TECHNIQUES FOR ASSESSING AND IMPROVING PERFORMANCE IN NAVIGATION AND WAYFINDING USING MOBILE AUGMENTED REALITY

by

BRIAN F. GOLDIEZ
B.S.A.E. University of Kansas, 1973
M.S. University of Central Florida, 1979

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Modeling and Simulation
in the College of Arts and Sciences
at the University of Central Florida
Orlando, FL

Fall Term
2004

ABSTRACT

Augmented reality is a field of technology in which the real world is overlaid with additional information from a computer generated display. Enhancements to augmented reality technology presently support limited mobility which is expected to increase in the future to provide much greater real world functionality. This work reports on a set of experiments that investigate performance in search and rescue navigating tasks using augmented reality. Augmentation consisted of a spatially and temporally registered map of a maze that was overlaid onto a real world maze. Participants were required to traverse the maze, answer spatially oriented questions in the maze, acquire a target object, and exit. Pre and post hoc questionnaires were administered. Time and accuracy data from one hundred twenty participants were collected across six treatments. The between subject treatments, which had an equal number of male and female participants, were a control condition with only a compass, a control condition with a paper map available prior to maze traversal and four experimental conditions consisting of combinations of egocentric and exocentric maps, and a continuously on and on demand map display. Data collected from each participant consisted of time to traverse the maze, percent of the maze covered, estimations of Euclidian distance and direction, estimations of Cardinal direction, and spatial recall. Data was also collected via pre and post hoc questionnaires. Results indicate that best performance with respect to time was in the control condition with a map. The small size of the maze could have facilitated this result through route memorization. Augmented reality can offer enhancement to performance as navigational tasks become more complex and saturate working memory. Augmented reality showed best performance in accuracy by facilitating participants' coverage of the maze. Exocentric maps generally exhibited better performance than egocentric maps. On demand displays also generally resulted in better performance than continuously on displays. Gender differences also were evident with males exhibiting better performance than females. Participants reporting an initial tendency to not rotate maps exhibited better performance than those reporting a tendency to rotate maps. Enhancements being made to augmented reality and related technologies will result in more features, improved form factor for users, and improved performance in the future. Guidelines provided in this work seek to ensure augmented reality systems continue to progress in enhancing performance.

ACKNOWLEDGMENTS

Many people and their support impacted the research reflected in this work. Three deserve particular note. They are my wife, Dr. Peter Hancock and Dr. Mansooreh Mollaghasemi. My wife has encouraged me to pursue this effort and has foregone our personal time, weekends, and vacations allowing me to complete this work. I am truly grateful for her support. Dr. Mollaghasemi was the principal force encouraging me to pursue a doctoral degree at this stage of my career. Her encouragement started me on a journey that has been most rewarding. Dr. Hancock has been the principal guide in that journey. I particularly appreciate his giving me wide breadth in my research, yet providing the needed focus, encouragement, and guidance when needed.

Many others deserve acknowledgement. Dr. Randall Shumaker, Director of the Institute for Simulation and Training, has given me the needed flexibility in executing my work tasks, thereby allowing me to conduct the research needed to complete this work. Dr. M. J. Soileau, Vice President of Research, provided encouragement to proceed with my doctorate. My dissertation committee has been supportive in many different ways, including ensuring that the research hypotheses are appropriate, that supporting data is relevant, and indicating where additional information is needed.

The agencies and people supporting this research have been most accommodating. I appreciate the support from Congressman Ric Keller, R-Florida, and Senator Bill Nelson D-Florida in their support of the modeling and simulation program, and associated augmented and virtual reality research at the University of Central Florida. The Naval Research Laboratory and Office of Naval Research not only administered the grant, but also provided guidance, interaction, and review of the research as it progressed. The support from Army Research Institute for Behavioral Sciences in Orlando in providing the augmented reality hardware and review of questionnaires has been superb.

Finally, but by no means last in moral and technical support, are the other graduate students who participated in this research. Ali Ahmad, Jeffrey Dawson, Daniela Kratchouvova, and Radhey Shah each brought individual expertise to support executing the research task. I am grateful for their help and hope the experience they gained on this work will serve them in their individual research pursuits.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

ANACOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AR	Augmented Reality
ARI	Army Research Institute
BARS	Battlefield Augmented Reality System
C_M	Control with Map
C_NM	Control with Compass –No Map
DV	Dependant Variable
E_C	Egocentric Map and Continuous Display
E_D	Egocentric Map and On Demand Display
FOV	Field of View
G-Z S-O	Guillford-Zimmerman Spatial Observation
G-Z S-V	Guillford-Zimmerman Spatial Visualization
GPS	Global Positioning System
GPU	Graphics Processing Unit
HMD	Head-Mounted Display
HUD	Head Up Display
ICF	Informed Consent Form
IE_time	Instrument Landing System
ITQ	Immersive Tendencies Questionnaire
IV	Independent Variable

LAN	Local Area Network
LCD	Liquid Crystal Display
LRS	Landmark-Route-Survey
PC	Personal Computer
PQ	Presence Questionnaire
SCR	Skin Conductance Response
SPSS	Statistical Package for Social Sciences
SUMCDir	Summation of Cardinal Direction Score
SUMEDir	Summation of Euclidian Direction Score
SUMEDis	Summation of Euclidian Distance Score
SUS	Slater, Usoh. Steed
TLX	Task Load Index
VE	Virtual Environment
VGA	Video Graphics Adapter
VSSP	Visuo-Spatial Sketchpad
WAN	Wide Area Network
Wt_T	Weighted Time
X_C	Exocentric Map and Continuous Display
X_D	Exocentric Map and On Demand Display

CHAPTER ONE: INTRODUCTION

Augmented reality (AR) describes a field of technology in which the real world is overlaid with additional information from a computer generated sensory display. The real world is the baseline upon which information is added, as contrasted with virtual reality where the desired state is to completely immerse the human's sensory systems within a computer created environment. Virtual reality's baseline is in a virtual or artificial environment the computer creates. As one adds more computer augmentation to a real world, the demarcation between virtual and augmented (as well as other types of realities) becomes blurred. Rapid advances in technology have contributed to this blurring process. This confluence of the various realities provides opportunities for some of the technologies to be easily adapted from one domain to another. The blurring also provides an opportunity to adapt human performance studies in one domain to better understand human performance in other related domains.

Technology has now reached a point of maturity where prototypical augmented reality systems are available and research can now begin optimizing the particular configurations and utility of those configurations that define various types of augmented reality. With this knowledge, particular types of augmented reality systems can be categorized and defined and methods for assessing their utility in different applications can be investigated. Closely linking research related assessments and development will accelerate augmented reality's maturation and introduction into society. Also, describing a human centered framework for augmented reality technology development and its use will facilitate a collaborative environment for assessment and development.

Research in augmented reality (AR) systems is timely and of interest because of AR's potential benefit to enhance the user's performance in the real world by having the ability to portray information that is not normally visible or immediately accessible to the user. AR systems provide intelligent amplification of computer utility as reported by Brooks (1996) in Azuma (1997). Augmented reality systems may also reduce the cognitive load associated with task performance by fracturing the limitations of human sensory system and information processing capacities (Neumann & Majoros, 1998). The super imposition of graphics onto real scenes through the use of display combining technology can therefore aid the user's performance of real world activities.

Augmented Reality Systems

Augmented reality has been defined by Barfield and Caudell (2001) as "a participant wears a see-through display (or views video of the real world with and opaque HMD) that allows graphics or text to be projected in the real world" (pg. 6). The participant is normally tethered to a computer and other hardware. The same authors characterize wearable computers, "where the user actually wears the computer and, as in virtual or augmented reality, wears the visual display" (pg 7). The wearable computer may be wirelessly connected to a local area network (LAN) or wide area network (WAN), thus allowing information to be accessed whenever and wherever the user is in the environment. Augmented reality systems are in the process of being merged with wearable computers providing increased mobility for the user. All augmented reality systems, though, rely on reality and add computer generated content to extend that reality as contrasted with virtual reality systems which mediate all sensory inputs. It is important to note that augmentation may also entail removal (negative addition) of selective information from the real world. For example, one might want to visualize a room without a bookcase or a theatre with a curtain drawn.

Current augmented reality systems have modest amounts of computer graphics as contrasted with virtual reality systems where relatively large amounts of computer graphics are employed. One can envision graphics information generated for augmented reality in two realms; added data pertinent to the task, but cognitively separable from the real world, or fused information that is pertinent to the task and difficult for the participant to distinguish between real and augmented information. This same separation should apply to other sensory modalities as the appropriate technology matures.

Augmented reality systems have distinctive features that characterize their functionality. Barfield and Caudell (2001) and Azuma, Baillot, Behringer, Feiner, Julier, and MacIntyre (2001) describe these functional characteristics that include blending the real and virtual in a real environment, real time interactivity, and 3D registration of information. Milgram, Takemura, Utsumi, and Kishino (1994) use the terms reproduction fidelity, extent of presence, and extent of world knowledge augmented to describe the functional characteristics of virtual and augmented reality systems. These terms represent respectively i) the graphics quality, ii) the degree to which the user forms an integrated mental model, and iii) the amount of real world knowledge available via the computer.

Head Up Displays as Augmented Reality Systems

Head Up Displays (HUDs) used in many aircraft are examples of augmented reality systems. They provide navigational and system aides to the pilot that augment out of the cockpit view. HUDs are one of the lowest orders of augmented reality systems since they have low reproduction fidelity, a small extent of presence and relatively small extension of world knowledge because of low graphics quality, marginal cognitive separation of computer and real world environments, and selective knowledge of the external world. HUD augmentation, though, has been an area of intense research in both aviation and automotive industries (see Wickens & Hollands, 2000). Therefore, an opportunity exists to explore how far HUD related research findings can be extended into higher fidelity augmented reality systems for navigation related tasks.

Virtual reality systems support, but also exhibit difficulties in various aspects of navigation, such as users becoming disoriented (Durlach & Mavor, 1995; Psotka, 1995). Careful comparison between the features and techniques of virtual and augmented reality systems and associated research provides some insights into designing effective augmented reality systems for navigation and wayfinding.

Augmented reality systems can be classified into three groups; augmentation that is not part of the environment, augmentation that is integrated and part of the environment, and augmentation that is part of the environment, but not perceived without augmentation. The heads up display is an example of augmentation that is not part of the environment. Augmentation that is integrated and part of the environment includes computer representations of physical objects that are added to the real environment. An example could be computer generated furnishings that are added to a real room. The final example of augmentation is where information is in the real environment, but cannot be perceived without augmentation. An example is an infrared sensor that is processed and fused with a real world scene. The technical approaches used to deliver augmentation through these various groups and which specific task that is best met by a specific type of augmentation has not been definitized.

Augmented Reality and Human Centered Research

AR is not a new area of research. However, to date most efforts have been focused on the technological barriers and few studies have been conducted on the human aspects of this emerging technology. More specifically, many questions remain unanswered concerning the perceptual and cognitive implications of the use of such technology for improving human performance (Barfield & Caudell, 2001). Some studies have recently been made on manipulative tasks, such as the work reported by Tang, Owen, Biocca and Mou (2003). These authors used an assembly task in their study and found that a 3D augmented reality overlaid on objects reduced errors by 82 percent when compared to printed manuals. Attention switching and spatial transformations were also reduced. One area of human performance that has not been studied in augmented reality is wayfinding. Wayfinding is characterized by acquiring landmark, route, and survey knowledge about an area as described by Siegle and White (1975). Wayfinding is one of the crucial application areas for augmented reality in areas such as public safety. Wearable computer technology makes real-world wayfinding now feasible.

The primary purpose of the present work is to investigate the impact of using different augmented display design approaches and display contents on wayfinding performance using mobile augmented reality. The secondary goal of this research includes validation of some existing subjective measures of performance in an augmented reality setting, including the existence of a sense of presence, such as described by Singer and Witmer (1996).

A specific wayfinding task called search and rescue is used to study augmented reality. Search and rescue wayfinding occurs when one must search an area, find an objective, exit by retracing a route learned during ingress, and recall the configuration of the space traversed. Any learning should normally occur during ingress. This area of execution of wayfinding is appropriate for two reasons. First, greater mobility is now feasible allowing for techniques to expand areas of coverage. In the foreseeable future, wearable augmented reality will be available as an assistive product for use by the military, public safety, entertainment, the elderly, and other constituencies. Secondly, the execution of wayfinding is a relevant issue in many contemporary settings in defense, homeland security, and public safety. One may not have an opportunity to train or practice in a virtual setting before needing to find their way to a point in the environment, such as a hostage rescue in a building. Additionally, an individual may have only one chance to properly execute the wayfinding task, say when looking for someone in a smoke-filled building. Wayfinding is an area that can be difficult to train and practice in anything other than in the real environment. For example, Stanney and Salvendy (1995) cite particular difficulties in acquiring wayfinding knowledge in virtual environments. Many of these difficulties are due to limited navigational metaphors in the virtual environment. It is therefore interesting to see how a mixture of real and virtual might be carefully integrated to facilitate wayfinding.

The expected outcome of this work is a quantification of our understanding of the trade offs between augmented reality display contents, persistence of the displayed images, egocentric, and exocentric display orientation in order to facilitate improving human performance in wayfinding. The work is significant in that it is the first comprehensive treatment that evaluates and recommends methods to exploit and guide technical developments to improve our understanding of the dynamics of using mobile augmented reality to enhance human performance in navigation and wayfinding tasks. It also confirms the utility of conducting human performance studies on systems that are early in their development cycle.

Conceptual Framework

A conceptual framework was developed in order to facilitate the investigation of human performance in wayfinding using AR. Given the rarity of data on the interaction of the constructs within this framework, this conceptualization represents a means by which the field can generate testable hypotheses in order to determine the direction of interactions associated with these factors within AR environments. Figure 1 is a graphical representation of the conceptual framework. The motivation for the development of the conceptual framework is to better understand the underlining processes of the interaction between technology and human performance, more specifically, the overarching goal is to assess the degree to which AR environments and the varying forms of information presentation they afford, impact both cognitive processes and human performance in the context of wayfinding and navigation. It is hypothesized the different components of working memory be differentially tapped by the task and technology factors and as a result human performance will also be affected. Thus, using this framework, and through the analyses of task performance, it will be possible to investigate how these factors affect information processing and predict performance changes associated with task and technological changes. Description of the primary components of the proposed conceptual framework follows.

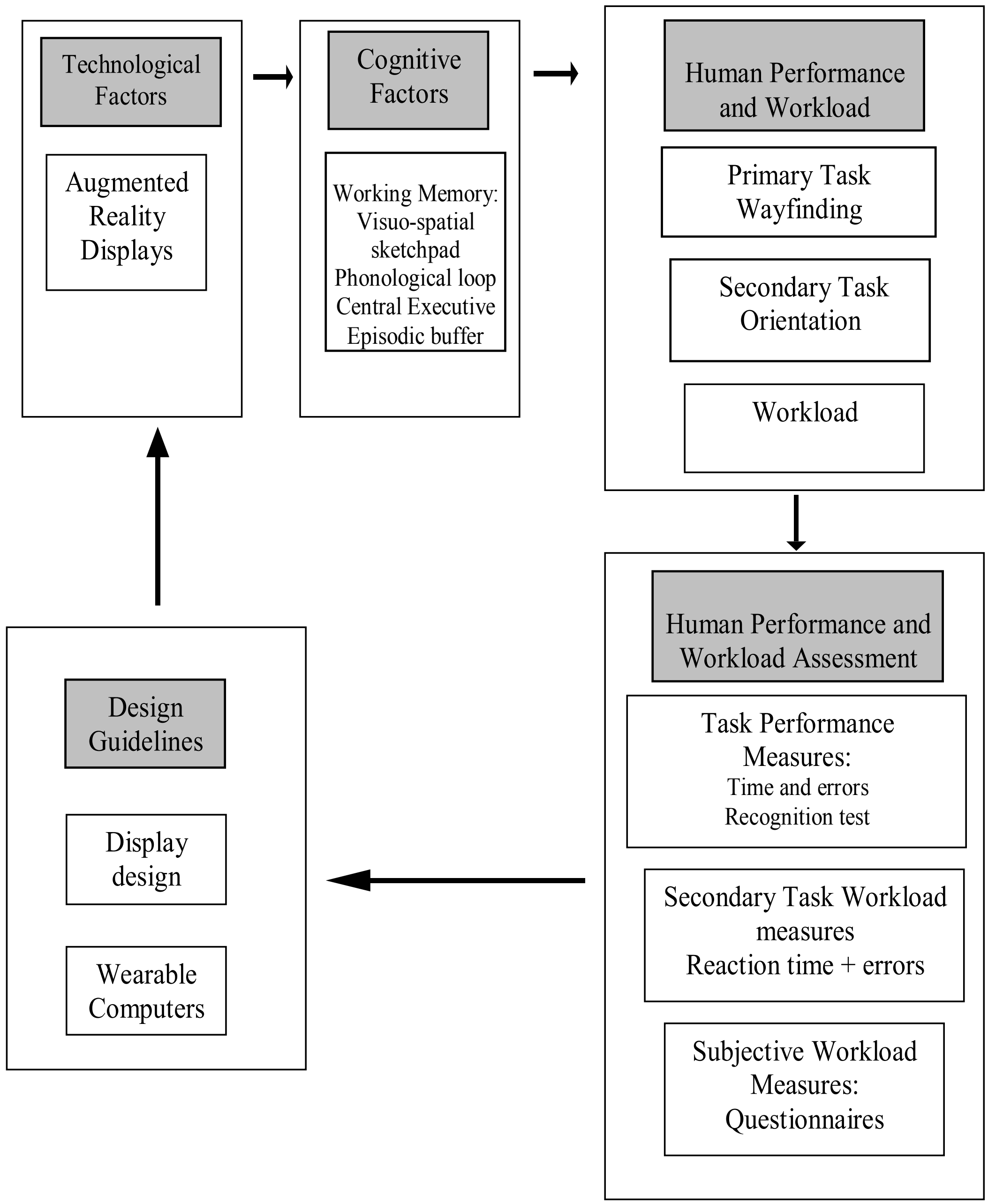


Figure 1: Conceptual framework for analyzing augmented reality performance

Technical Factors

AR technology and related developments are advancing at a rapid rate. Some functionality has been demonstrated with respect to manipulative tasks such as equipment repair and manufacturing. There are several technological hurdles being attacked that could greatly facilitate AR moving forward. These include work in tracking, registration, graphics, displays, and human factors. Tracking issues involve precision tracking over long distances with minimal external equipment and special markings. Registration involves insuring proper temporal and spatial alignment between simulated and real sensory inputs. Graphics and displays involve having the ability to smoothly integrate visual scenes from real and computer generated sources, especially as the real source characteristics dynamically change. Human factors issues involve form factor, usability, and designing for the proper information content, delivery, and display to optimize task performance. There is much research to do, in particular in the human factors area.

Augmented reality technology holds many promising applications, especially with the introduction of powerful PC technology. In the context of spatial cognition research, however, it is essential to determine which variables impact on the utility of AR for the many practical applications being investigated. Many studies have explored the types of stimuli and environmental exposure necessary in order to navigate in real and virtual environments, but none have investigated navigation and wayfinding in mobile AR. Furthermore, the cognitive factors that may mediate the effects of the information coming from stimuli and the exposure, have not received that much attention either. In short, there clearly exists a low level of technological sophistication necessary to fuse the virtual and the real and virtual in augmented reality. The issue is now to research how these factors individually and differentially impact the human information processing system and ultimately human performance in wayfinding. The components of the framework that describe the cognitive processes most likely to be impacted by the technology-enhanced environments and more specifically by mobile AR are discussed next.

Cognitive Factors

Recent research has demonstrated that working memory is composed of a number of subsystems, each responsible for uniquely processing inputs (e.g., Baddeley, 1986; 1992; Baddeley & Hitch, 1974). Additionally, it is worthwhile to investigate if external devices like AR systems contribute or become an extension of working memory. Figure 2 provides a graphical representation of a model of working memory. This figure has been modified by adding a linkage to long term memory, which may be accessed from the episodic buffer. This is consistent with interpretation by Haberlandt (1997) indicating investigations into additional linkages to working memory.

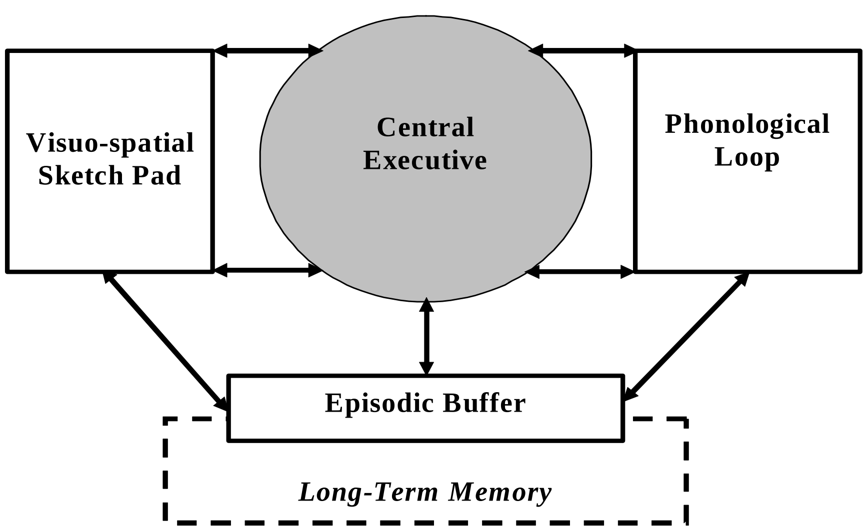


Figure 2: Baddeley's model of working memory

Through the use of this well-established model of working memory, a more productive approach for integrating human information processing into design and use of advanced technology can be achieved. This will allow system designers to appropriately and effectively develop augmented reality systems that can improve human performance in such environments. The working memory model is briefly described below.

Visuo-Spatial Sketchpad

What is generally characterized as "visual short-term memory" is known as the visuo-spatial sketchpad (VSSP) in Baddeley's model. This subsystem is particularly relevant to performance of complex tasks such as wayfinding because such tasks often require the integration of information that is either spatial in nature or visually based. The VSSP is thought to be responsible for encoding and maintaining information pertaining to the visual and spatial features of a given stimulus event. Behavioral studies illustrate that tasks requiring visual-spatial processing are interfered with from dual-tasks involving spatial information, but not verbal information. Research findings illustrate that attending to differing spatial locations hinders visuo-spatial memory (e.g., Smyth & Scholey, 1994).

Phonological Loop

What is generally characterized as "verbal short-term memory" is known as the phonological loop in Baddeley's model. The phonological loop is composed of a system for temporarily maintaining speech-based input (i.e., the phonological store) and a system for refreshing this information (i.e., an articulatory loop) as it decays (Baddeley, 1992).

Central Executive

Although the role of the central executive is the least researched component of Baddeley's model, it is thought to be responsible for allocating processing resources and in the ongoing revision or updating of information to the working memory subsystems (e.g., Baddeley, 2000; Kiss, Pisio, Francois, & Schopflocher, 1998). In essence, the central executive is suggested to be an "attentional control system responsible for strategy selection, control and co-ordination of various processing tasks" (Collette et al. 1999). As such, the primary role of this working memory system component is to both focus and divide available attentional capacity, as well as switch attention as needed, to perform the task at hand (Baddeley, 2001).

Episodic Buffer

Baddeley (2000; 2001) has expanded on the working memory model by suggesting the existence of a fourth component, the episodic buffer. This component may serve two main functions: 1) as the interface for combining information from the two separate subsystems (i.e., the phonological loop and the visuo- spatial sketchpad), and 2) as the interface between these working memory subsystems and long term memory. The episodic buffer allows for active maintenance and manipulation of multi-modal information such that integration of the differing representational formats from the two subsystems and long-term memory is possible (Baddeley, 2000; 2001; Baddeley & Wilson, 2002).

Human Performance and Workload

The nature of the wayfinding task being investigated and in many practical situations is at least dual and sometimes multi-task in nature. For example, imagine reading the newspaper as you stroll for a walk. In a dual-task situation, participants work simultaneously on a primary task in a certain modality, and on a secondary task in the same or other modality. Brooks (1968) showed that memory for visual information was diminished by visual responses, but not by verbal responses and that such interference provides evidence for the visual component of working memory within the dual-task paradigm. From all the components of the Baddeley's working memory model the least is known about the central executive (Haberlandt, 1997). However, by using the dual task strategy among other research strategies, new discoveries continue to be made. The same logic has been used to stipulate processes that place demands on the central executive. According to Baddeley's model some of the processes under the control of the central executive may include problem solving, comprehension, and reasoning activities. It is assumed that several secondary tasks and more specifically the random generation task produce decrements in the primary task of problem solving. Handling both tasks simultaneously is too much of a load for the central executive.

Aretz (1991) and Aretz and Wickens (1992) conducted studies on the utility of different map configurations by pilots when navigating. The work described in these papers is relevant to Baddeley's model because of the impact that mental rotations impose on working memory. The authors show that there is a cost imposed on working memory that is task-dependent. The task in this case represents the degree of mental rotation needed and whether the participant needs to determine relative location of points in space or cardinal directions of objects.

Mental Workload

Traditionally, mental workload has been assessed by using one of four different techniques: primary task measures, secondary task measures, physiological measures and subjective measures. The selection of a technique to measure workload is generally facilitated by the most measures that correlate with each other in discriminating low and high workload. However, the use of multiple measures is recommended where possible (Wickens, Gordon and Liu, 1997). The assessment of workload is relevant for this experiment since it can serve a very useful function of system evaluation and can contribute to the usability analysis of the system even if the performance with the system in question may be good (Wickens et al., 1997). However, if the workload experienced while using the system is excessive, the system may require improvement.

