locations. This combination of sounds adds to the ambiguity since the precise position and orientation of the ears is rarely consistent.

The interaction between visual and haptic cues for fratricide occurrence suggests that when each cue is presented alone, fratricide occurrence is diminished; however, when both cues are presented together, fratricide occurrence is amplified, even surpassing the level of occurrence with no cues. This interaction, although significant, had a fairly small effect, and therefore could be spurious. Furthermore, previous research has not reported an increase in false positives with multi-modal cueing. Although the author believes this to be a spurious finding, the following interpretation is provided for readers who disagree. The interaction could be caused by the participant's attentional resources being overwhelmed by the information from the cues. When few resources are available for the identification of the target as friend or foe, the subsequent decision to act (whether to shoot or not to shoot) is diminished (Wickens, 1992). Apparently, the mechanism that creates the benefit of the cues at orienting attention to the correct spatial location do not operate the same way when looking at the performance of incorrect responses, i.e. false positives, otherwise the expected performance trend for fratricide would be opposite to that found for target detection.

The overall results suggest that since fratricide occurs very infrequently, there is a general floor effect, with all groups having similar effects except for the auditory cue group. Perhaps the

inherent ambiguity of the auditory cues caused more confusion than the other, less ambiguous cues – including the group that received no cues. So instead of attention being the major factor for fratricide, it is ambiguity at work.

Workload Analysis

Results of the workload analysis indicate that, although the workload level affected target detection performance, the expected results were reversed, i.e. the low workload condition produced worse performance than the high workload condition. The levels of workload chosen for the target detection task were designed with the limitations of the research in mind, i.e. the amount of time the participant was able to be exposed to the environment was limited to eight minutes total (four two-minute scenarios). Thus, the participant either had very few potential targets in the low workload condition, or nearly one target every second in the high workload condition. More than one target at a time was not possible in the current research since the focus of this research was to allow every target to have a fair chance of being detected. Thus, the target presentation was as fast paced as it could possibly be.

The prediction that workload would interact with cueing modality was not supported in this research. Perhaps perceptual tunneling is nearly equivalent for all modalities, or the effect size was too small to detect with this experimental design. Another possibility is that the workload manipulation was ineffective. The expected performance trend is opposite of what was actually observed. The low workload condition was expected to allow the participant to conserve attentional resources for the target detection task; however, performance was worse, not better, in the low workload condition than in the high workload condition. Another possible interpretation could be related to motivation. The low workload condition might not have been engaging enough. The participants could have become bored and inattentive, thereby not performing at their best.

Interestingly, the workload analysis of the secondary task performance showed the expected performance trend. Performance of light extinguishing accuracy was higher in the low workload condition than the high workload condition, and the reaction time was faster in the low workload condition than in the high workload condition. This effect might be due to the participant's low engagement in the low workload condition. Target detection declined as participants focused more on the light extinguishing task. This finding is consistent with Treisman's (1998) revised feature integration theory, i.e., at low perceptual loads, late selection should occur, and at high perceptual loads, early selection should occur. If this were the case, when workload is low, people should be able to process more information, and be able to detect more targets. However, this only happens for the secondary task targets. The results of this research suggest that this theory might be slightly too simplistic, overly abbreviating a complex task. The assumed linear relationship between workload and performance might actually be more of a non-linear relationship where very low levels of workload yield early selection, but only for the less engaging aspects of the task.

In addition, the results of the workload effects on target detection performance were inconsistent with Jonides' (1976) findings for exogenous cueing. Perhaps exogenous cueing with a limited field of view (30 degrees, as studied in previous research) is not really exogenous cueing. These exogenous cues, although drawing the individual's attention to them, are more endogenous in that they are foveal, or near foveal and work from a top-down fashion.

Alternatively, the cues used in the current work (360 degrees total, but only 60 degrees seen at a time) might be misclassified as well. They were designed to be exogenous, but they could arguably be considered endogenous as well, especially the auditory and haptic cues. This would explain the findings that the visual cues did not benefit target detection; however, it could also be explained simply as the cue just not being detected when out of the field of view.