Secondary Task Methods

Secondary task performance provides a method of measuring reserve capacity (Wickens et al., 1997). If the performance on the primary task takes up a certain amount of cognitive resources then the assumption is that the secondary task will use whatever the residual resources are left. There are, however, some limitations associated with the use of secondary task that is not a part of the primary task (Wickens et al., 1997). These limitations can be avoided with the use of embedded secondary task, which is a task that is normally a part of the primary task but with lower priority. Within the context of Baddeley's working memory model, the use of embedded secondary task may provide some useful insights into how the different components of working memory are differentially tapped by the primary and secondary tasks. Furthermore, the combined influence of technological factors and cognitive workload factors may predict performance changes.

Subjective Measures

Subjective workload measures involve asking the operator to rate workload on a subjective scale. The best scales include explicit description of the high and low ends of the scale. For example the NASA Task Load Index (TLX) imposes five different subscales, with seven levels (Wickens et al., 1997). However, there are certain limitations as with most measures to subjective ratings. Even though easy to obtain, those measures are by definition subjective and do not always coincide with operator's performance.

CHAPTER TWO: REVIEW OF LITERATURE

Blending humans and machines has been in the dreams and nightmares of people for many years (Barfield and Caudell, 2001). The literature depicting experiences in blending the two can serve as a guide to the future.

Augmented Reality Systems

Augmented reality generally describes a technology in which computer graphics are overlaid on an actual scene. This general definition leads to several more definitive descriptions. Azuma (1997) defines augmented reality as a variation of virtual environments. Whereas virtual environments completely immerse the user, augmented reality allows the user to see the real world with computer generated images superimposed. Barfield and Caudell (2001) characterize augmented reality systems as overlaying computer graphics onto the real world scene through a helmet mounted display. These authors describe augmented reality systems as being tethered to a computer. Tethering reflects older and many current systems which require large computers or cumbersome graphics systems that require large amounts of power. A special case is combining wearable computers with augmented reality to provide mobility. Wearable computers are currently under development and provide performance approaching tethered systems while affording freedom of movement to the user. Both Barfield and Caudell (2001) and Azuma (1997) acknowledge that augmented reality systems allow information to be removed from the real world in addition to being added. They also acknowledge that augmentation may occur in additional sensory modalities, other than the visual.

Clark (2003) considers virtual and augmented reality from a more holistic perspective. Holistic in this context is expressed in qualitative terms from a synergistic viewpoint by considering the machine and human as an adaptable system. In particular, he views human information processing as not being bound by the physical brain, but including physical systems that augment or extend the processes of the brain. For example, a wrist watch may be considered as part of the human information processing system because of the information it provides that is available but not resident in the physical brain. A query, "Do you know what time it is?" is met with an affirmative answer followed by looking at the watch. In this sense, one may infer that cognition as described by Wickens and Hollands (2000) influences perception. Whereas, Wickens and Hollands (2000) describe a feedback loop where perception feeds cognition, Clark (2003) infers that the loop can go the other way; that is what one perceives is influenced by ones cognitive state. A concept of co-evolution which is pertinent in describing and tracing the origins of augmented reality is introduced by Clark (2003). Co-evolution refers to a temporally extended period of time when technology advances and individuals and society adjust to those changes often times resulting in minor or major changes to the technology. When extending this concept to augmented reality as Clark (2003) does, then one can look at earlier devices that extended human perception. Examples include the telescope, binoculars, but more relevantly eyeglasses. One can also imagine where today's early versions of augmented reality will influence and be influenced by human usage.

The earliest contemporary augmented reality system using computer graphics for visualization was actually created by Ivan Sutherland in the 1960's (Barfield & Caudell, 2001). Although Sutherland is credited with creating the first virtual worlds, he used an optical see-through head-mounted display (HMD) so his work was actually in augmented reality. Most work growing from Sutherland's research breakthrough, though, was focused on creating virtual environments until a little over ten years ago.

Clark (2003) points out that augmented reality had its first practical use in the early 1990's by a group of engineers and scientists at Boeing. The intent of exploring the use of augmented reality by Boeing was to help facilitate the installation of wiring in an aircraft. The workers would see the desired positioning of wires superimposed on the actual aircraft. The benefits of augmentation would be in facilitating manipulative tasks and reducing the need for paper drawings and guidelines in the production facility. These items would be available through computer augmentation.

Additional uses of augmented reality are being suggested as the technology matures. The literature suggests functional groupings of augmented reality uses in manipulative tasks, decision aides, and navigation. Azuma (1997) discusses several potential uses for augmented reality for medical, manufacturing, annotation and visualization, robotic path planning, entertainment, and military aircraft. For example, it is conjectured that medical applications could include assisting surgeons in seeing what might not be detectable with the naked eye. Manufacturing could include, for example, wiring guidelines as reported by a Boeing technical research program in Azuma (1997). A more contemporary form of augmented reality related to annotations and visualizations are finding their way into modern media with computer generated advertisements added to a real world sporting event background. In military aircraft head up displays and helmet mounted sights are examples of augmentation (Azuma, 1997). The above examples indicate that augmented reality has many potential applications that have been conjectured, prototyped, but not yet fully explored.

In a study on the features and use of augmented reality, Klinker and Ahlers (1997) provide an overview of augmented reality and its major components. They discuss the needs of precise models of the environment so that items can be merged seamlessly. The authors also discuss the need for realism in various forms; merging of graphics and background, graphics quality, and physical properties such as occlusion. However, the main focus is on the technical aspects of computer vision to detect objects. They have developed a system called GRASP, which is a real time video mixing system that combines video of a real scene with computer generated graphics. The level of maturity of GRASP is not specifically discussed. The authors, though, appear to have captured the salient requirements for proper augmented reality and project its usage in a variety of applications. GRASP is projected to apply to situations where several augmented reality systems may be spatially distributed, potentially over large areas. In this vein, the authors propose a replicated architecture and databases to facilitate timely retrieval of data. Other technical aspects of augmented reality, such as latency, interface devices, and tracking are discussed by Barfield and Caudell (2001), Azuma (1997), and Azuma, et al. (2001).

Some of the technical areas that constrain augmentation include display limitations, graphics, tracking, registration, mobility and the user interface. Display limitations occur in both video and optical see-through devices. Graphics are limiting from two distinct vantage points all related to rendering limitations. First, while the processing power of graphics cards are increasing rapidly, the portability of the graphics pipeline is proceeding at a somewhat slower pace resulting in either tethering of the user to a fixed computing platform or limitations in scene content. The second constraint is related to the amount of scene content that is actually needed. The composite visual scene should minimize clutter or be an integrated scene (McGee, 1999). Minimizing clutter requires careful investigation into what should be in the augmentation to facilitate task accomplishment. Likewise, fusing graphics with real world background requires significant advances in rendering quality such as texturing, lighting effects, radiosity, and polygon count.

Tracking also constrains augmented reality functionality. Generally, precise tracking can be accomplished over small distances (e.g., the movement of a surgical instrument) and becomes less precise over larger distances. Outdoor tracking over areas that are not instrumented or mapped in the computer is currently not possible. Registration has similar limitations to tracking. Whereas tracking provides information on where someone is located, registration uses the tracking information to ensure that images, ones position, and displays are properly aligned. Latency is a major issue constraining registration accuracy in high movement environments.

Mobility and user interfaces are the final items of interest. As previously mentioned, technical performance and mobility in augmented reality graphics as well as other display modalities are generally inversely proportional. The more mobility one desires, the less technical performance is generally available. Likewise user interfaces and performance generally have an inverse relationship. Lighter and less obtrusive equipment which connote a better user interface generally result in poorer technical performance. Rapid technological advances are tending to ameliorate these issues.

The limitations in augmented reality, discussed above, are being addressed in various research laboratories and help define a capability useful for exploring usability issues in various tasks. For example, studies in the benefits of different augmented realty display systems for evaluating wayfinding are limited by the available technology and should consider these limitations. The results of task specific assessments in augmented reality can help formalize notional ideas on the benefits of augmentation, and also help guide technological developments.

Full exploration of augmented reality involves understanding both its technical and human utility aspects. As previously stated, most activity to date in augmented reality has been in needed technological advancements. Azuma (1997) reports on many technical obstacles needing to be overcome for augmented reality to reach its full potential. These include registration of images and the real world, tracking, time lags, and portability. Also highlighted are issues with integrating other sensory displays, such as audition, and social or political issues regarding access to information that might not be available to others. Finally, Azuma (1997) and Tang, et al. (2003) generally point to the need for assessments in human performance when using augmented reality systems. These assessments are needed to better understand the utility of current augmented reality systems and can help guide the research community in addressing the most pressing needs. Broad based assessments will surely uncover new uses of augmented reality as well as help hone the areas where augmentation has shown some benefit.

To date, there have been limited evaluations of augmented reality utility with users. Two noteworthy exceptions are the works by McGee (1999) and Tang, et al. (2003). McGee (1999) explores techniques to form an integrated mental model of the merged real and computer generated images by fusing the visual scenes together spatially and through merged graphics. Merged graphics imply that images blend smoothly in color and texture. Tang, et al. (2003) conducted tests showing the benefits of manipulative and spatial (about a fixed table) tasks in assembly. The authors explore building structures out of Lego blocks. They discuss how various forms of augmented reality can reduce attentional switching between instructional media and the target system by placing information together, facilitate spatial docking of disparate information that supports spatial cognition, and eliminate short term memory demands by spatial superimposition. Tang, et al. (2003) evaluated these concepts in a manipulative and spatial task involving strategies for assembling Lego blocks. Their results regarding the benefits of augmented reality as opposed to other computer based instructions were somewhat mixed. It was noted that visual resolution limited some tasks. It was also found that attentional switching was reduced with augmented reality, but at a cost of cognitive tunneling, possibly due, in part to needed spatial rotations by users.

Other augmented reality evaluations are reported by Baird (1999) and Feiner, MacIntyre, Höllerer, and Webster (1997). Baird (1999) evaluates different types of instruction in assembly tasks. Four conditions are evaluated. One condition is with an opaque augmented reality display and another condition is with a see through augmented reality display. Baird reports benefits with augmented reality in terms of reduced errors and time to task completion and also reports important negative usability aspects of augmented reality as compared to other delivery strategies. Negative aspects are with respect to comfort and poor contrast. Feiner, et al. (1997) discussed software aspects of a navigational augmented reality system. The authors concentrate on software issues related to strategies for labeling buildings while navigating through the Columbia University campus. This is an important technical work and gives some general insights into usability and navigation using augmented reality. However, no principled evaluation of utility or strategy for human navigation using augmented reality has been found in the literature.

Augmented Reality's Position in Reality Continua

When considering usability in general and navigation, in particular, in the context of augmented reality, it is beneficial to investigate works in more established fields such as virtual reality because some accomplishments in virtual environments can be useful in augmented reality. Witmer and Sadowski (1998) explore how comparative performance in real and virtual environments (VE) can provide clues to needed enhancements in VE technology. The authors use a distance estimation task experiment to support this idea. The results of their experiments demonstrated better performance in estimating distance in the real compared to the virtual world. The experimental results provided important insights into needed improvements in binocular displays. The importance of this work is the global view of how comparisons between the real and virtual environments (or augmented reality in the case of this dissertation) can provide insights into technical improvements needed to enhance human performance. This strategy should be extendable to augmented reality systems. That is, with appropriate care for relevance, the needed enhancements to augmented reality systems can possibly be identified from evaluations of similar tasks in real or virtual environments.

However, it is also important to understand the differences in augmented and virtual realities because some systems, strategies, or studies might only be relevant to one technology. Augmented reality can be considered as a component of a continuum of various types of reality available to the human. Milgram, Takemura, Utsumi and Kishino (1994) describe a taxonomy that depicts the relationship between reality, virtual reality, and augmented reality. Figure 3 from Milgram, et al. (1994) portray this relationship that can be called the Mixed Reality spectrum. The spectrum is anchored on one end by reality and the other end by virtual reality.

Figure 3: Mixed reality spectrum showing relative placement of various types of reality

Virtual reality, virtual environment, virtual presence, and artificial reality are all similarly defined by Sheridan (1992) as something experienced by a person when sensory information is only generated by and within a computer and delivered through display technology. The author goes on to describe that the sensory experience must compel the individual to feel that they are present in an environment other the one they are actually in. He conjectures that in an ideal sense and with sufficiently good technology, an individual would not be able to distinguish between actual and virtual presence. Finally, Sheridan views the terms virtual reality and artificial reality as fashionable, but linguistically contradictory and recommends using virtual presence and virtual environment as better terms.

Virtual reality is summarized by Barfield and Caudell (2001) as using a helmet mounted display to provide an immersive representation of a computer generated simulation of a virtual world. The user does not view the real world. All sensory information provided to the user is mediated through the computer. External sensory cues are controlled by the virtual environment. For example, headphones are typically used to reduce external uncontrolled sounds as well as to add the environment created sounds.

The Mixed Reality spectrum includes augmented reality and augmented virtuality. Milgram, et al. (1994) describe three characteristics of mixed reality as 1) reproduction fidelity,

2) extent of presence metaphor, and 3) extent of world knowledge. Reproduction fidelity is used to qualitatively describe the graphical quality of the imagery and ranges from simple wire-frame to elaborate renderings. This area has been technically challenging in augmented reality systems due to processing limitations in meshing computer generated images into a complex and dynamically changing real world and mobility introduces additional computational power limitations. The presence metaphor is used to describe the degree to which the user is cognitively engaged in the total scenario composed of real and computer generated components. This factor has generally been low or zero in augmented reality systems because of lack of guidelines for creating the sense presence in all types of systems. The extent of world knowledge factor describes the extent that mixed reality systems have knowledge of the world they are depicting. In augmented reality systems the extent of world knowledge is a particularly difficult area technically because of the need to properly spatially register the augmented and real environments. These distinctions between augmented reality and augmented virtual realities will become blurred as processing power increases and as new sensory modalities are introduced.

Another way to conceptualize the area between reality and virtual reality depicted in Figure 3 is to consider the predominant type of reality being presented to the user. If the predominant reality being presented to the user is virtual, then adding some reality results in augmented virtuality. An example of augmented virtuality is the Pit Room described by Meehan, Insko, Whitton, and Brooks (2002). In the case of the Pit Room, the user is immersed in a virtual environment and passive haptics in terms of a small physical ledge reinforces the sense of being on a precipice over a room. Similarly, a Head Up Display can be considered an example of augmented reality, because the primary source of sensory stimulus is the real world.

Building upon the reality-virtuality continuum, one can conceptualize a different way to portray the holistic approach to augmented reality described by Clark (2003), the uneven technical advances to the field, and the variety of blending of virtual and augmented technologies. This dissertation suggests a new continuum as portrayed in Figure 4. This figure can be interpreted in the following way. First, only if areas of augmentation and perception intersect, is there an opportunity for reinforcement, expansion, or interference of human perception and sensory augmentation. The representation shown in Figure 4 is an example reflecting perfect interaction and an expanded reality between the augmentation and unaugmented reality. If there is an intersection between the two realms, the area that is larger predominates. This results in augmented reality or augmented virtuality. In another circumstance, if sensory augmentation completely surrounds human perception, without intersecting the area one has virtual reality (human perception is controlled through augmentation). Areas that abut but do not intersect form the augmented reality or augmented virtuality area within the anchors of Milgram's Continuum. Intersecting areas result in conflicting sensory cues to the participant. The approach taken puts augmented reality on both sides of the reality anchor of Milgram's continuum reflecting the additional sensory information that augmentation can provide.

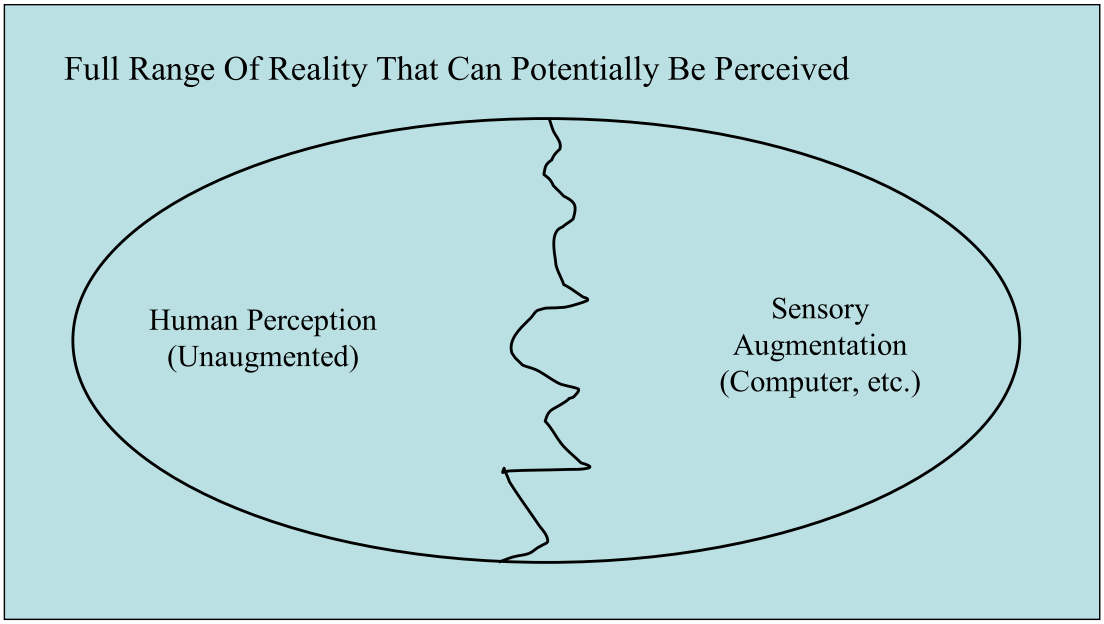


Figure 4: Venn diagram conceptualization of the interaction of various types of reality

Augmented Reality Display Technology and Human Performance

Current implementations of augmented reality systems deliver augmentation principally through the visual channel (Barfield and Caudell, 2001). It is important to review the different approaches to displays and discuss issues with them, because displays have been an area of technical interest and because many evaluations of augmented and virtual environments as well as wayfinding in virtual reality center on visualization. An overview of displays in augmented reality can be found in Rolland and Fuchs (2001). The two principal approaches in displays are video or optical see-through. Both approaches merge imagery from two sources. In video, images of the real world are merged with computer generated images electronically. In optical see-through, the images are merged optically. Both approaches have advantages and disadvantages. Video see-through offers the opportunity for better integration of real and electronic images because they are combined electronically and can provide some level of control for registration problems arising from latency. Optical systems have no means to control latency when trying to merge computer generated graphics with a real scene that might be moving because the user is moving. Video see-through, however, tends to be more physically cumbersome than optical see-through systems. Other issues are field of view and resolution. Optical see-through performance in both resolution and field of view is generally better than video. The point of this discussion is that there are trade-offs to consider with respect to the task being performed and the various effects on the human user of augmented reality displays.

Head up displays optically merge electronic images with a real world scene. McGee (1999) provides a good summary of head up displays as adding symbology in the forward field of view of the pilot. McGee (1999) distinguishes head up display from augmented reality as one involving approach rather than technical definition. Head up displays are often meant to be used in place of regular instruments. Augmented reality prefers to enhance and add to a real scene rather than replace information. McGee (1999) points out that head up displays are intended to clearly separate the real world and not intended to be integrally perceived. He discusses that a major issue in head up displays is the attentional demands caused by the display.

Wickens and Hollands (2000) and Carswell and Wickens (1987) take a different point of view from McGee (1999). They both look at approaches to have head up and other types of displays more integrated to mitigate attentional issues. However, McGee (1999) takes a point of view of perceptual fusion while Wickens and Hollands (2000) consider fusion of objects cognitively. Wickens and Hollands (2000) discuss spatial proximity from a spatial and perceptual point of view. They infer that one should link objects in a display if they have strong task relevance. Careful considerations must be taken in establishing these relationships from a cognitive perspective. McGee (1999) concentrates only on the spatial and graphical characteristics of displays as providing linkage. Wickens and Hollands (2000) and Carswell and Wickens (1987) approach, though riskier because design guidance is less specific than McGee (1999), seems to be more inclusive in scope and consistent with Clark's (2003) holistic approach that acknowledges the plasticity and adaptability of human behavior.

There are other aspects of augmented reality relevant to human performance issues. An excellent summary of recent advances is provided by Azuma, et al. (2001). They mention persistent problems with displays such as the dynamic range of brightness, limited field of view, and fixed contrast ratio that restrict the blending of real and synthetic images. However, displays for gaming are making advances across some of these areas while keeping the cost per unit reasonable for consumers. Likewise, support for occlusion in optical displays has been prototyped by including special LCD panels in the optical path. Display sizes for augmented displays are also decreasing with eyeglass type of displays becoming available.

Tracking technology is needed for effective augmented reality. Tracking is needed in order to represent proper occlusion, for interacting with real and virtual objects, and to support graphics and other sensory processing. Indoor tracking is now available, but expensive. Precise outdoor tracking is more difficult to achieve unless the outdoor area is prepared with fiducials, differential GPS, or precision maps. Azuma, et al. (2001) point to improved user interfaces being prototyped that allow for more mobility (untethered) and better form factor for the user. Other user interfaces include physical devices that facilitate manipulation of virtual items as well as techniques to facilitate interaction between users. They believe more studies are needed to help guide and prioritize augmented reality development.

Human Performance Characteristics in Augmented Reality

In considering human performance in various augmented reality systems, one can use an analogy from control systems design where the human and equipment form a closed-loop feedback system. The literature is focused on visual displays and how different display strategies affect various aspects of human performance.

Display Performance

Human performance is dependent, in part, on the technology and the task. Wickens and Hollands (2000) present a model of human information processing. The model is a feedback model of sensory processing, perception, cognition, and response. Cognition involves long term and working memory and attentional resources. Information presentation affects attention and memory, which impacts spatial knowledge. There is a large amount of variation in visual and display related content possible between the real world, virtual reality, and augmented reality that can affect sensor processing. The following literature discusses considerations and findings that bear on how an augmented reality display system should be designed to optimize human performance.

The literature provides some keen insights into merging various types of visual information as one moves from considering display hardware to their contents. The contents of displays can affect the user's attentional characteristics so care must be taken in the display design (Wickens & Hollands, 2000). They provide an overview of many issues needing to be addressed in properly designing displays. Concepts discussed include avoiding divided attention through proper display design, organization of like features on the display, the concept of spatial proximity and compatibility of spatial proximity with the demands of the task. Head up displays (HUD) are special examples of employing spatial proximity. The authors also introduce a concept of conformal symbology that is also pertinent to augmented reality. Conformal symbology is characters that are spatially fused with the real world scene. The use of conformal symbols helps reduce divided attention between the symbology and the background scene by keeping augmented and real spatial information aligned spatially and perspectively. McGee (1999) confirms the benefits of conformal, or as he calls them fused, symbols in certain augmented reality systems. McGee (1999) goes one step further in investigating and finding the benefits of computer generated graphics of relevant objects that become cognitively fused with the real world background. He also cites the relevance of looking at Head Up Displays as a specific case of augmented reality. His work is oriented towards cognitive tasks.

In related research, Edgar and Reeves (1997) studied the effect of clutter from the HUD on the far field and divided attention between the HUD and far field. The application environment was flight. The authors point to the advantage of close placement of information for reduced response times in aircraft. Differences in registration of less than 6.3 degrees and separation of similar information of up to 22.5 degrees, though for secondary tasks did not affect response time. However, users tended to fixate at one image or the other, an effect known as cognitive tunneling or cognitive fixation. Edgar and Reeves (1997) point to research by Weintraub and Ensing (1992) indicating that conformal characteristics of HUD imagery can mitigate divided attention and increased workload. Conformal characteristics are those HUD symbols that serve as a virtual analog to far field items. Edgar and Reeves (1997) also point to the work of Fischer (1979) where symbology does not adversely affect pilot attention to the far field, the reverse is not true. Conformal displays are facilitated by coherent and consistent motion and continuity between graphics and real world far field. These findings for far field displays can be used as a starting point for further investigations in the field of augmented reality where far field range is reduced in many situations such as wayfinding in a building or maze.

Edgar and Reeves (1997) conducted experiments under a variety of visibility conditions (full and zero visibility). They investigated response times in varying conditions of HUD clutter by introducing an unexpected discrete event. The HUD used has a field of view (FOV) of 15.4 degrees by 9.4 degrees (which is smaller than many monocular displays, such as one produced by Micro Optical Corporation which is 16 degrees x 20 degrees). Edgar and Reeves (1997) varied instrument position (head up versus head down). Their results indicate that in almost all cases head up is better than head down for minimizing divided attention (the only exception is when unexpected events occur). A second experiment by Edgar and Reeves (1997) investigates the benefits of conformal displays. For displays that are overlayed onto a background scene, there is an apparent accommodation effect even when the overlay and the background are set at the same distance. If the overlay image is not perceived as integral with the background, it has the effect of appearing closer than the background. The background foreground effect can result in divided attention.

Ellis and Menges (1998) examined errors in localization of nearby virtual objects in helmet mounted displays. In this regard they also disclose some useful ideas for designing helmet mounted display hardware and software. The suggestion is made that accommodation might be a useful depth cue. They also offer a concept where the display's accommodation distance can be set to important physical distances to provide a possible point where information might be ideally viewable offering an integrated display of real and virtual. Three display conditions: monocular, binocular, and stereoscopic are investigated. Binocular uses two exact images to avoid potential problems with visual rivalry. A rotating virtual pyramid was used as a stimulus and was positioned 58 cm from the user. Users were asked to position a cursor under the nadir of the pyramid. The results of experiments showed stereo and binocular with similar levels of performance. Monocular performance was inconsistent. One result that is of interest is that monocular performance seemed to drop back to the background wall that was 2.2 m from the subject. In a second experiment a rotating checkerboard was used with similar results. Given these experimental results by Ellis and Menges (1998) some of the relevant recommendations include:

1. Head movement of the subjects might introduce affects that were not tested. In particular, motion parallax effects could change the results.
2. Weight and cost considerations may encourage use of monocular displays. If used, monocular displays should allow for variable focus to accommodate a variety of user needs.
3. Biocular and stereo displays require bore sighting by the user prior to use.
4. Computer generated images or other images displayed binocularly should have tailored stereo disparity.

Research conducted by Edgar and Reeves (1997) also investigated visual accommodation when information is presented visually on a HUD or HMD and the real world aviation scenes. HMD's are being introduced and have similar characteristics as the HUD in an aircraft. The authors point to the mixed opinions of whether accommodation issues arise in these systems even when the computer generated images are projected at an apparent infinity. The results of studies show a wide range of effects depending on ambient conditions, but do not point out whether conformal displays are being used. These results indicate caution should be exercised in designing and implementing head up or helmet mounted displays. A degree of dynamic variability in display contents might be indicated depending on the task being performed and ambient conditions.

Martin-Emerson and Wickens (1997) studied conformal symbology, visual attention and superimposition in HUD versus head-down display performance. There is a similar relationship between computer generated maps used in AR and regular paper maps with respect to performance. However, adjustments are required for this dissertation because of differences in task complexity and the fact that participants will be traversing the space. Also, Martin-Emerson and Wickens (1997) used a HUD where the displayed information was superimposed onto the visual scene; information content added to the real scene in this dissertation is collocated and spatially registered, but not with the background scene. The research conducted by Martin-Emerson and Wickens (1997) compared superimposed HUD with an identical display in head-down position in varying visibility conditions. They investigated the extent that the characteristics of the HUD symbology support a division of attention by contrasting conformal symbology with traditional ILS symbology. The results showed minimized scanning between flight instruments and the far domain contributes substantially to performance advantage. Their hypothesis was that if the symbology forms an object with the far domain, attention may be divided between the superimposed image and its counterpart in the far domain was supported. Conformal displays in this context involve perception. The augmented reality approach being studied in this dissertation will use fixed map conformal displays from a cognition point of view in the augmentation's role in facilitating task performance.

Furmanski, Azuma, and Daily (2002) study and present human factors design guidelines for spatially representing augmented scene content in a real environment. The authors use a cognitive approach to develop guidelines for visualizing occluded objects in augmented reality. The authors are not concerned with fusing information cognitively, but are interested in strategies for retaining information that may become occluded by adding computer generated information, for allowing the user to determine which information is closer than other information, and for keeping information content uncluttered. The authors present some excellent guidelines for monocular or binocular displays as well as for moving or static images. The guidelines when using monocular systems include the selective use of transparency, occlusion, size scaling, texture, shadow gradients, and cross-referenced depth information. An important guideline for all types of augmented reality systems is to eliminate unneeded motion of the image, such as what might occur due to registration jitter. Pilot studies are conducted to assess some of the guidelines in a static setting. Post experiment findings supported most of the guidelines. One interesting outcome that should be considered in designing augmented reality systems, though, is that subjects tend to use occlusion as a primary cue for spatial relationships even when other cues present more relevant information.

Map Configurations

There is a great deal of information on map orientation and its relationship to wayfinding. Modern maps are either egocentric or exocentric, corresponding, respectively, to the users forward orientation being up (i.e., the map moves under a representation of the user) or a fixed position, normally north being up (i.e., the map orientation is fixed and the user's orientation moves). The literature synopsized below considers both points of view and seems somewhat mixed with respect to benefits in real world settings. However, careful investigation of the literature seems to suggest that an egocentric map representation is better when the user is directly represented and participating in the environment. An exocentric representation seems to be the better choice when the area being mapped is large or when the user is observing or a more passive participant in the environment. However, the literature is silent for situations where learning is not an explicit part of the task or when the environment has no specific enroute features. These situations can occur, for example, in a rescue situation.

Arthur (1996) reported that humans tend to create cognitive maps of areas based on how spatial information is obtained. There is an alignment effect when spatial data is obtained from a map as reported by Evans and Pedzek (1980) and Thorndyke and Hayes-Roth (1982). The alignment effect results in storing a fixed orientation of the space in memory. Participants then must mentally rotate their current position to align with the stored representation. Gillner and Mallot (1998) studied the degree of misalignment and reported on increasing times to align map representations as misalignment angles increased. On the other hand, mental representations acquired through navigation in an environment did not exhibit the alignment effect (Evans & Pedzek, 1980; Thorndyke & Hayes-Roth, 1982). However, these authors report increased errors for routes with many turns than simpler routes when navigation is the means of acquiring spatial knowledge. Their work confirms research conducted by Tarr and Pinker (1989) who proposed the multiple view theory, whereby humans accumulate multiple views of a space by navigating the space. The multiple view theory also suggests an increase in judgment time as the amount of mental rotation of the stimulus increased.

Arthur (1996) also reported that navigation in a virtual environment contains aspects of primary and secondary learning. Primary learning is acquiring information directly from the environment. Secondary learning is acquiring information from abstract sources such as a paper map. Using paper maps result in the acquisition of an orientation specific view, but that limited view only has an effect if learning is involved. That is, performance as measured by time to travel from a start point to an end point was effected if the map was learned. This dissertation asserts that augmented reality systems can have similar characteristics if the augmentation uses a map. Interestingly, Arthur (1996) confirmed Tarr and Pinker's multiple view theory in some virtual environment experiments involving 3D spaces, but was unable to attribute the theory to the number of perspectives available during acquisition of spatial knowledge in a virtual environment. Review of Arthur (1996) indicates issues with scene content resident in the virtual environment. There are some useful further investigations that can be performed in augmented reality where different map displays can be present in a real environment providing insights into the effects of primary and secondary learning.

Arthur and Hancock (2001) explored the use of maps, north up, and track up (the latter two, respectively, representing exocentric and egocentric views) in virtual environments. The authors point to findings from studies (such as Evans & Pezdek, 1980; Kulhavy, Schwartz, & Shaha, 1983; Presson, DeLange, & Hazelrigg, 1989; Presson & Hazelrigg, 1984, Thorndyke & Hayes-Roth, 1982) indicating that knowledge acquired from maps is qualitatively different than knowledge acquired by real world navigation and that learning from maps (secondary learning) leads to alignment effects where the traveler must perform a mental rotation to align the paper map with the mental map. Acquisition of knowledge from physical exploration (primary learning) does not result in alignment effects. Arthur and Hancock (2001) also point to the work of (Endsley, 1995; Smith & Hancock, 1995) indicating similar results with track up versus north up maps. Experiments conducted by Arthur and Hancock (2001) in a stationary setting confirm the benefits of navigation to acquire multiple views of an environment. This dissertation will confirm these findings in a mobile augmented reality setting where navigation occurs in a real world with several map options.

Research conducted by Hooper and Coury (1994) was concerned with the design of periscope workstations that display orientation information. The results of this research are relevant to this dissertation for several reasons. First, the authors suggest the results of similar studies on aircraft display principles and specifically those proposing that there are two reference frames for navigation: an ego-centered and world-centered (Aretz, 1991), are especially relevant to submarines. However, a very similar dichotomy also exists for displays used in any domain concerned with navigation and wayfinding. The results revealed that the judgments of orientation were best with north-up displays.

In research conducted by Aretz (1991), an experiment with 18 male pilots compared a map display that employs the principle of visual momentum with 2 traditional approaches, track-up and north-up. The data show that the advantage of a track-up alignment was its congruence with the ego-centered forward view; however, the experimental results demonstrated that the inconsistency of the rotating display hindered development of a cognitive map. Further studies conducted by Aretz and Wickens (1992) identified alignment issues between the egocentric and world reference frames (the latter being an example of an exocentric reference frame) resulting in the need for mental rotation as a central operation. Visual momentum (Woods, 1984) is the process by which computer display users integrate data across successive displays. Visual momentum codifies the integration process through display design guidelines for effective distribution of user attention. Wickens and Hollands (2000) report similar results when using egocentric displays that seem to be oriented and rotate in a different direction than self motion. The stability of a north-up alignment aided the acquisition of a cognitive map, but there was a cost associated with the mental rotation of the display to a track-up alignment for tasks involving the ego-centered forward view. The visual momentum design captured the benefits and reduced the costs associated with the traditional approaches. Results support the conceptualization of navigational awareness as a cognitive coupling between the perceptual view of the world (the ego-centered reference frame) and a map display (the world-centered reference frame).

The results of another study conducted by Wickens, et al. (1996) show a general advantage of rotating maps; however the authors suggest that fixed-map options be eliminated from user-selectable displays. Furthermore, they recommend that there are circumstances in which fixed-map displays may be preferred, for example, for planning. The conclusions are consistent with earlier work by Harwood and Wickens (1991) addressing the capability of different frames of reference for facilitating navigational problem solving. Matching of a particular frame of reference to certain task demands, they suggest, may increase the effectiveness of the map display for wayfinding. The pattern of map-task dependencies revealed by the results of the study suggests that no single map configuration is superior across all navigation tasks. According to Harwood and Wickens (1991), navigation and wayfinding tasks are complex and different task components can be supported by either north-up or track-up map configurations.

In another study, Werner and Schmidt (1999) found egocentric orientation of self and maps facilitate navigation. The authors point to numerous studies where allocentric (or exocentric) orientations, especially hierarchical ones lead to distortions on absolute position. The authors point to the work of Sholl (1995) who indicates that spatial information is stored in an allocentric reference system, but retrieved by overlaying an egocentric position that is dependent upon the user's physical and mental orientation. View oriented representations are tied to a specific location and heading and retrieval time is dependent on the degree of realignment needed. Map orientations are easier to comprehend when the user and map have similar orientations. The authors say nothing, though, about the impact of map orientations when one is moving, which can be disorienting (Wickens and Hollands, 2000) when the map and subject do not move in the same direction. This positive response is diminished for remote areas that are learned on a map and not familiar or visible to the subject.

Arthur and Hancock (2001) also conducted experiments on navigating by creating multiple mental models of spaces in real and virtual environments. The authors point out that spatial knowledge obtained from maps is different than that acquired from the real world route learning which is considered superior based on experiments. Route traversal allows one to obtain multiple views and therefore multiple representations of an area. Multiple representations are consistent with Tarr and Pinker (1989) studies. Part of the reason for superior performance in route learning could be based on kinesthetic cues not available in a map. Virtual environments offer the same benefits (but possibly not the same amount of benefit) albeit without full kinesthetic fidelity. Map learning is considered secondary learning. Multiple turns, though, can result in more errors in memory based route knowledge than using a map. There is an open research issue brought forward in the paper, though, in determining the optimal number of representations.

Navigation Using Augmented Reality

Navigation or wayfinding is a relevant application domain for augmented reality systems (Feiner et al. 1997; Azuma, 1997). Whereas virtual environments must provide the complete sensory environment for tasks such as wayfinding, augmented environments can provide maps or directions that can facilitate wayfinding while still retaining the needed sensory inputs to the user. Wayfinding is a pertinent area to investigate using augmented reality because prototypical systems are just becoming available that support mobility. Because the technology is just emerging, it is useful to consider approaches and issues with navigation in the real or virtual worlds.

Different terms have been used in the literature to describe navigation, wayfinding, and route learning, but all of them generally describe how people get from one point to another in a real or virtual environment. For example, navigation is defined as the aggregate task of wayfinding and motion, where wayfinding is the cognitive element of navigation and motion is the "motoric" element of navigation (Darken & Peterson, 2002). Navigation is a process inherently cognitive in nature and a good understanding of spatial knowledge and its use may prove beneficial to the design of training programs for large-scale system architectures. This dissertation involves a particular type of wayfinding that will be designated search and rescue. Search and rescue navigation involves searching an area to reach an objective, finding the objective and then exiting through a route covered while reaching the objective. Search and rescue is typical of what might be found in various public safety or military applications.

A widely accepted model of spatial knowledge acquisition during navigation has been the Landmark-Route-Survey (LRS) model (Colle & Reid, 1998; Siegel & White, 1975). This model defines three different spatial knowledge representation types. Landmark knowledge can be described as descriptive information of noticeable places within an environment. Route knowledge consists of knowledge of mentally defined routes between locations. Finally, survey knowledge can be defined as spatial knowledge in the form of a mental map of an environment.

These mental maps are thought to be metric representations of locations and objects that are in the environment (Colle & Reid, 1998).

Siegel and White (1975) suggest that the different knowledge that is gained from an environment occurs in sequence. First, knowledge is gained about landmarks that are visible. This knowledge allows people to recognize places but doesn't give them any information that helps them to traverse between the landmarks. Next, landmarks are located in reference to each other. This relationship between landmarks allows people to develop and learn routes through the environment. This route knowledge can be described as topological and consists of paths which use the previously gained landmarks as decision points and markers to reach specific sites or destinations (Colle & Reid, 1998). Over time, with support from the procedural knowledge that has been gained, a mental map of the environment can be constructed. This mental map is considered survey knowledge of the environment and contains metric representations of locations and objects (Siegel & White, 1975).

Although the LRS model of wayfinding is widely accepted, there are theories that build upon it and contradict Siegel and White's suggestion that different types of spatial knowledge is gained sequentially. Colle and Reid (1998) describe the dual-mode model, which proposes the idea that survey knowledge of local regions can be quickly gained. They assume that the local region is one that is within the immediate viewing range of the user and partitionable from other areas. The dual mode model assumes that there are two modes of spatial knowledge acquisition (Colle & Reid, 1998). The two modes are the gaze viewing mode and the route tour mode. Unlike the LRS model, the two proposed modes in the dual mode model can operate in at the same time and can both operate at any user experience level.

The gaze viewing mode is utilized when people acquire a spatial representation of the current local area. A local area can be defined as anything in the spatial span, which can be seen from the observer's viewpoint. By rotating the head or body, an observer can create an exocentric spatial map of this local area. The spatial knowledge acquired in this mode is referred to as direct imagery, or "a reconstruction of the three-dimensional perceptual space acquired from viewing objects within the spatial span" (Colle & Reid, 1998). The spatial knowledge gained in this mode is metrically coded and appears to represent Euclidean angular directions and distance.

The route tour mode is engaged when an observer is moving throughout a larger region of space. In this mode, participants don't acquire the knowledge of object positions that are acquired in the gaze viewing mode. Instead, they gain topologically coded knowledge of how to get from one place to another by connecting different decision points (landmarks) and turns. The information gained in this mode is more egocentric and cognitively constructed than gaze view knowledge. By employing the gaze view and the route tour modes, users can quickly gain survey knowledge about local areas while simultaneously connecting that knowledge to create a cognitive map of the larger region.

Wayfinding in virtual environments is somewhat problematic when compared to the real world. For example, learning to navigate in computer-based virtual environments has some characteristics similar to the acquisition of macro-spatial knowledge in real space including that the desired objective may not be readily visible. There are important differences in the fidelity of the real and virtual environments that can impact wayfinding (Waller, Hunt, & Knapp, 1998). They point to the work of Mosser (1988) which indicates that acquiring survey knowledge of complex spaces can take up to a year in the real world. The authors indicate that part of the problem in acquiring survey knowledge in VE's is the limited field of view in VE displays that interfere with acquisition of spatial knowledge

Waller, et al. (1998) point out that the differences in knowledge acquisition between the real and virtual worlds must be understood before the benefits of using simulation or virtual environments can be identified. This could be particularly relevant to augmented reality where there is a mixture between real and virtual and the ratio can be changed, possibly in situ. They discuss domain characteristics that facilitate transfer of knowledge. In this regard, they present three information domains in VE training; the real world environment, the training (or virtual) environment, and the trainees mental representation. They claim that these three can never be perfectly aligned, because of systemic differences between them. The mappings between these three environments are related to fidelity which is characterized as the extent to which the real and virtual worlds are indistinguishable. This mapping is also similar to the cognitive approach used to describe human computer interactions (Eberts & Eberts, 1989). The mapping is therefore applicable to designing augmented reality systems.

Witmer, Sadowski, and Finkelstein (2000) conducted navigation experiments using an enhanced virtual representation of an office building that had been used in previous virtual and real navigational experiments. The virtual reality rendition of the real building was enhanced with wider corridors to minimize collisions, additional rooms, and themed objects and sounds added to the environment. Knowledge of building configuration was determined by having participants take the shortest route between two points. Aerial maps were also provided to participants. Sixty four college students participated in the experiment along with eight experts who had multiple exposures to the building. The result of experiments showed that participants who received the enhanced cues performed better in training than those who used the virtual environment that was not enhanced, but not on tests of configuration knowledge. Only participants with aerial views during training performed better during one-week later retention tests. The authors concluded that aids must be employed with care. When used as a crutch, they are not effective on obtaining configuration knowledge. Similarly, aids lose their effectiveness when they increase workload beyond what a participant can handle. The navigation aids worked best when participants could organize their use. These insights are useful to consider in an AR setting and could be useful in creating design guidelines for navigational aids.

Measurement of Human Performance in Navigation

Navigation is an area having received considerable study with respect to human performance in virtual environments. The literature points to several types of navigational tasks as well as system strategies oriented to improving human performance in navigation tasks using virtual or real environments.

Navigation Performance in Mazes

Waller, et al. (1998) create a maze to further explore acquisition of spatial knowledge in VE's. There are several conditions explored including blind (no exposure to maze), real (one minute of free exploration), map review, VE exposure through a desk top maze (2 minutes), VE short exposure (2 minutes), VE (5 minutes). Subjects were not allowed to fly above the maze nor go through walls. Directional arrows were used to direct the user. The users had to touch certain objects in the maze and avoid touching the walls. During some repeated exposures the anticipated path was blocked forcing the user to rely on mental representations of the maze. A distracting task was also introduced at a latter exposure. After exposure, participants were given a number of questions including some regarding selecting among several representations of part of the maze. Having the user select among maze representations was used to determine interface fidelity or the difference between the user's mental model and the actual environment. Actual and mental map representations are used to illustrate the differences. Another significant result showed that relatively low fidelity VE's allow people to create useful representations of large scale spaces, however, short exposures to the VE are normally no better than using a map. There is also evidence that immersive VE's did not facilitate acquisition of survey knowledge. The authors conclude with a recommendation to study map usage within a virtual environment. These conclusions provide useful insights for studies in augmented reality systems.

In another study, Gillner and Mallot (1998) tested navigation in a virtual maze, by seeing how participants found short cuts, estimated distances, and drew sketch maps after their traversal. The authors discuss mental representation of spaces via cognitive maps. Three types of cognitive maps evolve: cognitive maps at a spatial reasoning stage, cognitive maps at a cue integration stage, and cognitive maps as goal independent memory of space.

In analyzing spatial representations, Gillner and Mallot (1998) mentioned that recognizing views requires long term memory and a given view or set of views. They discuss view-based (orientation specific) and place-based (independent of orientation) approaches to acquiring spatial data, both of which can be acquired without global knowledge. These views respectively correspond to fixed map and an egocentric view. A maze was created and was better represented from a view based as opposed to place-based perspective. Human performance assessments of navigation performance were made by analyzing paths and trajectories. The place based approach caused almost equal difficulty in return path navigation (so-called back tracking) as the original traversal for large spaces when using traversal distance measurements as criterion. Sketch maps were interesting to analyze with respect to added or omitted information. The authors devised a parametric system for measuring performance in the maze which was rather elaborate with respect to the number of orientations and trajectories available to the subject. This is not the case with the limited maze for this dissertation. However, some path analysis and map reconstruction might be beneficial. The view based (or fixed) approach is recommended by Gillner and Mallot (1998) due to the lack of clear results in the place based approach.

Gender is a consideration in spatial tasks. Watson and Kimura (1991) studied sex differences on spatial oriented tasks related to throwing and intercepting objects. Their findings showed large differences between males and females on motor tasks and weak trends on paper-and-pencil spatial tasks. The authors note that males have an advantage over females in tests of spatial abilities that involve mental manipulations of objects (referring to the work of Kimura & Hampson, 1990; Maccoby, 1996; Maccoby & Jacklin, 1974; and Wittig & Peterson 1979). Also, Lawton and Morrin (1999) found men consistently outperformed women in pointing accuracy when traversing a simulated maze. The authors point out that pointing accuracy declines as the number of turns increases, but that men consistently outperformed women in this task. The task consisted of the participant pointing towards where they believe the center of the simulated maze is located at certain points in the maze. Also, feedback was designed for the participant to focus their attention on directional information while traversing the maze. Men performed better than women in these tasks. Gender should be considered in wayfinding experiments to avoid a confounding variable.

Subjective Measures Including Presence

People and machines interact in a variety of ways. It is important to understand how machines and people interact in general, but in particular when machines are worn by people or in otherwise close proximity. The reason that it is important is so the machine is useful and does not impede the person's mobility, perceptions, or secondary task performance. Augmented reality is a technology that can influence these aforementioned interactions positively and negatively. Negative outcomes are indicators of where barriers need to be overcome. For example, AR equipment might not be used by the public if it is heavy and is not fashionable. While one can measure many aspects of the interactions through outcomes, some measures lend themselves to questionnaires or interviews. Several authors have addressed these types of issues in various writings.

Hancock (1997) discusses how humans and technology have shaped each other. Some technologies have moved from discretionary to mandatory (whether real or otherwise). The author points out differences in evolutionary characteristics between humans and machines such as life span which in humans is decades and in machines are years or months. However, there is an inextricable linkage between people and useful machines that often happens haphazardly, but causes mutual change to occur in a process of co-evolution.

Clark (2003) takes a similar point of view as Hancock (1997) in explaining co-evolution. The author points to successful co-evolution as extending the human cognitive processing, working memory, or perceptual systems. For example, the wrist watch can be considered an extension of the cognitive system, because when asked, "Do you know what time it is?" most will answer, "Yes". The watch becomes an adjunct to working memory by storing the time.

Hancock (2004) discusses the future of simulation, in part, by use of the Turing test. The Turing test is one where the participant cannot distinguish between inputs generated by a human or computer. Users of simulations in the future will not be able to distinguish between computer generated responses and those of the real system. Such a capability will allow scientists to evaluate concepts before committing to hardware as well as exploring aspects of human interactions with machines and other humans outside of the actual interactions. These types of interactions could be useful to the co-evolution of augmented reality systems as well as potential uses of developed systems in areas such as human evaluation of prototype products or to facilitating human to human interactions. These discussions are relevant and germane to augmented reality's current and future capabilities, but a more principled method is needed to understand and steer the dynamics of co-evolution. Presence may be the principled method.

Presence is a relatively new field of research that seems to have emerged through the musings of Sheridan (1992). Since Sheridan (1992), there have been numerous definitions offered along with measurement methods. These definitions and associated measurement methods have often times changed with changing requirements, research studies, and technological advances. Perhaps the most contemporary definition(s) of presence that is relevant to augmented reality can be found in International Society for Presence Research (2000) which in part states, "Presence (a shortened version of the term "telepresence") is a psychological state or subjective perception in which even though part or all of an individual's current experience is generated by and/or filtered through human-made technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience…". Key to this definition is recognition of the integral roles of technology and the human in presence. The synergistic interplay between the technology and the human is of interest in order to create more useful, compelling, and cost effective augmented reality systems. Based on this definition, presence is a relevant term to augmented reality and may be a useful tool to assess user acceptance of augmented reality technology. In particular, presence could be a catalyst to help accelerate augmented reality's development and integration into segments of society. In this context, presence is a facilitator of co-evolution of augmented reality.

Heeter (1992) studied presence by describing its three components: self, environment, and social presence. Representation of self in the environment includes visualization of self and sensory responses to the individual in the environment. The environmental component involves the ability of self to affect and be affected by the environment. Social aspects include the ability to interact with other entities in the environment. Heeter's discussion is oriented to virtual environments, but the concepts are applicable to augmented reality environments or any environment on Milgram's Continuum.

Presently, there is no single universally accepted method for measuring presence. However, there are a number of factors that have been suggested that influence presence. These factors are shown in Table 1 (Sadowski, 1999). Although no single approach for measuring presence has emerged, two different approaches have been proposed for measuring presence: subjective measures and objective corroborative measures. Objective and subjective measures are further explained below. Measurement methods are an area of active research and discussion among virtual environment researchers.

Table 1: Factors Affecting Presence

Factor	Guideline	Issue
Ease of Interaction	Provide seamless interaction to simulate the real world.	Poorly designed interaction takes focus away from the experience.
User-initiated control	Provide immediacy of system response, correspondence of user-initiated actions, and a natural mode of control.	Delays, discordance of users' versus effectors actions, and unnatural control devices hinder engagement in the VE.
Pictorial Realism	Provide continuity and consistency in presented stimuli.	Poorly designed visual interaction hinders engagement in the VE.
Length of Exposure	Provide sufficient exposure time to provide VE task proficiency, familiarity with the VE, and sensory adaptation.	Avoid unnecessarily prolonged exposures that could exacerbate cyber sickness.
Social Factors	Provide opportunities to interact with and communicate with others verbally or by gestures.	If one's presence is not acknowledged by others it may hinder the perception that they "exist" in the VE.
Internal Factors	Identify the types of individuals who will use a VE system and their preferred representational system.	Individual differences can render VE systems differentially effective.
System Factors	Provide multi-modal interaction input/output to facilitate presence.	Poorly designed systems can degrade the users' experience.
User Characteristics	User's perceptual, cognitive and motor abilities and prior experience should be taken into consideration.	It also depends on age, sex and mental health conditions.