The results also suggest that workload might have different effects on primary and secondary tasks. At low levels of workload, primary task performance was poor probably because the scenario might not have been engaging enough and the participants might have been bored and inattentive, causing them to hesitate when a cue and target finally appeared, but at the same time providing more attentional resources to be allocated to the secondary task, and thus making the secondary task more salient than the primary task. At moderate levels of workload, performance is better, participants are engaged, and there is a minimum amount of stress, and at the same time reducing the amount of attentional resources allocated to the secondary task. At high levels of workload, performance once again declines, most likely due to too much stress, and requiring more attentional resources than were available to perform all aspects of the task efficiently, and at the same time the secondary task performance still declines because the participants are very engaged and using up most attentional resources for the primary task.

Cue Specificity Effects

The hypothesis that cue specificity would affect target detection performance was supported. The performance of both accuracy and reaction time measures was better with the medium cues than all the others and the control condition. The inverted U function for small, medium, and large cues was obtained in the low workload condition. However, the prediction that this inverted U function will be moderated by workload was not completely supported. As discussed earlier, care should be taken when interpreting these workload results since the workload manipulation did not have the expected performance effects.

Implications

This research provides strong empirical support to the previous work related to attention and performance. The results indicate that there are bandwidth limitations in the human attention system, i.e. there is an immense amount of information bombarding the individual, but only a very limited amount can pass through to be processed. The findings are partially consistent with Broadbent's (1958) information processing model. There is a filter that chooses the information that will receive further processing; however, the current cueing system acts as an external filter that aids the internal filter in choosing where in the environment (the channel) the important information is coming from. Wickens' (1992) model explains that the modality of this cue may provide benefits over other modalities based on the characteristics of the task and environment. If the visual system is overloaded, as it is in the current work, then visual cues will benefit less than the other modality cues.

Inconsistent with Jonides' (1976) findings that only endogenous cueing was affected by workload; exogenous cueing was found to be affected by workload in the setting employed in the current work. This supports the approach that attentional resources are shared by a variety of tasks, including perception, attention, decision making, and even response execution.

The results of the present research also show much support for the haptic modality for an attentional cue. This supports the idea that when a target originates from another modality (e.g., vision, and perhaps audition), a haptic cue can be remapped to the other modality and aid in target detection (see Kennet, Eimer, Spence, & Driver, 2001; and Eimer & Driver, 2000).

In addition, these results provide support for the spotlight of attention theory. Attentional breadth could be cued using cues of different sizes and of different modalities. Therefore, the cue possessed two kinds of information about the location of the target in the 360 degree environment: direction and specificity. The direction information oriented the individual to the general area in the visual field the target resided, and the specificity information reduced the visual field into a smaller, workable area. Specificity, in a way, reduced the parallel information existing in the environment, and made the search task more serial in nature.

The findings of this experiment support the added benefits of multimodal cues in the orienting and performance of attention tasks. They further support and extend previous research on orienting of attention. In addition, these findings add a new piece of evidence regarding the benefits of multimodal aspect of augmented reality in a complete 360 degrees. Also, as new technologies are being developed for military operations, it is necessary to use such multimodal cues in order to reduce workload and better orient the Soldier's attention without causing distraction.

In recent years, The U.S. Army has made a considerable investment in new technologies aimed at helping the Soldiers become better trained and efficiently perform on the electronic battlefield. However, such technologies have resulted in higher levels of stress and workload. The present findings will serve as a basis to provide a variety of training and design recommendation to direct attention during military operations, cueing the Soldier to the location of hazards, and mitigating the effects of stress and workload.