Many believe that presence is primarily a subjective sensation and there are several methods researchers suggest for its measurement. One method suggested is known as Post-Test Rating Scales (Sadowski, 1999; Ijsselsteijn, 2000). In this method, the user answers a number of questions rating the naturalness of interaction in VE on a rating scale of 1-7. Witmer and Singer (1998) developed the PQ (Presence Questionnaire) to measure presence with respect to involvement/control, naturalness, and interface quality. The PQ seems to be a very popular rating method. Another variation of questionnaire is the so-called Subjective Reports suggested by Sadowski (1999). In this method, the participant has to answer the questionnaire in his own words. So the bias from questionnaire interpretation is reduced. But the final interpretation of results is difficult due to large response variability.

One important limitation of post-test subjective ratings is that they do not provide any measure of temporal variations in presence. Such variations are likely to occur through temporal changes in the stimulus or the participant's workload. Therefore, some researchers have applied a method of continuous presence assessment (Ijsselsteijn, 2000) to measure presence. In this method, the subject has to judge the amount of presence using hand held slider. This slider is connected to computer where we can see the temporal variation of presence measure with respect to input sensory information. The limitation of this method is that the subject's attention is diverted between display and online rating slider. So the alternative option is analyzing verbal reporting for temporal measurement.

Subjective measures of presence have the advantages of being generally easy to administer and interpret. However, the major drawback of subjective measures is the problem of inter-rater reliability. The participants must understand the concept of presence and interpret the questions uniformly. Subjective ratings of presence can be biased by previous experience, subject's interest and prior knowledge.

Objective measures of presence relate to user responses that are, in general, produced automatically and without much conscious deliberation. Ijsselsteijn (2000) is a compendium that addresses many of the salient issues. Some authors in Ijsselsteijn (2000) have suggested task performance as an objective measure of presence. Task performance is a valid form of assessment in many human factors and psychological tests as implied by Wickens and Hollands (2000).

A number of physiological indicators, such as heart rate and skin conductance response (SCR), have been suggested as objective corroborative measures of presence. They serve to isolate influences on presence and can be objectively measured. The main limitation is internal and external noises that affect the results. However, recent results cited by Meehan, et al. (2002) show benefits of using heart rate variability as a reliable measure of presence in a stressful setting induced by the so-called vertical cliff.

Finally, over 80 studies in various aspects of presence were reviewed by Younblut and Perrin (2002). They analyzed the predominant methods of assessing presence, namely the Slater, Usoh, Steed (SUS) questionnaires and the Witmer-Singer Presence Questionnaire (PQ). Interestingly, Youngblut and Perin (2002) find no reliability data in the literature for the SUS questionnaire. Both the SUS and PQ were unable to reliably distinguish between a real and virtual environment and further state that it is not clear that either are valid measures or measuring the same construct. The authors seem to allude to the fact that presence and performance might be related in some tasks, possibly those requiring high levels of performance. More research is needed in virtual environments.

The utility of measuring presence in augmented reality is of interest for its possible relevance to visual fusion as described by McGee (1999), conformal displays as described by Wickens and Hollands (2000), and co-evolution as described in the literature by Hancock (1997) and Clark (2003). It is clear that adjustments to existing measures might be needed to account for the affects noted by Heeter (1992).

The literature provides a solid basis of research on technical and human performance issues in navigation, the use of maps, virtual environments, and augmented reality. The literature also demonstrates a gap in research assessing human performance improvements when navigating with mobile augmented reality equipment.

CHAPTER THREE: EXPERIMENTAL METHOD

The understanding of the current state-of-the-art, derived from the present review of literature leads to the hypothesis that mobile augmented reality is beneficial to improving performance in navigational tasks. However, mobile augmented reality is a nascent technology with variability in equipment performance. This equipment variability coupled with human variability introduces additional uncertainty that is not found in many studies involving human performance or engineering-related research. As augmented reality technology matures, it is expected that variation in its performance will be controllable. The general hypothesis about AR spawns a number of specific hypotheses as enumerated, below.

Research Hypotheses

Hypothesis 1: Wayfinding using a real time, hands-free AR map display will improve human performance compared to the control conditions (there are two variants of control conditions).

Prediction 1: Performance will be improved by at least 50 percent as measured by the time needed to complete the primary way finding task as compared to both control conditions.

Prediction 2: Performance will be improved by at least 50 percent in the accuracy of the primary wayfinding task performance as compared to both control conditions.

Prediction 3: Performance will be improved by at least 50 percent in the accuracy of a secondary spatially oriented embedded task as compared to both control conditions.

Hypothesis 2: Within the AR condition, the use of egocentric moving map as compared to exocentric fixed North-up map will reduce the time and improve recall accuracy on a secondary task, in terms of Euclidian distance (in feet) and egocentric orientation, by providing real time egocentric view of the space.

Hypothesis 3: Within the AR conditions, the use of exocentric fixed North-up map as compared to egocentric moving map will reduce the time and will improve recall accuracy on the secondary task, in terms of cardinal directions, by providing real time North-up view of the space.

Hypothesis 4: Within the AR conditions, the use of an "on demand" map display as compared to a continuously-on map display will reduce the time and improve accuracy on the primary wayfinding task and the secondary embedded task by providing the participant with control over the times when the map is needed as a reference.

Hypothesis 5: Participants scoring higher in objective performance will positively correlate with higher subjective ratings of mobile AR utility. Co-evolution and presence require a threshold level of immersive fidelity and task performance requirements (e.g., heightened stress level) to be triggered.

These five hypotheses were formulated to better characterize the gains that can be expected and the issues that may arise in using mobile AR for certain wayfinding tasks. While the benefits of maps versus no maps might appear obvious, there are also detriments that could occur as a result of human factors issues such as display positioning, obstruction of the background, or the bulkiness of the mobile augmented reality system. These human factors issues could affect the sense of presence and utility of mobile AR. The participant's level of workload (passive to active) and augmented reality characteristics could affect the reliance on the AR equipment, its transparency, and therefore ones sense of presence. Activation of the sense of presence and passing the Turing test (Hancock, 2004) as well as utility issues could further impact co-evolution of the technology. Hypothesis 5 reflects these dynamics by establishing a baseline of the interaction between task involvement (passive to active continuum) and augmented reality fidelity (minimal to fully fused and integrated displays) leading to better insights into measuring the sense of presence and the Turing test for mobile AR systems. Objective measures related to task performance using mobile AR are also explored principally in hypotheses 1, 2, 3, and 4. These hypotheses are formulated to obtain baseline of the interplay between map orientation and persistence in specific workload and attentional situations.

Participants

136 students, faculty, and staff from area universities and schools volunteered to participate in the experiment. Eight participants were used for pilot studies, data from seven participants had to be discarded because of equipment failures, and one participant's data was discarded because of errors in the pre-test briefing by the experimenter. The resulting participant pool was 60 males and 60 females. The average age of the participants was 26.5 years (see Table 2 for details by treatment).

Volunteers were rewarded with extra credit, where practical, for participating in the experiment and were treated in accordance with the "Ethical Principles of Psychologists and Code of Conduct" (American Psychological Association, 2002). In addition, participants were informed of incentives based on performance. The best performing participant in each condition received a $20 gift certificate for use at a local restaurant.

Table 2: Age Distribution of Participants in Different Treatments

Treatment	Mean Age	N	Std. Deviation
Control with Map (C_M)	25.7	20	9.7
Control with Compass -No Map (C_NM)	29.8	20	11.5
Egocentric Map & Continuous Display (E_C)	25.6	20	10.3
Egocentric Map & On Demand Display (E_D)	24.1	20	7.8
Exocentric Map & Continuous Display (X_C)	27.0	20	10.9
Exocentric Map & On Demand Display (X_D)	26.8	20	9.0
Total Among Treatments	26.5	120	9.9

Apparatus

A mobile augmented reality system was used. This system, called the Battlefield Augmented Reality System (BARS), was integrated and configured by the Naval Research Laboratory for experimentation and provided for this research by the United States Army Research Institute for Behavioral Sciences. Images of the system can be seen in Figure 5. The BARS uses a Thermite computer from Quantum 3D as the central processing unit. The Thermite is a battery operated wearable computer. It uses a 1 G Hz Transmeta central processor, with supporting hardware from NVDIA GeForceFX Go GPU for graphics processing. The configuration of the Thermite for this research used the Windows operating system and provided several options for display outputs (VGA is used in BARS). Visuals are provided to the user with a MicroOptical SV-6 PC Viewer. The SV-6 has 640 pixels by 480 lines resolution, 18 bit color depth (262,144 colors), an approximate 16 degree x 20 degree field of view, 60 Hertz refresh, and adjustable focus from 2 to 15 feet. The setting of focus helps facilitate placing images at pre-selected distances to facilitate depth perception. An image of the SV-6 PC Viewer can be seen in Figure 6. There are two means for tracking human motion in the BARS. For these indoor experiments an InterSense IS-900 acoustic tracker was used. The tracker is fixed to the participant's shoulders providing a body fixed position and orientation. Outdoor tracking is accomplished through a differential GPS and inertial orientation sensor.

Figure 5: Front and side view of the battlefield augmented reality system (BARS)

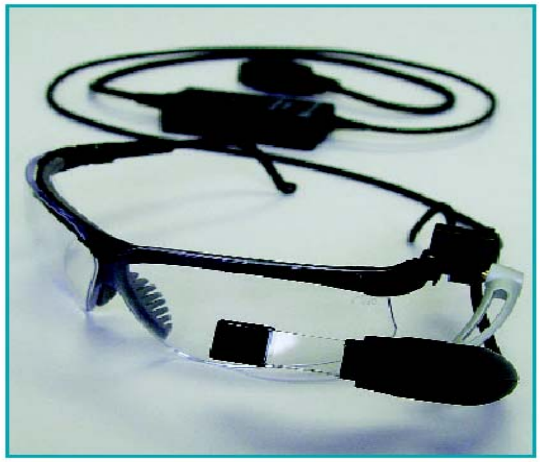


Figure 6: MicroOptical SV-6 display and glasses used with BARS

Spatially correlated virtual and physical mazes were constructed. The virtual plan view representation of the maze has been programmed on the BARS and is viewable through the SV-6 PC Viewer. The virtual maze occupies approximately 55 percent of the displayable area in the SV-6 PC Viewer. Reducing the displayed size of the maze represented a compromise to keep the complete maze viewable in nearly all egocentric viewing conditions. A small portion of the maze is not viewable if the participant is in one corner of the maze. The walls of the physical maze were opaque and the maze height was set at 78 inches. Hallways were 35 inches wide. The interior of the maze had no overt rooms or landmarks. Additionally, participants wore a visor limiting their tendency to look up to obtain spatial landmarks from the ceiling. Therefore, the map (displayable or paper) of the maze provided the only source of spatial information for the participant.

A scaled plan view representing the virtual maze and similar to that seen by the participants is provided below in Figure 7. The map seen by participants in the display was white on a black background. The blue, irregular shaped mark in Figure 7 indicates where a target object is placed. The placement of the target object, though, was not displayed to participants. Figure 7 also indicates where spatially oriented questions were placed for the participant to answer during maze traversal. As with the target object, these locations were not displayed to the participants. An oblique photograph of the physical maze is shown in Figure 8. The photograph and schematic have different orientations (north up for Figure 7 and north down for Figure 8) due to limitations in accessing the maze for photographs.

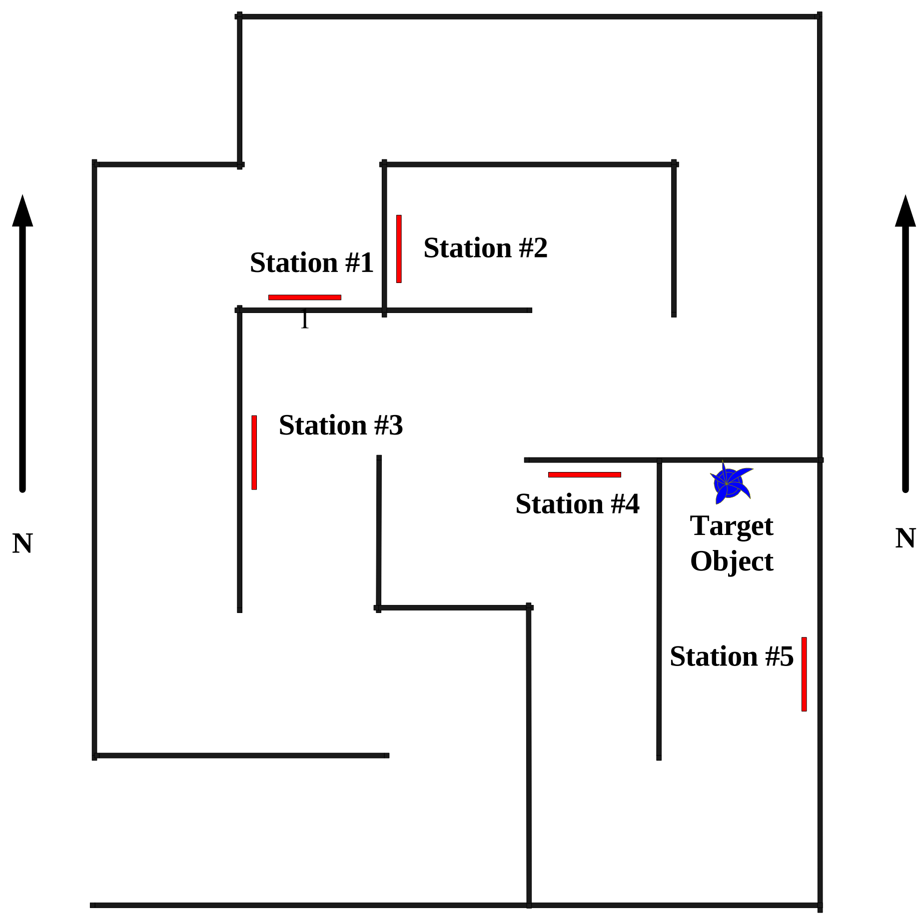


Figure 7: Schematic view of the maze

Figure 8: Oblique photograph of the maze

Experimental Task

Each participant completed two tasks: a primary and a secondary embedded task. The primary task was to completely traverse the maze (including dead ends) before retrieving the target object. The object's location is shown in Figure 7, but was not shown to the participant on their map. After the target object was retrieved the participant was to immediately find the shortest way back to the entrance of the maze. The secondary embedded task was to answer questions placed at five locations in the physical maze. The questions were placed in canvas bags and were designed to be of a similar nature (i.e. visual-spatial) to the primary task and involve spatial orientation questions. Appendix A contains the secondary task questions. The purpose of the orientation questions was to tap on the acquisition of landmark (the "stops" where the participants were asked to perform the secondary task may be considered as landmarks) knowledge and recall accuracy, thereby increasing mental workload. The secondary task also involved a subjective workload rating using a questionnaire. In order to assess if route and survey knowledge is acquired during maze traversal, after the completion of the primary and the secondary task the participants were asked to first, identify the maze they had traversed by presenting them with six options in a multiple-choice format; second, to identify the location within the maze of the target object. Appendix B contains the post hoc spatial recall questions.

As a part of investigation of co-evolution an assessment was made of the sense of presence after maze traversal. Questionnaires were used that were adapted from Singer and Witmer (1996) and Baird (1999) to be relevant to the areas of interest in this dissertation. The adapted Immersive Tendencies Questionnaire can be found in Appendix C. The adapted Presence Questionnaire can be found in Appendix D. The questionnaires were meant to consider presence and use related factors as they might affect co-evolution as augmented reality capabilities increase. Some question responses included a category ‘not applicable’ because the feature was not present in BARS or an experimental treatment.

Experimental Procedure

In a room separate from the maze, an informed consent form was completed and demographic information (see Appendix E) gathered from each participant. Participants were then tested for spatial abilities. Parts 5 and 6 of the Guilford-Zimmerman Aptitude Survey (as cited in Consulting Psychologists Press, 1976) were administered as pre-tests. These tests respectively measure a participant’s spatial orientation and spatial visualization abilities.

Participants were also given the Immersive Tendencies Questionnaire that was adapted for AR (Appendix C).

Prior to traversing the maze, the participant was briefed on the nature of the task. Participants were informed they should completely traverse the entire maze including dead ends, next retrieve the target object, and finally exit through the entry point. They were instructed to complete this task as quickly as possible without running. Participants were informed if they would be provided with a map on the head worn display, how the map would be enabled (i.e., continuous or on demand), and that they would be timed while traversing the maze. They were told that they would be penalized should they retrieve the object before completely traversing the maze and conducting the secondary task. They were also informed that the person performing best for a given condition would be rewarded with a gift certificate.

Participants were told that as a secondary task they would be asked workload and spatially oriented questions located at various stations during their traversal of the maze (Appendix A). Secondary task stations consisted of a numbered canvas bag containing questions that participants were to answer. The secondary task stations were to be visited in sequence before retrieving the target object. That is, questions in station 1 must be answered before answering questions in station 2, and so forth. Doing otherwise would result in a penalty. Additional spatially oriented questions would be asked at the completion of the experiment. Participants were informed that they should try to keep track of the route they had taken as they would be asked to complete a spatially oriented questionnaire after traversing the maze (Appendix B).

Participants were blindfolded and lead from the testing room to the maze room. They were allowed a short time to become comfortable with wearing the BARS equipment. The participant decided over which eye to position the SV-6 PC Viewer. The viewer was then focused and positioned at the most convenient position for them. The participants were randomly assigned to one of the experimental treatment groups described, in Table 3. They were instructed how to utilize the mobile AR equipment. After completion of all tasks the participants were debriefed.

The following actions represent a possible sequence of events for the participant, once pre-testing and accommodation to the equipment is completed;

1. Participant is positioned at the opening of the maze.
2. A quick review is made of the procedure the participant is to follow and that there are timing and accuracy considerations in their tasks.
3. The BARS is enabled by the experimenter, the participant is told to begin, and a timer is started (the timer is not visible to the participant).
4. The participant begins traversing the maze and viewing their map and area covered (if available as part of the experimental conditions).
5. A bag containing spatially oriented questions is found on the maze wall.
6. The participant answers the questions, returns the paper to the hanging bag and continues exploring the maze or finding the next bag.
7. Steps 4, 5, and 6 are repeated.
8. The participant finds and retrieves the target object (a wireless tracker wand) and exits the maze through the shortest route.
9. The experimenter stops the clock, ends the BARS exercise and retrieves the answers to the questions.

The participant removes the BARS and begins the post testing.

Various alternatives are possible to the above scenario. For example, the participant could find the object by chance first and then immediately leave the maze, leaving area that is not explored and not completing some or the entire secondary task. Or, the participant could spend time at a questioning station planning their next move. Or, the participant could find spatial questions that (s)he is to answer after finding the objective (which begins the exit activity), but are answered on the egress. In this case, the answers would be scored as incorrect. Other variations are possible.

A sample script of the procedure experimenters used to brief participants can be found in Appendix F. The script was used to minimize variation among the experimental staff.

The maze was derived from other mazes used in studying navigation in virtual environments (Waller, 1998; Jansen-Osmann, 2002; Barlow, 2001). The maze was positioned under an IS-900 tracking system. During traversal of the maze, real time position and orientation of the participant can be displayed.

Experimental Design

A between-subjects design was used for the study. There were two independent variables (IVs) in this experiment. There were also two different control groups used in the experiments. Table 3 describes the between subject design treatments.

Table 3: Matrix of Experimental Treatments

Control Group	On Demand Display	Continuous Display
w/o map (C_NM)	Fixed Map (X_D)	Fixed Map (X_C)
with Map (C_M)	Forward Up (E_D)	Forward Up (E_C)

Experiment Administration

Participants were randomly assigned to an experimental treatment. The population sample used was counterbalanced for gender across all treatments. It was also attempted to keep treatment assignments balanced by gender and by total participant distribution between treatments as testing progressed. Seven experimental treatments had to be rerun with different participants due to equipment failures. One participant's data was not used because of an error in instructions by the experimenter. 120 participants, equally divided by gender and distributed across the six conditions are used for analysis.

Control Groups

Two control groups were used for comparisons; a baseline control group and a baseline with paper map. The baseline group (C_NM) wore the BARS equipment, but had no computer augmentation. The map display was not used. Participants were provided with a compass. A variant of the baseline control group, Control Group with Map (C_M), were allowed to review a paper map prior to traversing the maze. Timing started when the map was given to this control group participant. The map was taken from the participant when they entered the maze and began traversal. No compass was provided to the participant in Control Group with Map.

Independent Variable 1: Map Orientation

The two levels of IV_1 (map orientation) were referred to as "Egocentric map display" (designated E in the first letter of the treatment abbreviation) and "Exocentric map display" (designated X in the first letter of the treatment abbreviation). Each display consisted of the same basic plan view of the maze except for the map orientation. The egocentric map was oriented to the participant's forward direction being up in the map viewed through the BARS display. The exocentric display was oriented in a fixed north up position. The north directional arrow appeared on all displays and maps.

Independent Variable 2: Map Display Availability

There were two levels of IV_2 (Map display availability). The two levels were referred to as "On Demand" and "Continuous". In the "On Demand" condition (designated D in the second letter of the treatment abbreviation), participants could turn the map display "on" by depressing a space bar on a keyboard attached to their forearm. The map would stay "on" for approximately 15 seconds. The participant could repeat the procedure as frequently as they desired. In the "Continuous" (designated C in the second letter of the treatment abbreviation) condition, the map display was continuously on.

Covariates

There were several potential covariates in this experiment that were obtained from the demographics information or questionnaire scores obtained prior to maze traversal. The covariates included age, gender, and Guilford-Zimmerman scores.

Evaluation of Performance

There were four broad dependent variables (DVs) in this experiment. All DVs were measurable and objective. The DVs were 1) total and weighted time (seconds) to complete the primary and secondary tasks; 2) maze area covered; 3) knowledge about various aspects of the participant's spatial orientation during maze traversal for the secondary task as reflected by answering orientation questions for the secondary task, and 4) spatial recall after maze traversal as reflected by recalling which maze was traversed and where the target object was in the selected maze.

The weighted time DV was computed as the total time to complete the primary and secondary tasks divided by the fractional area traversed before retrieving the target object. There are two paths to exit, one longer than the other, thereby requiring more time if the wrong route is selected. Overlapping areas already traversed are also reflected in total time. Maze area covered by the participant, expressed as a percentage of the total maze area traversed prior to retrieving the target object, is also considered as a dependent variable (the second DV) because of the relatively small maze area and the resulting possibility of small variations in traversal time with treatment regardless of the route traveled by the participant and because participants were instructed to completely cover the maze before retrieving the target object. The percentage of the area covered was made by observing the track of the participant through the maze and by capturing and analyzing the participant's location in the maze with the IS-900 tracker. The participant's position and orientation was captured at a 1 Hz. The stored information of a participant's track was overlaid onto the maze for analysis. A sample is shown in Figure 9.

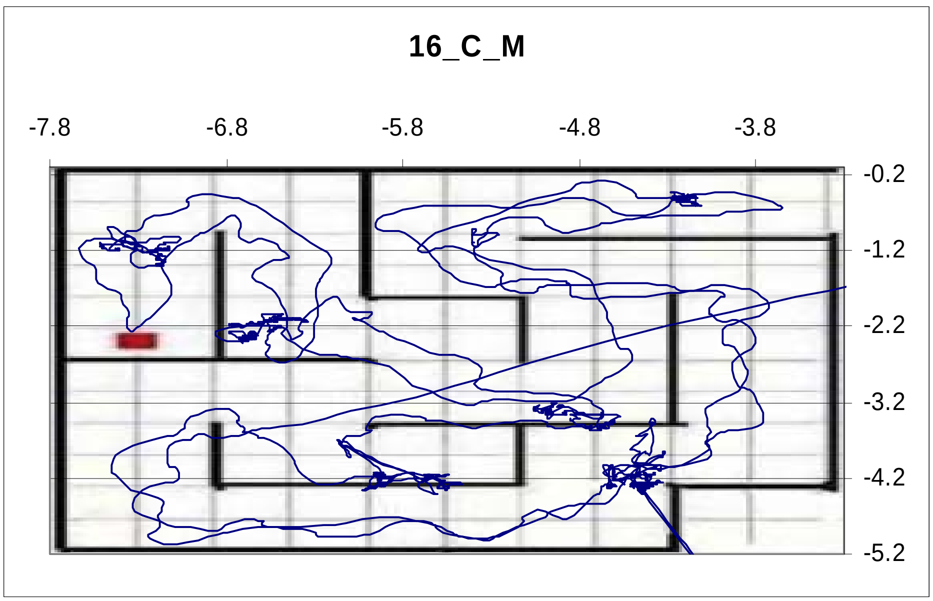


Figure 9: Trace of maze traversal by participant 16_C_M

The third dependent measures of performance were spatial orientation and distance estimation while traversing the maze. Participants were asked spatial and distance oriented questions at five different locations in the maze. Each of the five locations contained a numbered canvas bag with questions. The questions were structured to measure participant working memory ability to recall accuracy of Euclidian and Cardinal directions as well as the Euclidian distance to the maze entrance of the participant. All responses were keyed to be answered before retrieving the target object. Performance was decremented if the succession of bags visited was out of sequence, or if questions at each station were incorrectly or not answered. Answers were equally weighed across bag location. The five locations correspond to Figure 7 and the questions can be found at Appendix A.

After completion of both primary and secondary tasks the participants were given a spatial recall questionnaire where they asked to: 1) identify the maze they just traversed using multiple-choice format, and 2) draw a checkmark to identify the position of the target object before retrieval within the maze they identified as the one they just traversed. This fourth DV was based on the participant's spatial recall of the maze just traversed and represents the fourth dependent variable. No time limit was set on this latter test.

Subjective Assessment of Performance

A subjective assessment of performance was conducted using selective responses to questions from the Presence Questionnaire (Singer & Witmer, 1996) adapted for this experiment. This adapted version of the Presence Questionnaire can be found in Appendix D. The purpose of this evaluation was to assess the acceptability of BARS as a device for facilitating successful task performance from human factors and functional perspectives. This evaluation was also useful to determine if co-evolution and the Turing test for mobile AR (Hancock, 2004) was occurring. Due consideration was given to the participant's background as determined by participants completing an Immersive Tendency Questionnaire (Singer and Witmer, 1996) also adapted for this experiment. Changes to the questionnaires were necessary to remove questions strictly oriented to virtual environments and to add questions germane to navigation tasks in AR.

CHAPTER FOUR: RESULTS

A series of three experiments were conducted. The first was intended to ensure the testing methods and data collection techniques were sound as well as assess where any procedural changes were needed. These data have been analyzed as appropriate. Changes in the experimental design were made subsequent to this first procedure. The next two experiments had the same basic format. The second experiment consisted of 84 participants while the third experiment tested 36 participants. A total of 120 participants were tested divided equally between the treatment conditions and balanced across gender. SPSS, version 12.0, Grad Pack, was used for the subsequent statistical analysis.

Experiment 1 Results

Experiment 1 was conducted to qualitatively evaluate and refine procedures and address any issues with participant or equipment performance evaluation. This consisted of using conditions which represented controls for later procedures (no display/no map and no display/with paper map). A number of problems were identified that were subsequently resolved or mitigated. One problem occurred when entering the laboratory containing the maze in which participants saw the maze and its location. Having such views could provide participants with valuable spatial cues usable when they are in the maze. This problem was resolved by blindfolding participants in a separate room and before entering the maze laboratory, turning them around while blindfolded. A tent like structure was also installed around the maze to shield its configuration. The tent reduced the possibility of participants gathering any additional spatial cues external to the maze that might be visible. A second problem occurred when the experimenter and staff were talking in the maze area while the participant was traversing the maze. Discussions might have provided valuable spatial cues to the participants in terms of an audition landmark. This potential problem was minimized by posting signage in the maze laboratory to keep the area as quiet as possible.

Participants traversing the maze indicated the ease of the process during responses to questions at the various stations. If the secondary task was too simple, a ceiling effect might have resulted in the maze traversal performance due to its restricted dimensions. This potential problem was minimized by requiring participants to answer station questionnaires in numerical sequence, by positioning stations to minimize visibility between stations, as well as by blindfolding participants and turning them around before entering the maze area to avoid their having pre-traversal directional information. The blindfold was removed as the participant entered the maze.

The BARS system exhibited variability in activation causing uncertainty in accurately timing traversal. This problem was overcome by repeatedly measuring the time to activate BARS from a cold computer start and collecting samples of the time to activate the system for a new participant from an already running BARS. Times were collected, and analyzed for consistency within and between treatments as well as activation technique (restart versus cold start). There was a difference in activation timer between the Egocentric Continuous Display condition and all other conditions. The bias time to activate the BARS for all treatments other than the Egocentric Continuous Display treatment was 45 seconds from a cold start of BARS or 16 seconds for activation when BARS is running. The bias time for the Egocentric Control treatment was 55 seconds from a cold start of BARS or 21 seconds for activation when BARS is running. Bias times were subtracted from the subsequent performance times to reflect an accurate time for traversal of the maze.

BARS exhibited occasional problems due to tracking variations and freezing while operating. The tracking problem manifested itself in several ways including, losing tracking for short periods of time (typically less than two seconds), sudden movement in the angular orientation and position of the participant, noise in the system as manifested in erratic traces, and spinning maps. This problem was due, in part, to the tracker being placed on a pole attached to the BARS and at an elevation that was higher than the maze height resulting in a long moment arm of the tracker relative to its effective center of rotation. This tracker placement was necessary to avoid acoustic tunneling and reverberations that arise due to the maze's parallel and closely spaced walls. The problem was addressed by adjusting the tracker mount to the BARS to raise the effective center of rotation and instructing participants to stop for a few seconds if anomalous behavior of the tracker was objectionable. This resolved the problem to the greatest degree possible. Rotating maps were addressed by aborting the experiment for that participant. BARS freezing seemed to be due to a combination of memory leaks in the operating system and related problems with JAVA threads. This problem was addressed by restarting the BARS and its operating system on a daily basis.

Measuring Performance

Several variables were scored as measures of performance for the subsequent experiments. Scoring for the primary task of traversing the maze was measured in terms of total time, weighted time, and percentage of the maze covered. During traversal of the maze secondary tasks were added to increase workload by requiring the participant to use available resources to estimate Cardinal and Euclidian directions in the maze and Euclidian distance. Available resources could be augmentation (AR, a map, or a compass) or working memory. The secondary tasks were direct measures of the participant's spatial accuracy and recall ability. After traversing the maze, the participant conducted a spatial recall task that was also scored.

Other measures were taken prior to performance and subjected to analysis. These include data available from the Demographics Questionnaire (Appendix E), Parts 5 and 6 of the Guilford-Zimmerman Tests (as cited in Consulting Psychologists Press, 1976), and an adapted version of the Immersive Tendencies Questionnaire (Appendix C). Post hoc data reflecting subjective performance were taken from an adapted version of the Presence Questionnaire (Appendix D).

The following scoring scheme was used for Euclidian and Cardinal directions and Euclidian distances from questions in the five bags located throughout the maze:

Directions within maze (50 possible points)

- Euclidian Directions (worth 3 points)
  - exact: 3 points
  - ± 45° from exact: 2 points
  - more than 45° from exact: 0 points
- Euclidian Distance (worth 3 points)
  - exact or within ±1 foot: 3 points
  - $1 \text{ ft} < x \leq 2 \text{ ft}$ or $-2 \leq x < -1 \text{ ft}$: 2 points
  - otherwise: 0 points

- Cardinal Directions (worth 4 points)
    - exact: 4 points
    - ± 45° from exact: 2 points
    - more than 45° from exact: 0 points

The scoring scheme was developed to accommodate the variation in tracker performance and resulting graphical presentation of the participant's position and orientation in the maze. The scoring scheme also is intended to put a sufficient spread in scores thereby aiding statistical analysis. The scores for Euclidian direction and distance, and Cardinal direction are independent. Four points were used as a maximum for each Cardinal direction score to reflect greater difficulty in achieving an exact solution than estimation of Euclidian direction and the resulting Euclidian score. The increased difficulty arose from the alignment of stations along cardinal points, the previously mentioned variability in tracking accuracy, and the resulting requirement for participants to interpolate their directional orientation along the fixed Cardinal system. A total possible score of 50, summed from the bags at all of the stations, is equal to the total available score for the post hoc spatial recall question, thereby supporting comparative analysis of accuracy during and after the treatment. An aggregate accuracy score of 100 also supports any analysis of post hoc speed accuracy comparisons on an individual basis.

Station sequence effects were also considered in scoring. Answering bags out of order results in no points for those bags that are out of order. As an example, participants are to answer bags in the sequence 1, 2, 3, 4, and 5. A participant answering the station bags in the sequence 3,2,1,4, and 5, for example, would result in two bags being out of sequence. A score of zero for bags 1 and 3 would be recorded, regardless of the individual scores for Euclidian and Cardinal direction and Euclidian distance for those bags.

The scoring strategy for post hoc spatial recall was weighted equally (i.e., 50 points total) to the aggregate of Euclidian and Cardinal scores or could be scored categorically if other analytical methods, such as the Chi squared statistic, were to be applied. The numerical scoring used was as follows:

- Correct answer (correct maze, correct object location) – 50 points
  - (correct maze, wrong object location) - 10 points
  - (wrong maze, correct object location) – 10 points
  - (wrong maze, wrong object location) – 0 points

The strategy for allocating points was to spread the scores and to account for the fact that the location of the object was the same in all multiple choice responses (see Appendix B). Identifying the correct location of the object for the wrong reason was classified as an error.

Maze traversal time was scored as the actual time taken to traverse the maze, answer the questions in the bags, find the target object, and return to the entrance of the maze. The appropriate time taken for the BARS to become active was already subtracted from these total times which were then analyzed.

Maze coverage was expressed as a percentage of the entire maze covered prior to retrieving the target object. Participants were tracked using the IS-900 tracker. Their position was sampled at a 1 Hz and stored in an Excel spreadsheet. Traversal was followed in real time on a computer terminal external to the maze. Traversal was graphed onto the maze plan view from IS-900 tracker data allowing quick analysis of area covered as well as the order of the bags visited in the maze. Figure 9 in Chapter 3 provides a sample of graphical trace of traversal.

The final measure of performance was called the weighted time. The weighted time was the total time divided by the fraction of the maze traversed prior to retrieving the target object.

In the example shown above in Figure 9, the participant traversed 100 percent of the maze so their weighted time would be their actual time and their percent of the maze covered would be 100 %. The weighted time is meant to penalize a participant who does not completely traverse the maze prior to retrieving the target object.

In summary, measuring performance included dependent variables from the primary and secondary tasks. Other measures were taken using various instruments prior to performance. Table 4 summarizes the measures during and after traversing the maze.

Table 4: Description of Dependent Variables

Dependent Variable	Description
Bag Sequence	The ordinal succession of finding the bags in the maze
Cardinal Direction	A participant's estimation of maze orientation using 8 cardinal points (N, NE, E, SE, S, SW, W, NW)
Euclidian Direction	A participant's estimation of the maze entrance with respect to ones current orientation
Euclidian Distance	A participant's estimation of the distance between their location and the maze entrance
Percent Covered	The area of the maze covered by a participant
Spatial Recall	The participant's recall of the maze configuration and location of the target object
Total Time	The total time taken in traversing the maze
Weighted Time	Total time/Percentage of maze covered

Experiment 2 and 3 Results

A combination of Chi squared for categorical data, and ANOVA and ANACOVA analyses for the continuous variables were conducted on the collected measures of performance responses. In some cases the continuous data was converted to categorical data using a simple success/failure dichotomy. Data was analyzed at a level at which $p \leq .05$ was considered significant. There are also reports at larger values of $p \leq .10$ as marginally significant since many of the trends in this practically-oriented work were of design interest. This larger value is also reported in part because of the nature of the equipment used and the characteristics of the present research hypotheses. BARS is a prototype equipment with resulting variations that occur in its performance (e.g., start up time or tracking). These variations in system features thus results in the rationale for reporting a larger significance value. Moore (1994) further points out that using other values of significance depend on plausibility of H0 and the consequences of rejecting H0. This research is the first formal application of H0 to AR so there are no historical measures at present in this domain. Also, H0 for this research is the first comprehensive study of human performance with a prototype AR and intended to guide further developments and studies.

Analysis using Chi squared tests were performed for several data sets. The Chi squared test is useful for determining whether one variable has been influenced by other variables. Chi squared also supports multiple comparisons which is the situation in the present procedures. Chi squared analysis is an indicator, though, and more detailed analyses are often required using other statistical techniques supporting pairwise comparisons of data. ANOVA's and ANACOVA's were used to for further analysis of various treatments and DV's.

Total Time

The total time in the maze was analyzed using ANOVA with the following values, as shown in Table 5 and graphically in Figure 10.

Participants in the Control Map condition (C_M) spent considerably less time in the maze than either of the egocentric treatments, however it was not statistically significant when the treatments were considered on an aggregate basis, $F(5, 119) = 1.677$, $p = .146$. Pairwise comparison using ANOVA between C_M and the Egocentric Continuous (E_C) treatments had a mean difference of nearly 95 seconds which proved significant at less than the .05 level at $p=.039$ and in the Egocentric On Demand treatment (E_D) the difference was over 104 seconds at $p=.024$. No statistically significant differences were found between the C_M treatment and exocentric treatments or between the Control No Map (C_NM) treatment and any other treatment. Participants in the Exocentric Continuous (X_C) treatment exhibited better mean performance than Egocentric Continuous (E_D) treatment by a difference of over 81 seconds at a marginal $p=.075$ level.

Table 5: Total Time in the Maze Statistics

Treatment	Mean (seconds)	Standard Deviation (seconds)	N
C_M	347.4	88.2	20
C_NM	382.0	158.4	20
E_C	442.3	118.8	20
E_D	451.7	154.8	20
X_C	370.1	179.2	20
X_D	415.1	144.5	20
Total	401.5	145.8	120

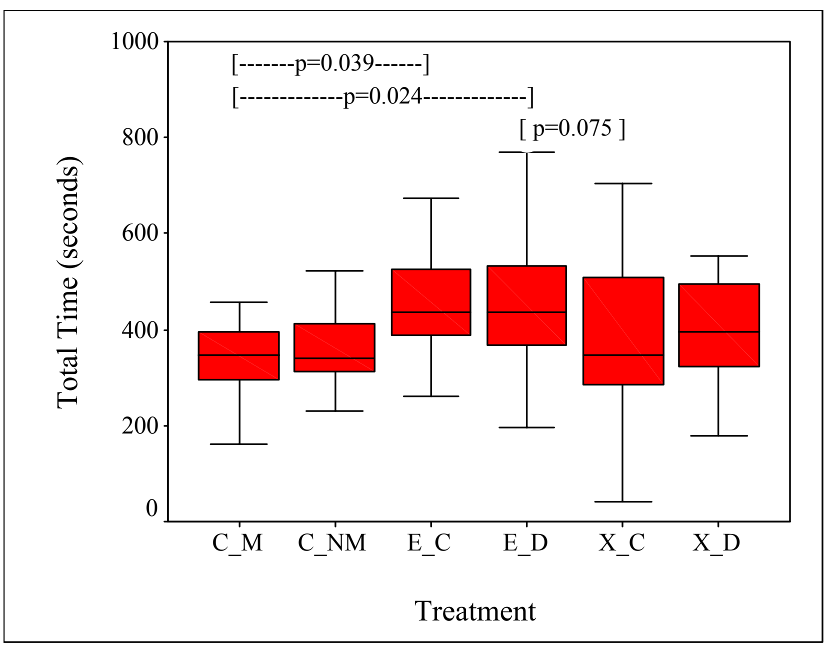


Figure 10: Boxplot of the total time in the maze statistics

Separate ANACOVAs were performed on total time in the maze using scores from the Guilford-Zimmerman Spatial Visualization, Guilford-Zimmerman Spatial Orientation and participant age. Results were significant for participant age, $F(1, 119) = 4.104$, $p < .05$. For age, Levene's Test of Equality of Variance had a significance value of .374 and a significance value of p=.045 for tests of between treatment effects. Table 6 shows the adjusted mean total time in the maze when adjusting for the age covariate and which are statistically significant. The ANACOVA results reveal superior mean performance of both egocentric display treatments when compared to the Exocentric Continuous treatment. The significance for the pairwise Egocentric Continuous and Exocentric Continuous treatments, though, is a marginal p=.094. Also of interest is the introduction of the better average performance of the Control No Map treatment when compared to the Egocentric On Demand treatment (p=.064). Age can affect speed performance.

Table 6: ANACOVA for Total Time in the Maze (Age is Covariate)

Treatment (I)	Treatment (J)	Mean Difference (I-J) (seconds)	P-value
C_M	E_C	-95.8	.035
C_M	E_D	-109.2	.017
C_NM	E_D	-85.0	.064
E_C	X_C	75.8	.094
E_D	X_C	89.3	.050

Data was analyzed on the affect of gender on total time in the maze by considering each treatment separately. In the Control No Map (C_NM) treatment, the average time for females exceeded the average time for males by over 127 seconds at a marginally significant p=.067. This was the only instance when gender had even marginal statistical significance.

Maze Traversal (%)

An alternative way to assess performance outcomes is through consideration of the percent of the maze traversed or covered by each participant traversal. They were instructed to traverse the maze completely, which included all dead ends, prior to retrieving the target object. Therefore, optimal performance is constituted by 100 % of the maze covered. Chi squared analysis was conducted on the distributions using a categorical analysis of success/failure. The success criterion was 100 percent traversal of the maze. A distribution of the success/failure across treatments is shown in Table 7. Pearson Chi squared showed significance at p=.016. The low p value indicates a statistically significant relationship exists between percent covered when considered categorically. Further scrutiny is indicated.

Table 7: Maze Coverage Success/Failure Criteria (Success=100 Percent Maze Traversal)

	Treatment						Total
	C_M	C_NM	E_C	E_D	X_C	X_D	
Failure	15	9	12	6	10	5	57
Success	5	11	8	14	10	15	63
Total	20	20	20	20	20	20	120

The ANOVA was applied to the percentage coverage directly as a continuous dependent variable to reveal any significance between pairwise conditions. The ANOVA reveals interesting and significant results. The between treatment effects were not significant, $F(1, 119) = 1.746$, $p = .130$. Table 8 provides the descriptive statistics (mean and standard deviation) for the overall percent achieved in each condition. As can be seen in Table 8, there was a ceiling effect and no large differences in the mean values. However, the standard deviation for each treatment was quite different revealing some variability for within treatment among participants. For example the standard deviation changed by a factor greater than three between the E_D and X_C treatments indicating greater variability between participants in the X_C treatment than the E_D treatment.

Table 8: Statistical Data for Percent of Maze Traversed

Treatment	Mean (%)	Standard Deviation (%)	N
C_M	93.4	7.4	20
C_NM	96.2	5.2	20
E_C	94.8	5.4	20
E_D	98.2	2.9	20
X_C	93.5	11.9	20
X_D	97.6	5.1	20
Total	95.6	7.0	120

A subsequent pot hoc test was conducted to distill which elements of the treatment conditions were significantly different. As can be seen in Table 9, these post hoc analyses distinguished four pairwise differences that can be subsequently analyzed.

Table 9: ANOVA for Percentage of Maze Traversed

Treatment (I)	Treatment (J)	Mean Difference (I-J) (percent)	p-value
C_M	E_D	-4.8	.030
C_M	X_D	-4.2	.057
E_D	X_C	4.7	.035
X_D	X_C	4.0	.067

Both On Demand display treatments show improved performance when compared to the Control Map treatment (Table 9). Participants in the E_D treatment showed average performance of 98.2% compared with the C_M treatment at 93.4% of the maze covered. Also, participants in the X_D treatment showed average performance of 97.6% compared with the C_M treatment of 93.4% of the maze covered, although the statistical significance was marginal ($p=.057$). Participants in the E_D treatment showed average performance of 98.2% compared with the X_C treatment at 93.5% of the maze covered. Finally, participants using the On Demand display treatment performed better by over 4% than those using the Continuous display treatment on the Exocentric map displays. Participants in the X_D treatment showed average performance of 97.6% of the maze covered compared to the X_C treatment at 93.5% of the maze covered. Statistical significance was marginal ($p=.067$).

Gender was investigated by conducting separate ANOVAs on individual treatments using gender as a fixed factor. The Egocentric Continuous treatment showed marginal statistical significance, $F(1,19) = 3.689$, $p = .071$. In this treatment, males had better average performance than females in covering the maze (98.3% versus 88.8%, respectively). The significance, however, was marginal, at $p=.071$ due to the respective levels of variability in performance with a treatment. This suggestive result is interesting, especially with respect to the known differences between the sexes in spatial processing (see Watson & Kimura, 1991).

Finally, an ANACOVA showed no affect when using scores from the Guilford-Zimmerman Spatial Visualization test, Guilford-Zimmerman Spatial Orientation test, or age.

Weighted Time

Weighted time is defined as the total time in the maze divided by the fractional part of the maze which the individual traversed during ingress (i.e., the percent traversed). The weighted time measure is intended to penalize participants for their failure to completely traverse the maze. It represents a composite measure which thus includes both speed and accuracy which have been only dealt with individually in the results to this point.

An aggregate measure of statistics across treatments for weighted time was conducted and was not significant, $F(5, 119) = 1.245$, $p = .293$. Additional statistical data for the weighted time in the maze by treatment is provided in Table 10 and graphically depicted in Figure 11. There are large variations between treatments in the range of standard deviations which were similarly noted in for Table 8. An ANOVA was conducted on this measure and showed significance. Post hoc tests revealed that the Control Map treatment had an average weighted time value equal to 371.5 seconds compared to 465.0 seconds for the Egocentric Continuous Map treatment (p=.053) and 459.3 seconds for the Egocentric On Demand Map treatment (p=.069). These results reflect poorer mean performance for traversing the maze in both egocentric treatments as compared to the C_M treatment. These differences were the only ones which showed important statistical differences.

Table 10: Weighted Time for Maze Traversal Statistics

Treatment	Mean (seconds)	Standard Deviation (seconds)	N
C_M	371.5	91.5	20
C_NM	397.3	167.8	20
E_C	465.0	117.7	20
E_D	459.3	154.0	20
X_C	394.5	203.3	20
X_D	427.1	148.0	20
Total	419.1	152.0	120

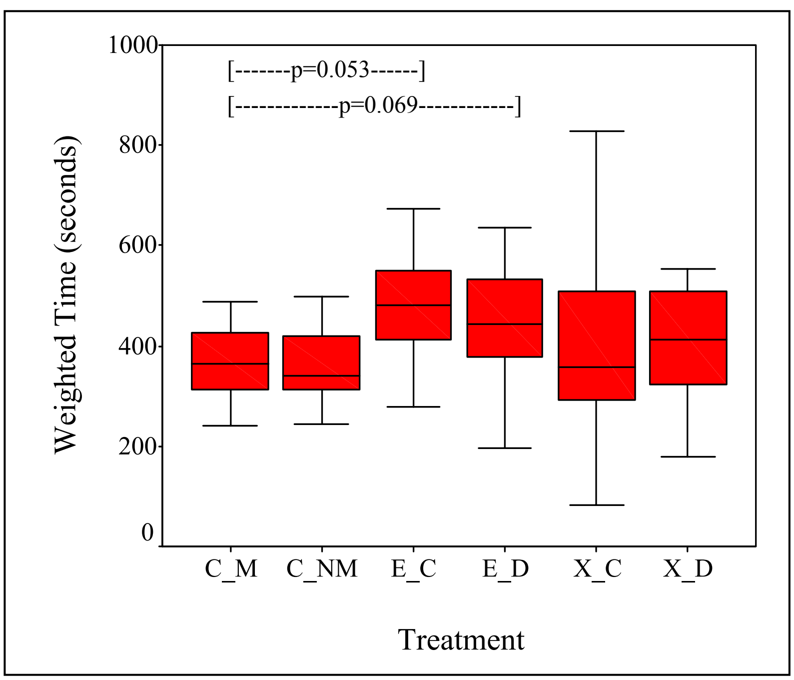


Figure 11: Boxplot of weighted time to traverse the maze statistics

The effect of gender was considered within each treatment using an ANACOVA. The only treatment showing any sex difference in performance was in the Control No Map treatment in which males (334.7 seconds) performed better than females (459.9 seconds) showing an average of more than 125 seconds. However, the statistical significance was a marginal value of p=.096. Although this work is primarily focused on differences between treatments, studying the variation in performance between treatments and variability within treatment scores are areas of potential interest using correlational psychology study methods (Cronbach, 1957).

Separate ANACOVA's were also performed on weighted time in the maze using scores from the Guilford-Zimmerman Spatial Visualization, Guilford-Zimmerman Spatial Orientation and participant age. Results were marginally significant for both participant Guilford-Zimmerman Spatial Visualization scores and age. For Guilford-Zimmerman Spatial Visualization, the Levene's Test of Equality had a significance value of .268 and a significance value of p=.007 for tests of between participant effect. Pairwise comparisons were not significant for Guilford-Zimmerman Spatial Visualization scores.

The ANACOVA for age showed marginal significance. Levene's Test of Equality had a significance value of .221 and a marginal significance value of p=.067 for tests of between participant effect. The estimated marginal means that consider age as the ANACOVA are shown in Table 11. Pairwise comparison of mean weighted times show better average performance by over 94 seconds between the Control Map treatment than the Egocentric Continuous treatment (p=.049). Control Map also had better performance than Egocentric On Demand treatment by over 92 seconds. This latter performance improvement is slightly out of the range for statistical significance (p=.054).

Table 11: ANACOVA for Weighted Time Marginal Mean with Age Covariate= 26.5 years

Treatment	Mean (seconds)	Standard Deviation (seconds)
C_M	373.1	33.5
C_NM	388.7	33.8
E_C	467.3	33.5
E_D	465.6	33.6
X_C	393.3	33.5
X_D	426.6	33.5

Measures of Response Accuracy

Several reflections of responses of accuracy were recorded. The participants had spatial related tasks to conduct both during and after traversing the maze which reflected their personal spatial accuracy assessments. The spatial tasks during traversal of the maze included answering various spatially related questions. Participants also answered spatial recall questions after completing the primary task that can be useful in assessing acquisition of spatially accurate information. Spatial tasks during maze traversal were intended assess spatial abilities and the use of spatially related features available to the participant using BARS. The spatial task in the maze was also intended to increase participant workload to mitigate a ceiling effect resulting from simply having a timed traversal in the maze.

One assessment of accuracy among the various treatments was accomplished by statistical analysis of performance with respect directional and distance estimation by participants during maze traversal. During traversal of the maze and before retrieving the target object, participants had to locate bags that contained spatially oriented questions. There were five bags which had to be visited in an ordinal sequence. The participants were asked questions regarding their Cardinal and Euclidian directional orientation and Euclidian distance to the maze entrance. A second assessment of accuracy comes from the participant's ability to recall which maze they traversed among six choices and to locate the target object within the selected maze. These differing reflections of spatial accuracy were subject to analysis which showed a number of important effects.

Estimation of Euclidian Direction

Statistical data for scores derived ($F(5, 119) = 1.788$, $p = .121$) from the participant's estimate of Euclidian directions (estimate of the direction of the entrance while facing each respective response bag) were aggregated across the five bags. The impacts of the various treatments are reported in Table 12.