One possible application of an attention cueing system as tested in this research might be for actual military combat. Although the testing procedure was very simplistic, it represented just that, a task in which the user must determine the location of the enemies and engage them as quickly as possible. As technologies advance and the military incorporates these technologies into the Soldier's standard equipment, the target cueing system could be of benefit. Technologies already exist and are being used in aircraft's head-up displays (HUDs). When an enemy aircraft is tracked, the HUD places a box around the target to aid the pilot in locating, identifying, and reacting to it. This technology shows the success of such a system incorporating the visual modality to cue attention, however this research shows that other modalities might aid also if not better. Particularly, the haptic cues would be of great benefit based on the current results. Furthermore, this modality is highly underused and provides an information avenue with an excess of attentional resources available to leverage.

Prior to using this cueing system in military combat, this type of system could be used during training. This may be a highly successful solution to the problem of training how, when, and what to look for during a target detection task. When first attempting a novel target detection task, the targets and noise blend together until the task becomes more familiar and the distinguishing features between them are more salient. During the training phase, the individual could be cued to these distinguishing features, highlighting what should be attended to during the visual search task, e.g. visual features, locations, or behaviors to be aware of. These skills could

be honed much more quickly and efficiently using the attention cueing system and could generalize to the real world task, improving performance even when the system is not being used.

Cueing Guidelines

The results of this experiment can help with the design and evaluation of systems incorporating augmented information in an effort to inform the user of the location of enemies, hazards, or targets out in the environment. Following is a list of design recommendations for an augmented reality system.

1. Tactile cues should be used as a way to inform users of the spatial locations of targets.
2. Auditory cues should be used sparingly, especially if incorrect identification has negative effects (e.g. fratricide).
3. If available, a combination of visual and tactile cues should be used to quickly inform the user with the tactile cue, and then to reduce the distracters and noise in the environment with the visual cue.
4. Medium sized cues should be used as much as possible, even if more specific location information is available, i.e. the smaller, more specific cues should be avoided as this reduces accuracy and increases reaction time.

Conclusion

The current research provides answers to some of the questions sought based on the previous literature reviewed; however, other questions still exist, and have been brought to light with the current findings. These questions should be explored to more clearly explain the complex nature of orienting attention and the many different situations that could alter the way the augmented information most efficiently orients.

More research is needed to investigate cue reliability and trust. The present research was designed to specifically understand the nature of modality cueing, cue specificity, and workload in an accurate and reliable system. However, systems rarely are perfect, and when the feedback is given of an incorrect cue, the consequences are stored in memory and are used later during the orienting and decision making processes. Also, the findings that auditory cueing, even in a perfectly reliable system, increased fratricide is an indication that ambiguity of the spatial location of the cue could affect the detection and identification of targets that were not even cued to. This effect might certainly be amplified when viewed from an unreliable system.

Future research should also be conducted to explore workload effects in more detail. The choice of workload levels in the current experiment did not produce the expected effects on performance because of motivational factors. Future research should further explore multiple levels of workload to better understand affects on performance as it relates to cueing modality and cue specificity.

References

- Ansorge, U. & Heumann, M. (2003). Top-down contingencies in peripheral cuing: The roles of color and location. *Journal of Experimental Psychology / Human Perception & Performance*, 29(5), 937-949.
- Atkinson, R. C. & Shiffrin (1968). Human memory: A proposed system and its control processes. In K. W. Spence. and J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in Research and Theory* (Vol. 2, pp. 89-105). New York: Academic Press.
- Bahri, T. (1989). Orienting of Attention and Vigilance. Unpublished Doctoral Dissertation. Department of Psychology. The Catholic University of America.
- Barfield, W. & Caudell, T. (2001). *Fundamentals of Wearable Computers and Augmented Reality*. Lawrence Erlbaum Associates, Inc. Mahwah, NJ, USA.
- Blanco, M. J. & Soto, D. (2002). Effects of spatial attention on detection and identification of oriented lines. *Acta Psychologica*, 109(2), 195-212.
- Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon.
- Bronkhorst, A.W., Veltman, J.A., & Breda, L. (1996). Application of a three-dimensional auditory display in a flight task. *Human Factors*, 38(1), 23-33.
- Cherry, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. *Journal of the Acoustical Society of America*, 25, 975-979.
- Crick, F. (1984) Function of the thalamic reticular complex: The searchlight hypothesis. *Proceedings of the National Academy of Sciences USA* 81: 4586-4590.
- Deutsch, J. & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, 70. 80-90.
- Eimer, M., & Driver, J. (2000). An event-related brain potential study of cross-modal links in spatial attention between vision and touch. *Psychophysiology*, 37, 697-705.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). G*Power: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, 28, 1-11.
- Feiner, S. K. (2004). Augmented reality: A new way of seeing. Scientific American. Retrieved from the World Wide Web, July 8, 2004, from:
<http://www.sciam.com/article.cfm?chanID=sa006&colID=1&articleID=0006378C-CDE1-1CC6-B4A8809EC588EEDF>