Post hoc analysis from the scores in Table 12 revealed the Control Map (C_M) condition exhibits better mean performance than X_C and X_D. The respective values are statistically significant at $p=.043$, and $p=.022$. There are two marginally significant statistical differences. Mean performance by participants in the C_M treatment is superior to mean performance in the E_C treatment ($p=.066$). Also, E_D is superior to X_D at $p=.062$. No statistically significant affect of gender on performance was found when considering each treatment separately. Separate ANACOVA were performed on Euclidian direction scores in the maze using scores from the Guilford-Zimmerman Spatial Visualization, Guilford-Zimmerman Spatial Orientation and participant age. Levene's Test of Equality of Variance for both scores was below .05 indicating that results were not significant because the variances within each covariate are not equal.

Table 12: Statistics of Scores of Participant Estimate of Euclidian Directions

Treatment	Mean Score* (maximum of 15)	Standard Deviation	N
C_M	11.6	4.4	20
C_NM	10.1	5.6	20
E_C	8.6	5.7	20
E_D	10.9	4.2	20
X_C	8.3	5.4	20
X_D	7.9	4.7	20
Total	9.5	5.1	120

* The score for a treatment is derived from a maximum of 3 points per bag. There are 5 bags.

Estimation of Euclidian Distance

The mean and standard deviation of the cumulative scores of participant estimates of Euclidian distance (the distance, in feet, between a particular bag in the maze and the maze entrance) aggregated from each of the five bags in the maze and averaged over each treatment are given in Table 13.

Table 13: Statistics of Scores of Participant Estimate of Euclidian Distance

Treatment	Mean score * (maximum of 15)	Standard Deviation	N
C_M	4.3	4.6	20
C_NM	2.7	2.6	20
E_C	3.8	3.8	20
E_D	1.2	1.8	20
X_C	2.4	2.9	20
X_D	1.9	3.5	20
Total	2.7	3.4	120

* The score for a treatment is derived from a maximum of 3 points per bag. There are 5 bags.

The ANOVA results show an overall $F(5, 119) = 2.504$, $p = .034$ indicating strong significance in the treatment. Further analysis reveals that participants in the Control Map (C_M) treatment exhibit better mean performance than the Exocentric (X_C or X_D) treatments (at $p=.079$ and $p=.023$, respectively). C_M participants also exhibited better average scores than Egocentric On Demand participants in estimating Euclidian distance ($p=.004$). Comparison of scores between the Egocentric treatments shows the Continuous map display results in better estimates of Euclidian distance than the On Demand treatment at $p=.013$. Additionally, average performance for participants in the E_C treatment was better than the X_D treatment, although the statistical significance was marginal ($p=.064$). Guilford-Zimmerman Spatial Visualization, Guilford-Zimmerman Spatial Orientation, and age were not statistically significant covariates. Gender was also not statistically significant when considered for each treatment. Also, the standard deviation for treatments was relatively large when considered with respect to the mean scores in each treatment.

Estimation of Cardinal Directions

While visiting each of the five stations in the maze, participants were asked their Cardinal direction while facing the bag at each station. Responses to questions were scored and then summed across all five stations. The overall results were significant, F (5, 119) = 2.350, p = .045. The relevant statistical data for each treatment is provided below in Table 13.

ANOVA shows significance between the treatment pairs. Post hoc analysis of this data revealed that participants in the Control Map treatment exhibited better average performance than participants in either Egocentric treatment on estimating Cardinal directions (Continuous display at p=.031 and On Demand display at p=.013). Participants using the Exocentric On Demand (X_D) display exhibited better average performance than participants using the Egocentric Continuous (E_C) display (p=.028). Both exocentric treatments, though, exhibited better average performance than the Egocentric On Demand treatment at estimating Cardinal directions. (X_C at p=.067 and X_D at p=.012). The standard deviations in the derived composite scores were generally as large as the mean values among all treatments when compared to the mean indicating large within treatment variations. Guilford-Zimmerman Spatial Visualization, Guilford-Zimmerman Spatial Orientation, and age were not statistically significant covariates.

Table 13: Statistics of Scores of Participant Estimate of Cardinal Directions

Treatment	Mean score * (maximum of 20)	Standard Deviation	N
C_M	9.8	9.2	20
C_NM	7.4	7.2	20
E_C	4.6	6.7	20
E_D	3.8	5.2	20
X_C	8.2	7.5	20
X_D	9.9	8.6	20
Total	7.3	7.7	120

* The score for a treatment is derived from a maximum of 4 points per bag. There are 5 bags.

Consideration of gender in performance within a treatment showed females performing considerable poorer than males in the Egocentric Continuous (E_C) treatment in ANACOVA's. The mean score for females in E_C was 1.6 while the mean score for males was 7.6. The difference was statistically significant at $p=.043$. No other treatments were statistically significant although the overall pattern within each non-controlled experimental treatment was suggestive of better male performance compared to females, as shown in Table 14.

Table 14: Statistics for Gender Affects on Cardinal Direction Scores

	Treatment	Mean	Standard Deviation	N
Female	C_M	10.4	8.9	10
	C_NM	8.0	6.7	10
	E_C	1.6	1.8	10
	E_D	3.8	3.8	10
	X_C	7.2	7.8	10
	X_D	7.6	8.9	10
	Total	6.4	7.2	60
Male	C_M	9.2	10.0	10
	C_NM	6.8	8.0	10
	E_C	7.6	8.5	10
	E_D	3.8	6.5	10
	X_C	9.2	7.6	10
	X_D	12.2	8.1.	10
	Total	8.1	8.2	60
Total	C_M	9.8	9.2	20
	C_NM	7.4	7.2	20
	E_C	4.6	6.7	20
	E_D	3.8	5.2	20
	X_C	8.2	7.5	20
	X_D	9.9	8.6	20

Spatial Recall

Spatial recall scores were analyzed using the Chi squared procedures. The data was viewed as categorical within treatment. Three Chi squared analyses were conducted on different distributions of spatial recall:

- Identification of the correct maze among the six possibilities in the spatial recall multiple choice test
- correctly identifying the object location in the maze (regardless of whether the correct maze was chosen)
- Identifying both the correct maze and object location in the maze

No Chi squared results were significant. Pearson Chi squared significance was, respectively, p=.975, p=.851, and p=.987. Additionally, the second test for object location had 50 % of the cells having less than the minimum count of five entries in the cell.

Gender and Spatial Performance

During the course of this work a number of trends were observed. Principal among these was a gender and spatial condition interaction which has been previously commented on (Watson & Kamura, 1991). Guilford Zimmerman tests of spatial visualization and spatial orientation were given to all participants prior to traversing the maze. Males perform better than females in spatial visualization and orientation tasks. Table 15 provides the descriptive statistics for the tests. The overall statistic for the tests of between subjects effects for Guillford-Zimmer Spatial Visualization is $F(1, 119) = 12.267$, p = .001. The statistic for the Guillford-Zimmerman

Spatial Orientation is *F*(1, 119) = 14.031, $p < .001$. Of note is the similar between standard deviation for any of the scores and the difference in means between sexes for a specific Guillford-Zimmerman test score.

Table 15: Guillford Zimmerman Descriptive Statistics With Respect to Gender

	Gender	Mean	Standard Deviation	N
G-Z_SV	Female	14.28	6.783	60
	Male	19.05	8.071	60
	Total	16.67	7.799	120
G-Z_SO	Female	16.67	7.557	60
	Male	22.70	9.927	60
	Total	19.68	9.293	120

Correlations were conducted on combinations of participant scores on estimating Euclidian distance, estimating Euclidian direction, estimating Cardinal direction, maze traversal time, percent of maze covered, Guilford-Zimmerman Spatial Visualization, and Guilford-Zimmerman Spatial Orientation (Table 16). The strongest correlation was between Guilford-Zimmerman Spatial Visualization and Orientation. This pair had Pearson Correlation = .625 with p=.000. There were several other correlations indicating a slight linear dependence such as the weak correlation between Cardinal and Euclidian scores (Pearson Correlation = .388, $p \leq .001$.

Table 16: Correlation Statistics for Dependent Variables

	SumEDir	SumEDis	SumCDir	Total Time	% Covered	G-Z SV	G-Z SO
Euclidian Dir. (SUMEDir)	1	.204*	.388**	.195*	.179*	.270**	.189*
Euclidian Dist. (SUMEDis)	.204*	1	.082	.090	.062	.166	.159
Cardinal Dir. (SUMCDir)	.388**	.082	1	.150	.118	.316**	.296**
Total Time	.195*	.090	.150	1	.215*	-.284**	-.211*
% Covered	.179*	.062	.118	.215*	1	-.002	.021
Guillford-Zimmerman (G-Z_SV)	.270**	.166	.316**	-.284**	-.002	1	.625**
Guillford-Zimmerman (G-Z_SO)	.189*	.159	.296**	-.211*	.021	.625**	1

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).
N = 120 for each factor

Map Rotation

Question 40 of the adapted Immersive Tendency Questionnaire asked participants their tendency to rotate maps. A seven point Lickert scale was used with 1 representing no rotation and 7 representing always rotating maps. An ANOVA was conducted using the rotation scores as fixed factors and total traversal time, percent of the maze covered, and weighted time as dependent variables. Treatment type was not considered. Statistically significant differences in scoring on total time and weighted time in the maze were seen between those responding to pairs of score 2 and 6 (that is, nearly never rotating and rotating quite often). Differences in total time in the maze benefited those not tending to rotate maps (score of 2) by over 107 seconds as compared to those tending to rotate maps (score of 6) at $p=.030$. Those indicating a score of 2 had a mean total time of 358.1 seconds and those indicating a score of 6 having a mean time of 465.2 seconds. Similarly, when considering the weighted time through the maze, those indicating a score of 2 had an average of over 119 seconds faster time than those indicating a score of 6 at $p=.020$. Those scoring 2 had an average weighted time of 372.1 seconds while those indicating a 6 had an average time of 491.5 seconds. The above analysis is suggestive of users having a fixed mental model(s) of maps and their reporting those maps as spatially fixed. The results also suggest the possibility for creating training programs for map users and to develop new map presentation methods that facilitate building dynamic mental models.

Immersive Tendencies Questionnaire/Presence Questionnaire

Analysis was accomplished on question pairs from the adapted Immersive Tendencies Questionnaire (ITQ) and Presence Questionnaire (PQ). Three questions from each questionnaire were used to conduct pairwise comparisons. The questions were as follows:

- ITQ #1: Do you easily become deeply involved in movies or TV dramas?
- ITQ #9: How good are you at blocking out external distractions when you are involved in something?
- ITQ #10: When watching sports, do you ever become so involved in the game that you react as if you were one of the players?
- PQ #22: How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?
- PQ #30: Were there moments during the experience when you felt completely focused on the task or environment?
- PQ #37: Did augmentation interfere with your tasks?

Combinations of each ITQ and PQ were run using Chi squared and statistical correlation methods. No results were found to be statistically significant.

Additional correlations were performed by comparing scores on Euclidian distance, Cardinal and Euclidian directions, traversal time, percent of the maze covered, and weighted time with the adapted Presence Questionnaire numbers 22, 30, and 37. No statistical significant results were found.

Addressing Hypotheses

Hypotheses were formulated on both quantitative and qualitative aspects of expected improvements in human performance when using the BARS. While the primary task was maze traversal time and accuracy, the experimental method was structured to allow consideration of hypotheses which could support evaluation of selective aspects of human performance during and after maze traversal.

Wayfinding Performance

The first hypothesis was that wayfinding performance using a real time, hands-free AR map display will improve human performance compared to the control conditions (there are two variants of control conditions).

Prediction 1: Performance will be improved by reducing the time by at least 50 percent the weighted time needed to complete the primary way finding task as compared to both control conditions.

Prediction 2: Performance will be improved by at least 50 percent in the accuracy of the primary wayfinding task performance as compared to both control conditions.

Prediction 3: Performance will be improved by at least 50 percent in the accuracy of a secondary spatially oriented embedded task as compared to both control conditions.

Traversal Time Performance

Performance with respect to time (Prediction 1) was not met. The Control Map treatment was statistically better than either the Egocentric Continuous or Egocentric On Demand display treatments. This performance difference was evident in both the traversal time and the subsequent derived weighted time. Recall that weighted time is the total time divided by the percent of the maze traversed by each participant. When age was factored into the analysis as a covariate, results using an ANACOVA showed that the Control No Map had better average performance than the Egocentric On Demand display treatment at a marginal statistical significance level ($p=.064$). Also, performance in the Exocentric Continuous treatment was better than either Egocentric treatment, although the statistical significance was also marginally significant ($p=.094$).

Weighted time, the total time divided by the percent of the maze traversed before retrieving the target object, is intended to penalize those participants that do not completely traverse the maze. As with the total time in the maze, participants in the Control Map treatment exhibited statistically better mean performance than either egocentric treatment. When an ANACOVA considering age was used, the Egocentric On Demand treatment had marginal significance.

There are several issues that might have influenced the results. First, the size of the maze was small due to constraints in tracking equipment. Maze size limited the number of turns and alternative paths which could be implemented. It was possible that participants in the Control Map treatment could memorize the maze and therefore negate the advantage of technology. This possibility is crucial since the use of memory as a strategy becomes progressively less effective as the environment increases in complexity. Secondly, by receiving the map at the entrance of the maze, participants could immediately begin planning a route. They could have done the same thing with the experimental treatments, but this was not required because the display was always available. Some levels of route planning are beneficial in supporting navigation (Thorndyke & Hayes-Roth, 1982). A short period of route planning, therefore, might be beneficial to those using AR equipment. Consequently, training with AR equipment to facilitate performance might take both general and technology specific forms.

Regardless of route planning, though, larger, more complex mazes would eventually saturate working memory limiting the memorization of routes. Other secondary tasks could also serve to saturate working memory. This saturation point is where the performance benefits of AR equipment would become more apparent and it is expected that such a pattern would emerge during real world deployment. It is also noteworthy with respect to speed that in no circumstance was the Egocentric condition better than an Exocentric or Control treatment. Therefore, it is recommended that Egocentric treatments might not be best when the time to traverse an area is important.

Wayfinding Accuracy

AR was beneficial when considering the second prediction that the accuracy of the navigating and wayfinding task would show improvements when compared to the control condition. The percent of the maze covered before retrieving the target object reflects wayfinding accuracy. Both the Egocentric and Exocentric On Demand treatments showed significantly higher performance when compared to the Control Map treatment. The Exocentric

On Demand treatment had a marginal statistical significance (p=.057). The average improvements in absolute terms were respectively 4.8% and 4.2%. Although these gains may seem modest, they occurred against a situation where a ceiling effect exhibited a strong impact. The baseline Control Map treatment had an average coverage of the maze of 93.4%. The 4.8% and 4.2% improvements exceed the 50% prediction if only the remaining 6.6 percent (=100-93.4) is considered available for improvements from the baseline treatment. There are several interesting aspects of the second prediction. First, routes and coverage are different. While a participant may traverse a small area faster by looking at a map versus AR, the participant will not cover the entire area as well. The second interesting conclusion is that the continuous display condition performs worse than the on demand condition. In addition to the above treatment pairs, the On Demand condition is statistically better in either the Egocentric or the Exocentric when compared to the Exocentric Continuous treatment (p=.035 and p=067, respectively). The cause of the better performance in on demand treatments could be due to divided attention issues which might arise when using the continuous treatment.

Spatial Accuracy

The third prediction from the first hypothesis was that performance would be improved by at least 50% in the accuracy of a secondary spatially oriented embedded task as compared to both control conditions. Spatial accuracy was measured by participant responses to spatially oriented questions in the bags during traversal and two spatial recall questions after maze traversal. The spatially oriented questions during maze traversal included questions on Euclidian and Cardinal directions and Euclidian distance. A question regarding Cardinal distance was not scored because it was not clearly worded making responses and evaluations problematic.

The predictions for improved spatial performance using AR were not met. In no treatment was spatial accuracy task performance statistically better in AR than the Control Map treatment. As with maze traversal, Control Map might be a superior solution in this case because of the simplicity of the maze. There are a limited number of bags and a limited number of turns in this experiment resulting in one's ability to keep location in working memory. The large standard deviation, though, causes concern about future activity oriented to a different experimental approach either with a more complicated maze or a different strategy for gathering spatial data during traversal.

Euclidian Accuracy

The second hypothesis is that within the AR condition, the use of egocentric moving map as compared to an exocentric fixed North-up map would reduce the time and improve recall accuracy on secondary task. This hypothesis was supported with respect to accuracy as reflected by scoring on Euclidian questions in the bags for selective treatment pairs. The results, though, were not specific with respect to display activation. The Egocentric On Demand display mode improved average performance when compared to Exocentric On Demand treatment. Alternatively, the Egocentric Continuous display modem improved average performance when compared to Exocentric On Demand Treatment. These results, however, are highly consistent with the work of Arthur and Hancock (2001) with respect to map orientation, but extended from virtual environments to AR. It is also noteworthy that the Egocentric Continuous treatment showed better performance than the Egocentric On Demand treatment for participants estimation of distance. These differing results cause some uncertainty with respect to which Egocentric treatment would be better if AR were to be deployed. These inconsistencies could be due, in part, to aliasing and other graphics anomalies visible in rendering egocentric graphics..

Cardinal Accuracy

The third hypothesis is that within the AR conditions, the use of exocentric fixed North-up map as compared to a egocentric moving map will reduce the time and will improve recall accuracy on secondary task, in terms of cardinal directions, by providing real time North-up view of the space. The hypothesis was not supported with respect to traversal time. The accuracy portion of this hypothesis was evaluated on the basis of directional information in the bags. Distance estimation was not used due to ambiguity in the question and measurement method. This hypothesis was also substantially met on directional values with both the Exocentric On Demand and Exocentric Continuous modes performing better than the Egocentric On Demand. Also, the Exocentric On Demand performed better than the Egocentric Continuous treatment. This hypothesis also confirms and extends the work of Arthur and Hancock (2001) to AR and confirms the problems with the Egocentric Continuous treatment. The conclusions are all statistically significant, except that the Exocentric Continuous and Egocentric On Demand treatments had marginal significance (p=.067).

On Demand Versus Continuous Displays

The fourth hypothesis was that within the AR conditions, the use of an On Demand map display as compared to a continuously on map display will reduce the time and improve accuracy on the primary wayfinding task and the secondary embedded task by providing the participant with control over the times when the map is needed as a reference. This hypothesis was partially met. The percentage of the maze covered condition showed improved performance by the Exocentric On Demand condition when compared to the Exocentric Continuous treatment. However, Euclidian distance estimation showed the Egocentric Continuous map treatment had improved performance compared to the Egocentric On Demand treatment. There were some instances where a continuous display mode resulted in better average performance across display orientations. For example, the Exocentric Continuous treatment had better average performance than the Egocentric On Demand treatment for the total time in the maze (average difference of over 89 seconds when age is considered as a covariate). This comparison might not be totally fair because of anomalies in the Egocentric Continuous treatment. Aliasing and graphics anomalies could have been a distraction to participants. The better performance resulting from an On Demand map provides participants with increased decision latitude and reduced mental strain which is consistent with improved performance (Karasek, 1979).

Presence

It was hypothesized that participants scoring higher in objective performance would positively correlate with higher subjective ratings of mobile AR utility. Co-evolution and presence require a threshold level of immersive fidelity and task performance requirements (e.g., heightened stress level) to be triggered. This hypothesis was not met. No statistically significant results emerged from selective pairing of the adapted Immersive Tendencies Questionnaire (ITQ) and adapted Presence Questionnaire (PQ), nor between the PQ and objective performance as measured by the traversal time, percent of maze covered, weighted time, Cardinal or Euclidian directions, or Euclidian distance scores. Co-evolution or presence did not appear to occur. This could be due to many factors, including the form factor of the BARS equipment, but more likely was task related. Presence for AR will be influenced by social interactions in the AR environment and the ability to influence and be influenced by the environment. There were no social interactions and the task likely did not trigger interactions with the environment. This should be overcome in time with better technology and with more compelling tasks.

Gender

As noted in the literature, Watson and Kimura (1991) studied sex differences on a specific spatial orientation task of throwing and intercepting objects. Males performed better than females. Lawton and Morin (1999) confirmed this finding for pointing accuracy. It was not known, though, at the onset of this research that these findings applied to the types of navigation experiments and technology used here. The results of this research indicate that where there is statistical significance, males perform better than females in this AR experiment. In most cases, though, the statistical significance was marginal, possibly due to variability in the BARS equipment. For example, on average, males took less time to cover the maze and covered a higher percentage of the maze than females. In traversal time, the difference was 127 seconds (p=.067) and nearly 10 percent more absolute coverage of the maze (p=.071). Males also performed better than females in weighted time by over 125 seconds (p=.096). Finally, males performed better than females on both Guilford-Zimmerman Spatial Orientation and Guilford-Zimmerman Spatial Visualization. Guilford-Zimmerman test scores might be a good indicator of navigational performance.

Map Rotation

The adapted Immersive Tendencies Questionnaire (Appendix C) asked whether participants have a tendency to rotate maps. A score of one indicated the participant never rotates a map and score of seven indicated one always rotates maps. There was a significant correlation between participants indicating scores of two and scores of six with benefits going to those who indicated that they do not rotate maps for total time to traverse the maze (107 seconds better performance than those who rotate maps) and the weighted time in the maze (119 seconds better performance for those who do not rotate maps). This result would tend to favor the Exocentric treatment. This result would also suggest some training strategies in using fixed maps for map rotators as an alternate means for creating more effective cognitive models of spaces.

CHAPTER FIVE: DISCUSSION

This study examined the changes in performance in navigation and wayfinding when using mobile augmented reality (AR) as well as two non-augmented control conditions. The navigation and wayfinding tasks represented a particular type of navigation this work has termed search and rescue. In a search and rescue task, individuals are not afforded a separate opportunity to study and learn the space they are to traverse. The time the individual spends learning their traversal route is considered part of the total time to find their objective. Also, search and rescue tasks require the individual to consider accuracy in traversing their space as well as the time element. In this respect, search and rescue has similarities to the sport of orienteering. AR could become a useful tool for orienteering and orienteering could be used as environment for evaluating the type of AR used in this work.

The mobile augmented reality system used in this study is known as the Battlefield Augmented Reality System (BARS). It was developed by the Naval Research Laboratory and provided for this research effort by the Army Research Institute (ARI). The BARS implementation used a monocular display for augmentation by portraying a maze to the participant in various formats. A physical maze was also constructed. The maze had no overt landmarks. As defined by this work, the augmented reality (AR) system implemented was on the low end of the augmented reality spectrum of performance. The system was spatially registered and operated in real time, but only two dimensional information was needed for the participant to conduct their task. The map displayed information that was distinct and not graphically fused with the real maze.

As a primary task, participants were asked to completely traverse the maze, find the target object and exit the maze. They were informed that they would be measured on time and accuracy. Additional secondary tasks included answering spatially oriented questions during maze traversal and after traversing the maze. Maze information was presented to participants in this research in combinations of egocentric and exocentric formats and displayed continuously and on demand. Two control conditions were used; one with a paper map and one with only a compass.

Several significant findings can be concluded from this work. The findings are based upon the specific configuration of AR used and the search and rescue task defined in this work. Principal among these findings is the benefit of the On Demand display mode compared to a Continuous display. This benefit was particularly pronounced in the percentage of the maze covered treatment where all statistically significant results showed benefits of an On Demand Display. The same benefit from an On Demand display mode was shown in the statistically relevant scores of accuracy in Euclidian and Cardinal direction estimation. The Continuous display mode was beneficial in a limited number of treatments and circumstances. In particular, it was beneficial to participant scores resulting from estimating Euclidian distance between different egocentric treatments. This effect could be due to observed anomalies in the rendering and display of the egocentric treatment. The anomalies consisted of aliasing and image jerkiness.

Another significant finding resulting from this work is the relationship between performance and a participant noting whether or not they have a tendency to rotate maps. Those participants reporting a tendency to use fix maps scored better on maze traversal time treatments than those who reported to rotate maps. These results are independent of the participant being presented with a fixed or rotating map in the experimental treatments. This finding appears to indicate a link between a participant's mental model and their performance on navigational tasks. It is consistent with the work of Arthur (1996) and Arthur and Hancock (2001). These results extend those works by indicating a predisposition by some individuals to store spatial information in fixed views thereby facilitating better navigational performance than those who are not predisposed to store fixed views. Additionally, this finding may be suggestive of investigating whether training may be beneficial to the formation better mental models of spatial relationships as well as investigating how AR might be an effective training or memory tool.

The effect of gender on performance using AR was pervasive in this work with males scoring better than females in all areas of statistical significance. This gender effect was evident in both navigational tasks and tests of spatial orientation and spatial visualization as well as on the percentage of the maze traversed conditions. Males also exhibited better performance than females on weighted time through the maze in the Control No Map treatment where only a compass is available. There was not, though, a statistical correlation between pre hoc questionnaires and time or accuracy performance indicating limited utility of the Guillford Zimmerman tests as a predictor of navigational performance. This could be due the orientation of the Guillford Zimmerman tests towards evaluating spatial abilities using small three dimensional items or static scenes as contrasted with navigation tasks which entail locomotion.

There was no quantitative evidence of presence resulting from this work. This is not totally unexpected considering both the components that can affect presence and the theoretical framework suggested by Goldiez and Dawson (2004). Heeter (1992) describes three components of presence; self, social, and environmental. The primary emphasis in presence research in virtual environments has been oriented to the representation of self in the virtual environment. Self presence is not an issue in AR because one is located in the real world. Therefore, the other two factors are hypothesized to be critical to introducing presence (Goldiez & Dawson, 2004). This work did not include social interactions or environmental affects between the participant and the augmented or real environment. The compelling nature of the task can also cause the onset of presence even using modest equipment. The search and rescue task, including an inducement of a reward for the best performer in a treatment was intended to be compelling and introduced observed activities by some participants that might qualitatively indicate or trigger presence. For example, one participant ran through the maze completing the task in approximately 45 seconds. Others expressed stress in needing to complete tasks quickly and accurately. Another factor that can influence presence is the naturalness and ease of use for the equipment. BARS has made great improvements in its physical design from previous versions, but it is still not a natural system for many users to wear or use. Continued improvements in BARS functionality and performance will mitigate these limitations, but such limitations will exist until AR systems achieve wider deployment. Familiarization with BARS might also be helpful.

The maze used in this work was limited in size allowing memorization of routes. While participants were not informed of the need to memorize their route, those participants given the paper map had to study and relinquish before proceeding, while participants wearing the BARS proceeded immediately to traverse the maze when allowed to and generally spent no time studying the augmented reality map before traversing the maze. There is an implied requirement for the participant to memorize the paper map. This difference in strategy could have affected results with BARS users not planning their route in advance and employing a strategy of repeatedly planning their route and assessing their progress. Pre-training using BARS might mitigate these differences. In any case, increased maze and navigational complexity will show the benefits of AR technology as working memory becomes saturated.

Design Guidelines

The following are design guidelines for the implementation from this research:

- Spatial memory load should be minimized, especially for older users.
- On demand displays should be used for AR to minimize divided attention and improve performance.
- The task should determine the type of display orientation used. If Cardinal directions are needed to be communicated to others, the Exocentric orientation is preferred. If Euclidian directions and distances need to be communicated to others, an Egocentric display mode should be used. If the task is mixed, it is better to use an Exocentric display.
- Augmented display contents should be either integral to the scene or be minimized to avoid divided attention issues for the user. The real world will usually prevail to a participant confronted with and augmented and real image. Therefore, integral augmented scene contents should be indistinguishable from the actual scene. This will minimize issues with divided attention.
- Visual anomalies should be minimized in AR. This includes aliasing, image tearing, and image popping.
- Training is needed on use of AR equipment in general and to accommodate variation in gender, in particular.

- Training for AR should consider technology specific instructions in conjunction with environmental familiarization and specific task.

Future Research

Research should be considered in the context of availability of prototypes and the symbiosis between technologically oriented and human performance oriented research. The categorization of AR's in this work into three groups is conducive for research and application planning purposes. The categories include augmentation that provides information that is in the environment but not visible without augmentation; augmentation where computer generated information is added to the real world and is intentionally distinguishable from the real world; and augmentation where computer generated information is added to the real world and indistinguishable from it. The categorization scheme relates to the type of AR technology employed, research focus, and potential utilization. For example, this work has involved investigations into the second category of AR, where information is added but not part of the real world. The availability of technology in this work allows one to focus on interface issues. Conversely, the third category, where information is added to and indistinguishable from the real world would cause a research focus oriented to breakthrough technological factors related to, for example, rendering processes to achieve sufficient graphics quality. Wizard of Oz methods can be carefully employed on the third category as a means to involve human users, but extreme care should be employed to keep user and technological research in balance.

At this time several areas of research would benefit mobile augmented reality's maturation in the area of better facilitating human performance in navigational tasks using BARS type of AR equipment. One area is hands free operation. Because On Demand displays result in better overall performance, there should be a mechanism that facilitates hands free operation. The implementation used in this research was with a keyboard attached to the participant's forearm. This configuration is actually a two handed operation because the opposite hand is needed to depress the display activation key on the user's forearm. Approaches should investigate one handed devices as well as automated tools, such as providing a map when the participant is stationary.

Multi modal interactions could also be beneficial in AR. The implementations of AR, to date, have been visual. Other modalities, such as audio are worthwhile to investigate from both technological and human performance perspectives. There are potential benefits with respect to operational considerations, impacts on working memory, and with respect to enhanced human computer interaction. Audio, in particular, can be accomplished with minimal adverse technical impact, but could impact usability with respect to hearing ambient sounds.

Tracking technologies continue to be problematic. This area was noted by Barfield and Caudell (2001) as well as others. Simply stated, the approach to tracking should become more robust and pervasive. Systems either have range and lack precision or have limited range with good precision. Improvements are needed in both axes. Also strategies for placement on the user and reduction in system noise warrant exploration.

Survey instruments that capture participant experiences in AR need further development. This is especially important in the area of presence which is largely unexplored in AR and can be used as a guide for some aspects of AR effectiveness. Surveys should cover a sufficent range of technologies and experimental tasks to be useful. In particular, experimental tasks involving real world situations should be able to be efficiently captured in surveys.

Research in AR specific applications or application domains should be further explored. This dissertation suggests that virtual environments (VE) can be considered a subset of AR. VE is used in training, engineering design, and entertainment. Therefore, one could configure an AR system as fully immersive and conduct evaluations in these various application domains. However, AR offers other options that have not been fully explored, such as inserting artificial entities into the real world and using these entities to enhance the real world in, for example, training. Mobile AR offers the potential for a smaller footprint and improved mobility when compared to VEs.

The AR research this work has conducted involves experiments where human performance improvements that can be realized using mobile AR for search and rescue navigation. Data from this and other experiments can form the basis of research in creating predictive models of human performance in navigation or other tasks as well as considering variations in AR technologies. Such research can yield methods to efficiently pre-screen AR users to automatically provide them with the features most useful to the task and user characteristics. These features might be particularly useful where age and gender are variations are expected.

Navigation using AR in large spaces is unexplored. The maze used in this experiment was small due to available equipment. The boundaries of where human working memory becomes saturated, how egocentric and exocentric displays can be used in large areas, the use of landmarks (virtual and real), and other strategies need further research. In a closely related area are issues with maps of large areas. Map design for automatic manipulation, optimal size, the use of ancillary features such as zoom, and feature types (fixed and those that the user can insert) are all open research issues.

Practical Applications

Mobile augmented reality is in its infancy with respect to applications. This research has shown that it can be an effective tool in search and rescue navigation tasks. It offers the opportunity to have hands free operation and provide an alternative to using working memory for navigational tasks.

In hypothesizing potential applications for AR, it is important to consider the work of Christensen (2003) who indicates that the initial performance of new technologies is often inferior to current systems or approaches. This is the situation in some aspects of mobile AR studied in this research, such as speed of traversal in the maze with augmentation when compared to a paper map. However, Christensen (2003) points out that many innovations have far more growth potential than the current approach being use. This appears to be obvious in the case of mobile AR, due to natural limitations in maps and human working memory.

Potential application domains for mobile AR are for public safety, security, military, entertainment, consumer products, training, and medicine. One can envision a military person, fireman, or police officer using mobile augmented reality to navigate in a smoke filled building, building clearing operation, or hostage rescue. Hands free operation would be needed to hold a weapon, hose, or other appliance. Likewise security personnel could use mobile AR to navigate a large perimeter or complex of structures. Entertainment and consumer applications could involve self guided tours through historical areas or navigating to find specific items in large stores. Training applications include military operations in urban terrain and team training. There are several medical applications for AR. One is to overlay of information during medical procedures. AR offers the possibility to help guide surgeons to their objective through overlaying of pathways or by adding imagery to augment an open wound site. AR also offers the opportunity to augment cognitive or perceptual processes, such as sight or navigation that can become diminished by aging or by physiological events.

AR is the next technological step from virtual reality. Advances in computing, packaging, and measuring human performance offer the opportunity for AR to subsume more traditional computer simulations and virtual reality. Many of the problems affecting traditional simulation and virtual reality systems relate to limitations in replicating the real world. AR affords the simulation developer the opportunity to selectively use the benefits of the real world and add information and interactivity in targeted areas. Improvements in ubiquitous computing, separate from AR, should benefit its development. Included here are Radio Frequency Identification (so-called RFID) technology, wireless computer networking, and global positioning systems. Coupling thee improvements in ubiquitous computing with packaging improvements and proper human factors design offers human worn AR that is available on demand without the need for a special infrastructure supporting the environment.

Mobile augmented reality is an exciting technology with a large potential to improve human performance in many tasks. The technology builds upon a solid foundation in virtual environments. This work has provided evidence of AR's benefits in the area of search and rescue wayfinding. It has also provided a method for further experimentation across a variety of fronts.

APPENDIX A: QUESTIONS IN MAZE STATION BAGS AND ANSWERS

Participant # _______________   Date: _________________

So far, how *easy or difficult* have you found it to traverse the maze using the augmented reality equipment?

1	2	3	4	5	6	7

VERY DIFFICULT			FAIRLY EASY			VERY EASY

____

Participant # _______________ Date: _________________

Please answer the following multiple-choice question:

1. While facing the wall where the bag with the clipboard is placed, indicate your heading (the direction of where your nose is pointing) in terms of Cardinal directions (East, West, North, South, South-East, etc.):

    a. East
    b. West
    c. North
    d. South
    e. North-East
    f. South-East
    g. North-West
    h. South-West

2. What is the total distance you have traveled (in feet) since entering the maze?

    ____

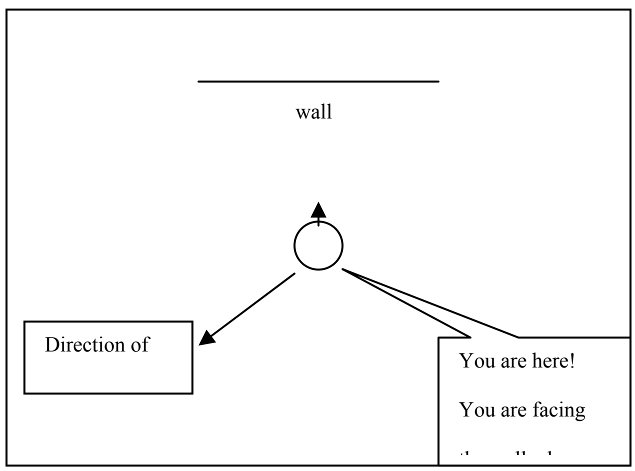
Participant # _____________ __   Date: ___________________

Please answer the following two questions:

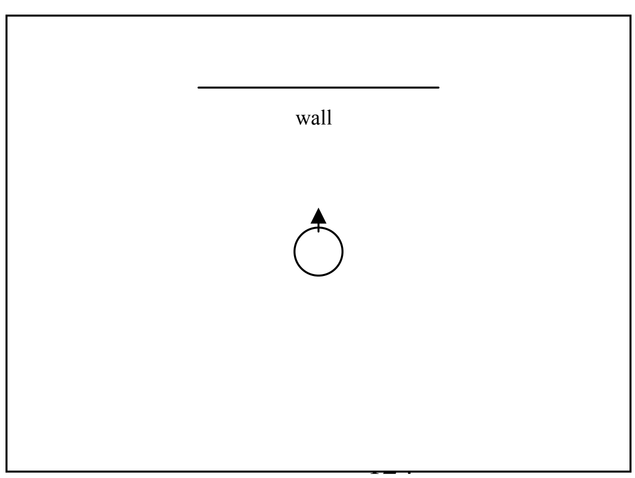
3. In the box below indicate the approximate distance (straight line or 'as the crow flies') to the entrance of the maze.

[   ] ft.

4. In the space provided below while facing the wall where the bag with the clipboard is placed, draw an arrow toward the entrance of the maze (see example below).



Example!



ANSWER KEY CONDITION:___________________

	Q2 (Eu. Dist.)	Q3 (Eu. Dir.)	Q4 (Card. Dir)
St 1	12 ft	↗	South
St 2	14 ft	↖	West
St 3	10 ft	↖	West
St 4	12.5 ft	↙	North
St 5	14 ft	↘	East

traversal sequence: okay ____

other_____________

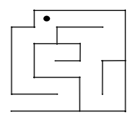
Maze question	#5 is the correct maze design

APPENDIX B: SPATIAL RECALL QUESTIONS

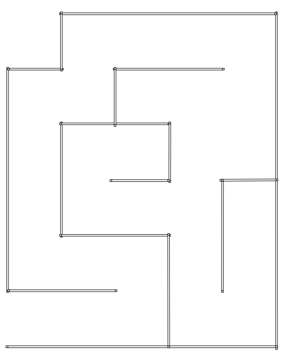
Participant #: Date:

1. Please circle the maze number to indicate which one of the mazes shown below is the one you just traversed:

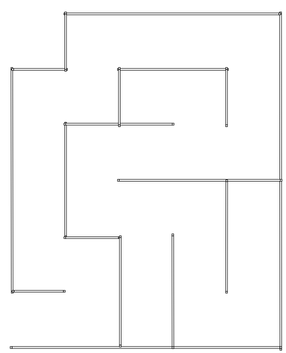
#1 #2 #3 #4 #5 #6



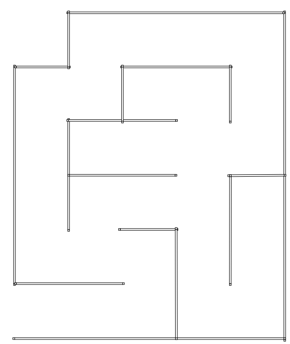
Example



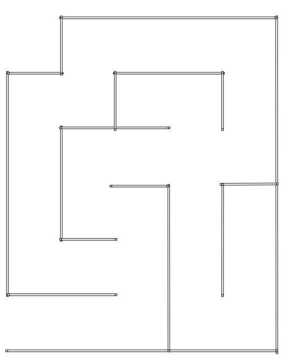
1



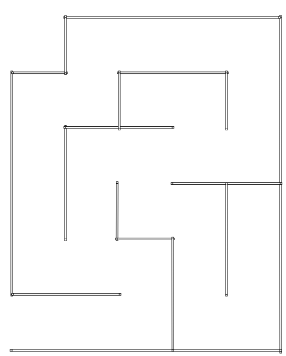
2



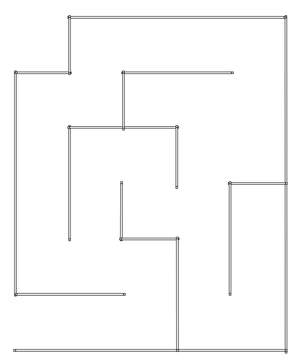
3



4



5



6

APPENDIX C: ADAPTED IMMERSIVE TENDENCIES QUESTIONNAIRE (ITQ)

IMMERSIVE TENDENCIES QUESTIONNAIRE

(Witmer & Singer, Version 3.01, September 1996. Adapted by B. Goldiez, May, 2004)

Indicate your preferred answer by marking an "X" in the appropriate box of the seven point scale. Please consider the entire scale when making your responses, as the intermediate levels may apply. For example, if your response is once or twice, the second box from the left should be marked. If your response is many times but not extremely often, then the sixth (or second box from the right) should be marked.

1. Do you easily become deeply involved in movies or TV dramas?

NEVER OCCASIONALLY OFTEN

2. Do you ever become so involved in a television program, movie or book that people have problems getting your attention?

NEVER OCCASIONALLY OFTEN

3. How mentally alert do you feel at the present time?

NOT ALERT MODERATELY FULLY ALERT

~~4. Do you ever become so involved in a movie that you are not aware of things happening around you?~~

~~5. How frequently do you find yourself closely identifying with the characters in a story line?~~

~~OCCASIONALLY OFTEN~~

6. Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?

NEVER OCCASIONALLY OFTEN

7. What kind of books do you read most frequently? (CIRCLE ONE ITEM ONLY!)

Spy novels	Fantasies	Science fiction
Adventure novels	Romance novels	Historical novels
Westerns	Mysteries	Other fiction
Biographies	Autobiographies	Other non-fiction

8. How physically fit do you feel today?

NOT FIT MODERATELY FIT EXTREMELY FIT

9. How good are you at blocking out external distractions when you are involved in something?

NOT VERY GOOD SOMEWHAT GOOD VERY GOOD

10. When watching sports, do you ever become so involved in the game that you react as if you were one of the players?

NEVER OCCASIONALLY OFTEN

11. Do you ever become so involved in a daydream that you are not aware of things happening around you?

NEVER OCCASIONALLY OFTEN

12. Do you ever have dreams that are so real that you fell disoriented when you awake?

NEVER OCCASIONALLY OFTEN

13. When playing sports, do you become so involved in the game that you lose track of time?

NEVER OCCASIONALLY OFTEN

14. How well do you concentrate on enjoyable activities?

NOT AT ALL MODERATLY WELL VERY WELL

15. How often do you play arcade or video games? (OFTEN should be taken to mean every day or every two days, on average)

NEVER OCCASIONALLY OFTEN

~~16. Have you ever gotten excited during a chase or fight scene on TV or in the movies?~~

~~NEVER OCCASIONALLY OFTEN~~

17. Have you ever gotten scared by something happening on a TV show or in a movie?

NEVER OCCASIONALLY OFTEN

18. Have you ever remained apprehensive of fearful long after watching a scary movie?

NEVER OCCASIONALLY OFTEN

19. Do you ever become so involved in doing something that you lose all track of time?

NEVER OCCASIONALLY OFTEN

20. On average, how many books do you read for enjoyment in a month?

NONE ONE TWO THREE FOUR FIVE MORE

21. Do you ever get involved in projects or tasks, to the exclusion of other activities?

NEVER OCCASIONALLY OFTEN

22. How easily can you switch from the activity in which you are currently involved to a new and completely different activity?

NOT SO EASILY	FAIRLY EASILY	QUITE EASILY

23. How often do you try new restaurants or new foods when presented with the opportunity?

NEVER	OCCASIONALLY	FREQUENTLY

24. How frequently do you volunteer to serve on committees, planning groups, or other civic or social groups?

NEVER	SOMETIMES	FREQUENTLY

25. How often do you try new things or seek out new experiences?

NEVER	OCCASIONALLY	OFTEN

26. Given the opportunity, would you travel to a country with a different culture and a different language?

NEVER	MAYBE	ABSOLUTELY

27. Do you go on carnival rides or participate in other leisure activities (horse back riding, bungee jumping, snow skiing, water sports) for the excitement of thrills that they provide?

NEVER	OCCASIONALLY	OFTEN

28. How well do you concentrate on disagreeable tasks?

NOT AT ALL	MODERATELY WELL	VERY WELL

29. How often do you play games on computers?

NOT AT ALL	OCCASIONALLY	FREQUENTLY

30. How many different video, computer or arcade games have you become reasonably good at playing?

NONE	ONE	TWO	THREE	FOUR	FIVE	SIX OR MORE

~~31. Have you ever felt completely caught up in an experience, aware of everything going on and completely open to all of it?~~
~~NEVER OCCASIONALLY FREQUENTLY~~

~~32. Have you ever felt completely focused on something, so wrapped up in that one activity that nothing could distract you?~~
~~OT AT ALL OCCASIONALLY FREQUENTLY~~

33. How frequently do you get emotionally involved (angry, sad or happy) in news stories that you see, read, or hear?

NEVER OCCASIONALLY OFTEN

34. Are you easily distracted when involved in an activity or working on a task?

NEVER OCCASIONALLY OFTEN

35. Do you use a backpack?

NEVER OCCASIONALLY OFTEN

36. Do you play games requiring you to find things or other people?

NEVER OCCASIONALLY OFTEN

37. Do you get anxious in confined spaces?

NEVER OCCASIONALLY OFTEN

38. Does your job or recreational activity require you to wear equipment unique to the activity (e.g., scuba tanks for a diver or gun belt for a police officer?

NEVER OCCASIONALLY OFTEN

39. What is your decision making style?

QUICK SOMEWHAT DELIBERATE

40. Do you tend to rotate maps when reading them?

NEVER OCCASIONALLY OFTEN

APPENDIX D: ADAPTED PRESENCE QUESTIONNAIRE (PQ)

PRESENCE QUESTIONNAIRE
(Witmer & Singer, Vs. 3.0, Nov. 1994, Revised by B. Goldiez, April, 2004)

Characterize your experience in the environment, by marking an "X" in the appropriate box of the 7-point scale, in accordance with the question content and descriptive labels. Please consider the entire scale when making your responses, as the intermediate levels may apply. Answer the questions independently in the order that they appear. Do not skip questions or return to a previous question to change your answer.

WITH REGARD TO THE EXPERIENCED ENVIRONMENT

1. How much were you able to control events?

NOT AT ALL SOMEWHAT COMPLETELY

~~2. How responsive was the environment to actions that you initiated (or performed)?~~

~~NOT RESPONSIVE MODERATELY RESPONSIVE COMPLETELY RESPONSIVE~~

3. How natural did your interactions with the environment seem?

EXTREMELY ARTIFICIAL BORDERLINE COMPLETELY NATURAL

4. How much did the visual aspects of the environment involve you?

NOT AT ALL SOMEWHAT COMPLETELY

~~5. How much did the auditory aspects of the environment involve you?~~

~~NOT AT ALL SOEWHAT COMPLETELY~~

~~6. How natural was the mechanism which controlled movement through the environment?~~

~~EXTREMLY ARTIFICIAL BORDERLINE COMPLETELY NATURAL~~

~~7. How compelling was your sense of objects moving through space?~~
~~NOT AT ALL MODERATELY COMPELLING VERY~~

~~8. How much did your experiences in the virtual environment seem consistent with your real world experiences?~~

~~NOT CONSISTENT — MODERATELY CONSISTENT — VERY CONSISTENT~~

~~9. Were you able to anticipate what would happen next in response to the actions that you performed?~~

~~NOT AT ALL — SOEWHAT — COMPLETELY~~

10. How completely were you able to actively survey or search the environment using vision?

|	|	|	|	|	|	|	|

NOT AT ALL SOMEWHAT COMPLETELY

~~11. How well could you identify sounds?~~

~~NOT AT ALL — SOMEWHAT — COMPLETELY~~

~~12. How well could you localize sounds?~~

~~NOT AT ALL — SOMEWHAT — COMPLETELY~~

13. How well could you actively survey or search the environment using touch?

NOT AT ALL SOMEWHAT COMPLETELY

14. How compelling was your sense of moving around inside the environment?

NOT COMPELLING MODERATELY COMPELLING VERY COMPELLING

~~15. How closely were you able to examine objects?~~

~~NOT AT ALL PRTTY CLOSELY VERY CLOSELY~~

~~16. How well could you examine objects from multiple viewpoints?~~

~~NOT AT ALL SOMEWHAT EXTENSIVELY~~

17. How well could you move or manipulate objects in the environment?

NOT AT ALL SOMEWHAT EXTENSIVELY

18. How involved were you in the experience presented in the environment?

NOT INVOLVED MILDLY INVOLVED COMPELTELY ENGROSSED

19. How much delay did you experience between your actions and expected outcomes?

NO DELAYS MODERATE DELAYS LONG DELAYS

20. How quickly did you adjust to the environment experience?

NOT AT ALL SLOWLY LESS THAN ONE MINUTE

21. How proficient in moving and interacting with the test environment did you feel at the end of the experience?

NOT PROFICIENT REASONABLE PROFICIENT VERY PROFICIENT

22. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

NOT AT ALL INTERFERED SOMEWHAT PREVENTED TASK PERFORMANCE ENHANCED

23. How much did the control devices interfere with the performance of assigned tasks or with other activities?

NOT AT ALL	INTERFERED SOMEWHAT	INTERFERED GREATLY

24. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

NOT AT ALL	SOMEWHAT	COMPLETELY

~~25. How completely were your senses engaged in this experience?~~

~~NOT ENGAGED — MILDLY ENGAGED — COMPLETELY ENGAGED~~

26. To what extent did events occurring outside the environment distract from you experience in the virtual environment?

NOT AT ALL	MODERATELY	VERY MUCH

~~27. Overall, how much did you focus on using the display and control devices instead of the virtual experience and experimental tasks?~~

~~NOT AT ALL — SOEWHAT — VERY MUCH~~

28. Were you involved in the experimental task to the extent that you lost track of time?

NOT AT ALL	SOMEWHAT	COMPLETELY

~~29. How easy was it to identify objects through physical interaction: like touching an object, walking over a surface, or bumping into a wall or object?~~

~~IMPOSSIBLE — MODERATELY DIFFICULT — VERY EASY~~

30. Were there moments during the experience when you felt completely focused on the task or environment?

NONE	OCCASIONALLY	FREQUENTLY

31. How easily did you adjust to the control devices used to interact with the virtual environment?

DIFFICULT MODERATE EASILY

~~32. Was the information provided through different senses in the virtual environment (e.g., vision, hearing, touch) consistent?~~

~~NOT CONSISTENT SOMEWHAT CONSISTENT VERY CONSISTENT~~

(Additional questions)

33. During the experience was the equipment burdensome (e.g., heavy, bulky, restricted movement) such that it interfered with your performing the tasks?

NEVER OCCASIONALLY OFTEN

34. During the experience did you experience any anxiety from feeling lost?

NEVER OCCASIONALLY OFTEN

35. During the experience did you experience anxiety from the confined space?

NEVER OCCASIONALLY OFTEN

36. Was the augmentation supportive in accomplishing your tasks?

NEVER OCCASIONALLY OFTEN

37. Did augmentation interfere with your tasks?

NEVER OCCASIONALLY OFTEN

38. Were you involved in the experimental task to the extent that you lost track of being in a laboratory?

NEVER SOMEWHAT COMPLETELY

39. This task possibly required you to balance speed in going through the maze with the accuracy of your response to questions in the bag. Was this balance difficult?