- Ferlazzo, F. , Padovani, T. , Couyoumdjian, A. , Olivetti, M. (2002). Head-centered meridian effect on auditory spatial attention orienting. *The Quarterly Journal of Experimental Psychology*, 55A, 937-963.
- Friesen, C. K. (2001). Attentional Effects of Gaze Direction. Unpublished Doctoral Dissertation. Department of Psychology. University of Alberta.
- Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload*, 139-183. Amsterdam, The Netherlands: Elsevier.
- Hatfield, G. (1998). Attention in early scientific psychology. In R. D. Wright (Ed.), *Visual Attention* (pp. 3-25). New York: Oxford University Press.
- Hillyard, S. A., Luck, S. J., Mouloua, M., Downing, C. J., & Woodward, D. P., (1990). Visual attention modulates signal detectability. *Journal of Experimental Psychology: Human Perception and Performance*. 16(4). 802-811.
- Hopkin, V. D. (1995). *Human Factors in Air Traffic Control*. Taylor & Frances, London, UK
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11). 1254 - 1259
- Jerome, C. J., Helmick, J. A., Mouloua, M., & Hancock, P. A. (2002, March). The effects of navigation system use on driver's performance. Poster presented at the South Eastern Psychological Association annual meeting. Orlando, FL.
- Jerome, C. J., Witmer, B., & Mouloua, M. (2005). Spatial orienting attention using augmented reality. Paper presented at the Augmented Cognition Conference, July 27, Las Vegas, NV.
- Johnston, W. A., & Heinz, S. P. (1978). Flexibility and capacity demands of attention. *Journal of Experimental Psychology: General*, 107, 420-435.
- Jonides, J. (1976). Voluntary versus reflexive control of the mind's eye's movement. *Paper presented at the meeting of the Psychonomic Society, St. Louis, MO*.
- Jonides, J. (1980). Toward a model of the mind's eye's movement. *Canadian Journal of Psychology*, 34, 103-112.
- Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall
- Kellogg, R. S., Kennedy, R. S., & Graybiel, A. (1965). Motion sickness symptomatology of labyrinthine defective and normal subjects during zero gravity maneuvers. *Aerospace Medicine*, 4, 315-318.

- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). A simulator sickness questionnaire (SSQ): A new method for quantifying simulator sickness. *International Journal of Aviation Psychology*, 3(3), 203-220.
- Kennet, S., Eimer, M., Spence, C., & Driver, D. (2001). Tactile-visual links in exogenous spatial attention under different posters: convergent evidence from psychophysics and ERPs. *Journal of Cognitive Neuroscience*, 13(4), 462-478.
- Kindratenko, V. (2001). A Comparison of the Accuracy of an Electromagnetic and a Hybrid Ultrasound-Inertia Position Tracking System. *Presence: Teleoperators and Virtual Environments*, 10(6), 657-663.
- Lloyd, D. M., Bolanowski, S. J., Howard, L., & McGlone, F., (1999). Mechanisms of attention in touch. *Somatosensory & Motor Research*, 16(1), 3-10.
- Luck, S. J., Hillyard, S. A., Mouloua, M., & Hawkins, H. L. (1996). Mechanisms of visual-spatial attention: resource allocation or uncertainty reduction? *Journal of Experimental Psychology*, 22(3), 725-737.
- Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effects of spatial cueing on luminance detectability: psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology*, 20(4), 887-904.
- McKnight, A.J. & McKnight, A.S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention*, 25(3), 259-265.
- Mosier, K. L. & Skitka, L. J. (1996). Human Decision Makers and Automated Decision Aids: Made for Each Other? In R. Parasuramen & M. Mouloua (Eds.) *Automation and Human Performance*. Lawrence Erlbaum Associates, Mahwah, New Jersey.
- Neumann, U. & Majoros, A. (1998). Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. In Proceedings of IEEE VRAIS'98, 4-11, 1998.
- Niebur, E., Itti, L., & Koch, C., (2001). Controlling the Focus of Visual Selective Attention, In: *Models of Neural Networks IV*, (L. Van Hemmen, E. Domany, J. Cowan Ed.), Springer Verlag, Aug 2001.
- Niebur, E. and Koch, C. (1996). Control of selective visual attention: Modeling the "where" pathway. In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E., (Eds.) *Advances in Neural Information Processing Systems*, volume 8, pages 802-808. MIT Press, Cambridge, MA.
- Norman, D. A. and Bobrow, D. G. (1975). On data-limited and resource limited processes. *Cognitive Psychology*, 7, 44-64.

- Posner, M. I. (1978). *Chronometric Explorations of Mind*. Hillsdale, N. J.: Erlbaum.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160-174.
- Pratt, J. and McAuliffe, J. (2002). Determining whether attentional control settings are inclusive or exclusive. *Perception & Psychophysics*, 64(8), 1361-1370.
- Prime, D. J. and Ward, L. M. (2002). Auditory frequency-based inhibition differs from spatial IOR. *Perception & Psychophysics*, 64(5), 771-784.
- Proctor, R.W. and Proctor, J.D. (1997). Sensation and perception. In G. Salvendy (Ed.) *Handbook of Human Factors and Ergonomics* (2nd ed., pp. 43-88). New York, NY: John Wiley & Sons, Inc.
- Reisberg, (1997). *Cognition: Exploring the science of the mind*. W. W. Norton & Company, New York.
- Rheingold, H. (2004). Augmented windshields? Retrieved from the World Wide Web, July 8, 2004 from: <http://www.thefeature.com/print?articleid=100752>
- Rosenthal D. M. (1998). Introspection. Retrieved from the World Wide Web, November 15, 2004. <http://web.gc.cuny.edu/cogsci/introspn.htm>
- Sach, A. J., Hill, N. I., & Bailey, P. J., (2000). Auditory spatial attention using intramural time differences. *Journal of Experimental Psychology: Human Perception and Performance*, 26(2), 717-729.
- Schmitt, M., Postma, A., & DeHaan, E. (2000). Interactions between exogenous auditory and visual spatial attention. *The Quarterly Journal of Experimental Psychology*, 53A(1), 105-130.
- Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1005-1030.
- Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 555-574.
- Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, 59, 1-22.
- Tabachnick, B. G., and Fidell, L. S. (2007). *Using Multivariate Statistics*, 5th ed. Boston: Allyn and Bacon.

- Taylor, J. G. (2004). Models of attention. Retrieved from the World Wide Web, July 13, 2004, from <http://www.epub.org.br/cm/n16/mente/attention.htm>.
- Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, 9(2), 314-318.
- Treisman, A., & Gormican, S., 1988. Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15-48.
- Treisman, A. (1998). Feature binding, attention and object perception. In *Philosophical Transactions of the Royal Society of London*, 353, 1295-1306.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Van der Heijden, A. H. C. (1992). *Selective attention in vision*. London: Routledge.
- Yeh, M. & Wickens, C. D. (2000). Attention and trust biases in the design of augmented reality displays. Technical Report # ARL-00-3/FED-LAB-00-1. U. S. Army Research Laboratory, Interactive Displays Federated Laboratory.
- Ward, L. M. (1994). Supramodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. *Canadian Journal of Experimental Psychology*, 48, 242-259.
- Ward, L. M., McDonald, J. J., & Lin, D. (2000). On asymmetries in cross modal spatial attention orienting. *Perception and Psychophysics*, 62, 1258-1264.
- Wickens, C. D. (1992). *Engineering psychology and human performance*, 2nd Ed. Columbus, OH: Merrill.
- Wickens, C. D. & Liu, Y. (1988). Codes and modalities: A success and a qualification. *Human Factors*, 30(5), 599-616.
- Witmer, B. G., & Singer, M. J. (1994). *Measuring presence in virtual environments*. (Technical Report 1014). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. (ADA 286183)
- Witmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), 225-240.

APPENDIX A:
TASK SCENARIO CHARACTERS

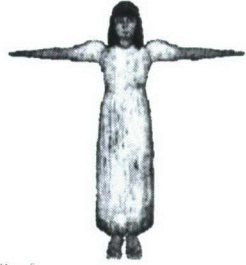
 Headband (enemy)	 Intel(enemy)	 Militia(enemy)	 Glasses(enemy)
 FFW(enemy)	 Bald(enemy)	 Sniper(enemy)	 Girl(civilian)
 Hostage(civilian)	 Mechanic(civilian)	 Farmer(civilian)	 Peasant(civilian)

Figure A-1. Task scenario characters (enemy and civilian).

APPENDIX B:
MR MOUT PORTAL MAPS



Figure B-1. MR MOUT portal maps.

APPENDIX C:
BIOGRAPHICAL QUESTIONNAIRE

Research Participant Information Questionnaire

Keyboard Directions: Position the cursor over the response that you want to select for a given question, then click the left mouse button to select it. If applicable, type in your answer. Use the scroll bar or PgDn button to move to the next set of (off-screen) questions. Please tell the experimenter when you are finished.

Instructions: Please click on the appropriate response.

1. Please type in your age.

_____ Years Old

2. What is your gender?

Female ____ Male ____

3. Are you currently in your usual state of good fitness?

No ____ Yes ____

4. Type in the number of hours sleep you had last night. Use a decimal format, e.g., 7.5, 8.0, etc.

_____ Hours Sleep

5. Have you ever experienced car or motion sickness?

No ____ Yes ____

6. How susceptible to motion or car sickness do you feel you are?

_____	_____	_____	_____	_____	_____
Not	Very Mildly		Average		Very Highly
Susceptible					

7. Do you have a good sense of direction?

No ____ Yes ____

8. Type in the number of hours per week that you use a computer. Use a decimal format, e.g., 7.5, 8.0, etc.

_____ Hours per Week

9. My level of confidence in using computers is:

_____	_____	_____	_____
Low	Average	High	

10. I enjoy playing video games (home or arcade):

_____	_____	_____	_____
Disagree	Unsure	Agree	

11. I am _____ at playing video games:

_____	_____	_____	_____
Bad		Average	Good

12. Type in the number of hours per week that you play video games. Use a decimal format, e.g., 7.5, 8.0, etc.

_____ Hours per Week

13. How many times in the last year have you experienced a virtual reality game or entertainment?

_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
0	1	2	3	4	5	6	7	8	9	10	+10

14. Do you have a history of epilepsy or seizures?

No ___ Yes ___

15. Do you have normal or corrected to normal 20/20 vision?

No ___ Yes ___

16. Are you color blind?

No ___ Yes ___

END Research Participant Information Questionnaire Form: Please inform the experimenter that you are finished. DO NOT click any of the buttons located below the red line.