NOT AT ALL SOMEWHAT COMPLETELY

40. Which task did you consider primary?

TIME NEITHER ACCURACY

NEVER OCCASIONALLY OFTEN

NEVER OCCASIONALLY OFTEN

NEVER OCCASIONALLY OFTEN

APPENDIX E: DEMOGRAPHICS QUESTIONNAIRE

Demographics Questionnaire

1. Please circle your gender: Female Male

2. Which is your predominate hand? Please circle: Right Left

3. What is your major (if applicable)? ____

4. What is your profession (if applicable)? ____

5. What is your age in years? ____

6. Do you have experience in virtual environments? _________________________________

If so, please explain: ____

____

7. Do you play video games? ____

If so, please estimate the number of hours per week you play these games: _____

What type of games are they (e.g., strategy, fighting, racing):_________________

---

8. Do you wear prescription glasses or corrective contact lenses? ___ Yes ___ No

If yes, are you wearing them now? ___ Yes ___ No

If yes, do you have ___ nearsightedness (myopia) or ___ farsightedness (hypermetropia)?

APPENDIX F: SAMPLE EXPERIMENTERS SCRIPT

Experimental Procedures (Treatment E_D)

(Read this to each participant)

The Experimenter: "Thank you for participating in this experiment. This experiment is part of an Augmented Reality research project sponsored by the Office of Naval Research. In general, augmented reality systems add information to the real world using a computer. You will be wearing a mobile augmented reality system that weighs about 10 pounds. It fits on a vest that you will wear. There is an eye piece that is attached to a pair of clear glasses. The glasses are designed to fit over most eyeglasses. The display's position is adjustable and we will help you in adjusting it to a comfortable position. The entire test should take less than one hour.

Your task involves going through a maze and is similar to what a police officer might do in a search and rescue operation. Before putting on the equipment and performing the task you will be asked to fill out an informed consent form, a demographics sheet and take three surveys. Two of the surveys are used to evaluate your spatial abilities. The third survey is a questionnaire that seeks to find out about your tendency to be immersed in a task. Do your best on the tests, go in sequence through the questions, and do not go back and redo questions unless you have finished before time is called. There is no grade and the test results and other data we collect will be kept strictly confidential. Before entering the maze, you will be blindfolded and turned around (similar to the children's game "Pin the Tail on the Donkey"). We do this to facilitate 'erasing' your spatial orientation memory.

After completing the task of traversing the maze you will be asked to fill out two more surveys: one oriented to your ability to recall spatial relationships in the maze and a second to evaluate how involved you were in the tasks.

If at any time you feel uncomfortable with the situation and want to stop the experiment, please verbalize your intent and someone will come and help you out. Otherwise, we will not respond to questions or comments during your traversal of the maze or any other tasks.

We would like you to do your best during this experiment and are asking that you provide some contact information (email, phone, address) because the best performer in each test condition will receive a gift certificate to a local restaurant (Bennigan's, Applebee's or similar). Do you have any questions so far?"

NOTE to the Experimenter:

Give the participant the Informed Consent Form (ICF) and the demographics sheet.

The Experimenter: "Please fill out the ICF and the demographics sheet".

NOTE to the Experimenter:

After the participant is done with filling out the ICF and the demographics sheet, administer Spatial Visualization Measure (make sure you use the correctly labeled scantron with the participant number!) Hand participant spatial orientation measure.

The Experimenter: "Please complete the following exercise. Read over the instruction page carefully and complete the practice session on pages 1, 2, and 3. Please do **NOT** record your responses to the **practice** items. When you have finished the practice session and are ready to begin, please let me know. Please do **NOT** begin working on the remainder of the exercise until instructed to do so".

NOTE to the Experimenter:

Make sure that you tell them *NOT* ***TO WRITE IN THE BOOKLET and record their answers ONLY on the scantron provided!!!!!!! After they let you know that they are done going over the sample items give them a scantron and ask them again to record their answers on the scantron ONLY! Hand them the scantron.***

The Experimenter: "Please record your answers on the scantron ONLY. DO NOT WRITE ON THE BOOKLET! You have 10 minutes to work on the test. Do not spend too much on one item. If you finished before the time is called, you may go back and check your work. If you are not sure about the answer to any item, you may guess, but avoid wild guessing. Your score will be the number of correct answers minus a fraction of the number wrong. WAIT FOR THE SIGNAL TO BEGIN".

At the end of 10 minutes: "Please stop working on the exercise."

NOTE to the Experimenter:

Ask him/her to give you the scantron and. Collect the scantron and test booklet.

Administer Spatial Orientation Measure (make sure you use the correctly labeled scantron!) Hand participant spatial orientation measure.

The experimenter: "Please complete the following exercise. Read over the instruction page carefully and complete the practice session on pages 1, 2, and 3. When you have finished the practice session and are ready to begin, please let me know. Please do **NOT** begin working on the remainder of the exercise until instructed to do so."

After the practice session: *Hand them the scantron.*

The experimenter: "You will now have 10 minutes to work on this exercise. Please remember, do **NOT** make any marks in the booklet and only record your responses on the scantron provided."

At the end of 10 minutes:

The experimenter: "Please stop working on the exercise."

NOTE to the Experimenter: Collect the scantron and test booklet*.*

The experimenter: "You will now be given a short test that assesses your ability to become immersed in a task. Please mark your answers directly on the form."

NOTE to the Experimenter: *Administer the Immersive Tendency Questionnaire and retrieve the test when completed.*

The experimenter: "Please listen carefully to the following instructions.

- Put on the mobile augmented reality system and get comfortable with it. In this test condition, the display will only be active when you depress the space bar on the keyboard attached to your forearm. The display will go off after 15 seconds and can be turned on by depressing the key again. The map you see will be fixed in a north up position. Remember, too, that we will blindfold you before taking you into the maze area and turn you around several times.

NOTE to experimenter: Let the subject put on the equipment in the testing room and get comfortable with the weight and use of the glasses. When comfortable, blindfold the participant. Escort them into the lab are and place them in the maze before removing the blindfold. Insert earplugs before beginning.

- Stand at the entrance to the maze, but do not begin until instructed.
- You will wear a visor.
- You will be timed on how long it takes you to traverse the maze, find the target object and exit the maze. The exit and entrance of the maze are the same.
- You should traverse the maze COMPLETELY before retrieving the object (this includes dead ends). Doing otherwise will result in a penalty. The target object is easy to recognize.
- As you are traversing the maze you are to stop at all stations where you see a bag with a clipboard in it. There are five (5) bags placed at different locations in the maze. You are to stop at **all of them in sequence** and answer all questions. That is, find and answer the questions to bag 1 before proceeding to find and answer the questions in bag 2, etc. Answering the questions in the bags out of sequence will result in a penalty.
- The questions in the bags address your workload and your current location in the maze. The bags are numbered from 1 to 5. Three answer sheets are provided on each clipboard. After answering **all** the questions, please return the clipboard along with

the three answer sheets to the bag. Not answering the questions in a bag will result in a penalty.

- When you retrieve the target object verbalize it by saying: “Found the object!” Please note that you will not receive a verbal confirmation from us.
- Take the object with you and exit the maze through the most expedient (shortest) route.
- Remember; try to completely traverse the maze before retrieving the object and exiting.
- Please be aware of some equipment limitations. If you notice any anomalies in the display’s operation, such as erratic deviations of the user’s symbol, please stop moving and within a couple of seconds the display will be back on track.

Any questions before we begin?

NOTE to the experimenter:

After the participant exits the maze, collect the target object and ask him/her to take off the wearable computer and the display piece.

The experimenter: Now, you will be asked fill out two more surveys: one oriented to your ability to recall spatial relationships in the maze and a second to evaluate how involved you were in the tasks. Please fill out the first survey (***Note to the experimenter: this is the maze question. After he/she is done with it, collect the answer sheet and give them the Presence questionnaire).***

The experimenter: Please fill out the following Presence questionnaire.

Note to the experimenter: After the participant is done collect the questionnaire.

The experimenter: The principle objective of this study is to determine what type of augmented reality display information best facilitates navigation performance in areas, which people are not familiar with, and how seamlessly the augmented reality system works for them. Two control conditions are used in this experiment: no augmentation with no map and no augmentation with a map available prior to entering the maze. Two types of display orientations: (egocentric and exocentric);and two display availability conditions(continuous and an on demand display) are evaluated in the augmentation condition. It is hypothesized that:

1. Performance using a real time, hands-free AR map display will improve compared to the control conditions;
2. Within the AR condition, the use of egocentric moving map will reduce the time and improve recall accuracy on secondary task, in terms of Euclidian distance and egocentric orientation, by providing real time egocentric view of the space
3. Within the AR conditions, the use of exocentric fixed North-up map will reduce the time and will improve recall accuracy on secondary task, in terms of cardinal directions, by providing real time North-up view of the space
4. Within the AR conditions, the use of an “on demand” map display as compared to a continuously on map display will reduce the time and improve accuracy on the primary task and the secondary embedded task by providing the participant with control over the times when the map is needed as a reference

Please feel free to ask any questions at this time about the procedure or the experiment in general. Should you desire to learn more about the study or receive the results of the experiment when they become available, please contact the principle investigator Brian Goldiez.

Thank you for your participation.

Experimenter Guidelines

1. Each participant is to have a file that will contain a unique identifier that will be coded to the experimental condition they are to be engaged in. The file will contain the signed informed consent, demographics, ITQ, PQ, pre and post hoc spatial tests, contact information (optional), and scoring sheets. Each sheet in the folder should contain the unique identifier for the participant.
2. Assist the participant with the equipment. They will wear and have to depress a key on the forearm keyboard of BARS, ask which arm they prefer. Also, the participant will need help positioning the display on the proper eye and with proper orientation (if you switch eyes the display is inverted and must be set to proper orientation). You should ask this question and make the orientation adjustment BEFORE the participant puts on the BARS. The display works best if positioned slightly below the horizontal position when the eye is at rest.

3. If someone is tested in the control condition (no augmentation), they are to still wear the BARS (with a blank display). After all testing, they should be afforded the opportunity to explore the maze with BARS (but no data will be collected) so that they can have the same fun as others.
4. When the participant is positioned at the entrance to the maze, the BARS display should NOT be on. When you say "BEGIN" you should start the BARS system, which might take a few seconds to begin (please convey this to the participant). This delay will introduce a bias in the results which we will measure and account for in our analysis.
5. As the participant traverses the maze, keep away from the edges to avoid the participant receiving information on your position. Also do not speak and ensure that the lab stays quiet (POST SIGNS).
6. Do not answer questions while the participant is doing their work. If a question is raised during one of the tests (e.g., I don't understand this question), tell the participant to respond in the best way they can.
7. After a participant completes their test collect their tests, forms, and sheets in the maze. Download the data from the AR computer to a file on the lab computer and label the file with the participant identifier.

- Begin appendix text on the page following the buffer page.
- Continue Arabic pagination; do not restart page numbering.
- Use the same style and format for buffer page headings as you do for other body chapter headings.
- Letter, don’t number, appendixes.
- If you have only one appendix, do not letter it at all.

Experimental Condition: mazeExoDemand (Exo-centric On Demand Map)

Operating BARS to run participants:

1. Turn on InterSense equipment: The power button is at the back at right hand side of the equipment. It will take around one minute to boot. The blue lights on the sensor track (at the ceiling) turn on for a while and then turn off.
2. Switch ON the wireless receiver for the tracker. The wireless receiver is located on yellow pole. After switching this on, the blue lights on sensor track will start blinking one by one.
3. Put the batteries into the pockets on the mobile augmented reality system. The battery pockets are on the left side of the system. Do NOT attach the battery clips yet.
4. Attach the base power adapter. This is the external power adapter near the battery pockets.
5. Turn ON the thermite (small computer). A thermite is on the back side of mobile virtual reality system. A green indicator should be visible.

On the BlackLab1 computer:

1. Take out the Wireless Ethernet from the top right hand side pocket of the mobile augmented reality system and plug it in any of the USB ports of BlackLab1 computer.
2. Log in using "maze" as username, "the.maz3" as password and "BlackLab1" as the domain.
3. Double click on "ISDemo" → Click on "Detect" → (After establishing connection) Click on "Accept" → Tools → UDP Broadcast Server → Broadcast.
4. Minimize the two windows and let them be there. Do NOT close them.
5. Double Click on "Thermite". Give the password as the.maz3. A window showing desktop of thermite will appear.

6. Before Positioning the participant:

    1. Attach the battery clip. Use one battery only.
    2. Detach the base power adapter.
    3. Disconnect the monitor's VGA cable and connect the micro display device to that socket of thermite.

4. Install the charged battery in the battery holder for the display device. Turn on the Display device by pressing Power/Mode button on the installed battery for the display device. The green LED labeled BRT will illuminate if valid signal is present. This LED will extinguish after 5 seconds to save power.
5. Make orientation adjustments of display.

7. Double click on "Cygwin". A dos window will appear with a prompt.
8. If you type "ls" you will see the different conditions. "mazeExoDemand" gives the experimental condition exo-centric On Demand Map.
9. Whenever the participant is ready, type following after the prompt:
   ./mazeExoDemand {participant number} Do not include curly brackets.

   (for example: <DOT SLASH>mazeExoDemand <SPACE> 3_X_D)

10. Whenever the participant is ready press enter after the above command. At the same time start the timer to record the time system takes from pressing "ENTER" button and displaying the map. Stop the timer as soon as you see the map on the screen.
11. Minimize the dos window (cygwin) to see the tracking. Ask participant to start as soon as you see the map on the screen.
12. Monitor the participant's progress through the maze. When he or she retrieves the target object, press "SPACE BAR" on the Thermite's wireless keyboard.
13. Press "Esc" when the participant exits from the maze. Maximize the cygwin window on the desktop screen of thermite. You should be able to see the prompt again.
14. After one participant is done with the experiment, switch OFF the wireless receiver, plug in the external power, take the clip off the battery, remove the display and turn the display OFF.
15. After several participants (number to be determined), place the used battery in the battery charger.
16. At the end of the day after running all participants,
    1. Type "mv *.dat Data" at the command prompt of cygwin window. This will move all the dat files for different participants in "Data" folder from root folder. You can close the cygwin window now.
    2. Open the Data folder from the desktop of BlackLab1 computer by using "Administrator" as user name and no password. Copy all the files into My Documents/experimentData folder.
    3. In the ViaCT window (the screen for thermite), click on Start→Shutdown. This will shut down the thermite.
    4. Switch OFF the thermite by pressing the button and holding it till the green light goes OFF.
    5. Stop the wireless Ethernet device on BlackLab1. Right click on  icon at the right side of taskbar on the bottom of the screen. Click on 'Safely Remove D-Link Air DWL-122'. After few moments a note will appear saying "Safe to remove hardware". Remove the wireless Ethernet device from the USB port and keep it in the top right pocket of mobile augmented reality system.

6. Quit from the two minimized windows on BlackLab1 computer, one is ISDemo-Configured for IS-900 Series Device and other is UDP Broadcast Server.
7. Log off the BlackLab1 computer.
8. Detach the power adapter from Thermite.
9. Plug the charger into the InterSense wireless receiver. Make sure other end point is plugged into the socket.

Switch off the InterSense

LIST OF REFERENCES

American Psychological Association (2002). Ethical principles of psychologists and code of conduct. Retrieved September 21, 2004, from http://www.apa.org/ethics/code2002.html

Aretz, A. J. (1991). The design of electronic map displays. *Human Factors, 33* (1), 85-101.

Aretz, A. J., & Wickens, C. D. (1992). The mental rotation of map displays. *Human Performance, 5* (4), 303-328.

Arthur, E. J. (1996). *Orientation specificity in the mental representation of three dimensional environments*. Unpublished Doctoral Dissertation, University of Minnesota, MN.

Arthur, E. J., & Hancock, P. A. (2001). Navigation training in virtual environments. *International Journal of Cognitive Ergonomics, 5* (4), 387-400.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments, 6* (4), 355-385.

Azuma, R. T. (2001). Augmented reality: Approaches and technical challenges. In W. Barfield & T. Caudell (Eds.), *Fundamentals of wearable computers & augmented reality* (pp. 27-63). Mahwah, NJ: Lawrence Erlbaum and Associates.

Azuma, R.T., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics & Applications*, *IEEE, 21* (6), 34-47.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D. (1990). *Human memory: Theory & practice*. Hove, UK: Lawrence Erlbaum Associates.

Baddeley, A. D. (1992). Working memory. *Science, 255*, 566-569.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A*, 5-28.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory, *Trends in Cognitive Science, 4*, 417-423.

Baddeley, A. D. (2001). Is working memory still working?, *American Psychologist, 56*, 849-864.

Baddeley, A. D., Eldridge, M., & Lewis, V. (1981). The role if subvocalization in reading. *Quarterly Journal of Experimental Psychology, 33A*, 439-454.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning & motivation, Vol. 8* (pp. 47-90). London, UK: Academic Press.

Baddeley, A. D. & Wilson, B. A. (2002). Prose recall and amnesia: Implications for the structure of working memory. *Neuropsychologia, 40*, 1737-1743.

Baird, K. M. (1999). *Evaluating the effectiveness of augmented reality and wearable computing for a manufacturing assembly task*. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Barfield, W., & Caudell, T. (2001). Basic concepts in wearable computers and augmented reality. In W. Barfield and T. Caudell (Eds.), *Fundamentals of wearable computers and augmented reality* (pp. 3-26). Mahwah, NJ: Lawrence Erlbaum and Associates.

Barlow, T. (2001). Acquisition of route knowledge through navigation in a virtual environment. *International Journal of Cognitive Ergonomics, 5*(3),279-295.

Campbell, C. S., & Maglio, P. P. (1999). Facilitating navigation in information spaces: Road-signs on the World Wide Web. *International Journal of Human-Computer Studies, 50*, 309-327.

Carswell, C. M., & Wickens, C. D. (1987). Information integration and the object display: An interaction of task demands and display superiority. *Ergonomics, 30*(3), 511-527.

Christensen, C. M. (2003). *The innovator's dilemma*. New York, NY: HarperBusiness.

Clark, A. (2003). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. New York, NY: Oxford University Press.

Colle H. A., & Reid G. B. (1998). The room effect: Metric spatial knowledge of local and separated regions. *Presence: Teleoperators & Virtual Environments, 7* (2), 116-129.

Collette, F., Salmon, E., van der Linden, M., Chicherio, C., Belleville, S., Degueldre, C., Delfiore, G., & Franck, G. (1999). Regional brain activity during tasks devoted to the central executive of working memory. *Cognitive Brain Research 7*, 411–417.

Consulting Psychological Press. (1976). Palo Alto, CA: Consulting Psychological Press.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*, 1957, 671-684.

Darken, R. P., & Peterson, B. (2002). Spatial orientation, wayfinding, and representation. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation and application* (pp. 493-518). Mahwah, NJ: Lawrence Erlbaum Associates.

Durlach, N. I., & Mavor, A. S. (Eds). (1995). *Virtual reality: Scientific and technological challenges*. Washington, DC: National Academy Press.

Eberts, R. E., & Eberts, C. G. (1989). Four approaches to human computer interaction, in P. A. Hancock and M. H. Chigell (Eds.), *Intelligent interfaces: Theory, research and design* (pp. 69-127). North-Holland, Netherlands: Elsevier Science Publishers.

Edgar, G. K., & Reeves, C. A. (1997). Visual accommodation and virtual images: Do attentional factors mediate the interacting effects of perceived distance, mental workload, and stimulus presentation modality?. *Human Factors, 39* (3), 374-381.

Ellis, S. R., & Menges, B. M. (1998). Localization of virtual objects in the near visual field. *Human Factors, 40* (3), 415-431.

Evans, G. W., & Pezdek, K. (1980). Cognitive mapping: Knowledge of real-world distance and location information. *Journal of Experimental Psychology: Human Learning and Memory, 6* (1), 13-24.

Feiner, S., MacIntyre, B., Höllerer, T., & Webster, A. (1997). A touring machine: Prototyping 3D mobile augmented reality systems for exploring urban terrain. *Proceedings of the International Symposium on Wearable Computing, Cambridge, MA.* 74-81.

Furmanski, C., Azuma, R., & Daily, M. (2002). Augmented-realty visualizations guided by cognition: Perceptual heuristics combining visible and obscured information. *Proceedings of IEEE and ACM International Symposium on Mixed and Augmented Reality, Darmstadt, Germany,* 215-224.

Gillner, S., & Mallot, H. (1998). Navigation and acquisition of spatial knowledge in a virtual maze. *Journal of Cognitive Neuroscience, 10* (4), 445-463.

Goldiez, B. F., & Dawson, J. W. (2004). Is presence present in augmented reality systems? *Proceedings of the Seventh Annual International Workshop on Presence 2004, Valencia, Spain,* 294-297.

Haberlandt, K. (1997). *Cognitive psychology*. Boston, MA: Allyn and Bacon.

Hancock, P. A. (1997). *Essays on the future of human-machine systems*. Eden Prairie, MN. Banta Information Services Group.

Hancock, P. A. (2004). The future of simulation. In D. Vincenzi & J. Wise (Eds.), *Human factors in simulation and training*. Mahwah, NJ: Erlbaum and Associates.

Harwood, K., & Wickens, C. D. (1991). Frames of reference for helicopter electronic maps: The relevance of spatial cognition and componential analysis. *International Journal of Aviation Psychology*. *I* (1), 5-23.

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators & Virtual Environments, 1*, 262-271.

Hooper E., & Coury B. G. (1994). Graphical displays for orientation information. *Human Factors*, 36 (1), 62-78.

Ijsselsteijn W., Ridder H., Freeman J., & Avons S. E. (2000). Presence: Concept, determinants and measurement. Retrieved on September 21, 2004 from http:// www.presence-research.org/papers/SPIE_HVEI_2000.pdf

International Society for Presence Research (ISPR) website. Retrieved on September 21, 2004 from http://lombardresearch.temple.edu/ispr.

Jansen-Osmann, P. (2002). Using desktop virtual environments to investigate the role of landmarks. *Computers in Human Behavior, 18*, 427-436.

Karasek, R. A. (1979). Job demands, job decision latititude, and mental strain: Implications for job redesign. *Administrative Science Quarterly,* 24, 2, 285-308.

Kiss, I., Pisio, C., Francois, A., & Schopflocher, D. (1998). Central executive function in working memory: event-related potential studies. *Cognitive Brain Research, 6*, 235-247

Klinker, G. J., & Ahlers, K. H. (1997). Confluence of computer vision and interactive graphics for augmented reality. *Presence: Teleoperators & Virtual Environments, 6* (4), 433-451.

Lawton, C. A., & Morrin, K. A. (1999). Gender differences in pointing accuracy in computer-simulated 3D mazes. *Sex Roles, 40* (1-2). 73-92.

Martin-Emerson, R., & Wickens, C. D.(1997). Superimposition, symbology, visual attention, and the head-up display. *Human Factors, 39* (4), 581-601.

McGee, M. K. (1999). *Integral Perception in Augmented Reality*. Unpublished Doctoral Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Meehan, M., Insko, B., Whitton, M., & Brooks Jr., F. (2002). Physiological measures of presence in stressful virtual environments. In *ACM Siggraph* (pp. 645-652). San Antonio, TX.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. *Proceeding of Society for Photographic and Information Engineering, Telemanipulator and Telepresence Technologies, Vol. 2351*, unknown location, 282-292.

Moore, D. (1994). *The basic practice of statistics*. New York, NY: W. H. Freeman and Co.

Navon, D., & Gopher, D. (1979). On the economy of the human information processing system. *Psychological Review, 86*, 214-255.

Neumann, U. & Majoros, A. (1998). Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. *Proceedings of IEEE Virtual Reality Annual International Symposium '98.* Atlanta, GA, 4-11.

Psotka, J. (1995). Immersive training systems: Virtual reality and education & training. *Instructional Science, 25*, 405-431.

Rekimoto, J. (1997). Navicam: A magnifying glass approach to augmented reality. *Presence: Teleoperators & Virtual Environments, 6* (4), 399-412.

Rolland, J. P., & Fuchs, H. (2001). Optical versus video see-through head-mounted displays. In W. Barfield and T. Caudell (Eds.)*, Fundamentals of wearable computers and augmented reality* (pp. 113-156). Mahwah, NJ: Lawrence Erlbaum Associates.

Sadowski, W., & Stanney, K. (1999). Measuring and managing presence in virtual environments. University of Central Florida. Retrieved on September 21, 2004 from http://vehand.engr.ucf.edu/handbook/Chapters/Chapter45.html.

Sheridan, T. (1992). Defining our terms. *Presence: Teleoperators & Virtual Environments, 1* (24), 272-274.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. W. Reese (Ed.)*, Advances in child development and behavior* (pp. 9-55). New York, NY: Academic Press.

Singer, M. J., & Witmer B. G. (1996). *Presence measures for virtual environments: Background & development*. (United States Army Research Institute for the Behavioral and Social Sciences). Alexandria, VA: Army Research Institute.

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory & Cognition, 22*, 1-13.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments: Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual environments, 6*, 603-616.

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory & Cognition, 22*, 1-13.

Stanney, K. M., & Salvendy, G. (1995). Information visualization: assisting low spatial individuals with information access tasks through the use of visual mediators. *Ergonomics, 38*, 1184-1198.

Tang, A., Owen, C., Biocca, F., Mou, W. (April 5-10, 2003). Comparative effectiveness of augmented reality in object assembly. *Proceedings of Association for Computing Machinery Computer Human Interaction.* Ft. Lauderdale, FL. 73-80.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*, 233-282.

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology, 14*, 560-589.

Waller, D., Hunt, E., & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators & Virtual Environments, 7* (2), 129-136.

Watson, N. V., & Kimura, D. (1991). Nontrivial sex differences in throwing and intercepting: Relation to psychometrically-defined spatial function. *Individual Differences, 12* (5), 375-385.

Werner, S., & Hable, C. (1999). Spatial reference systems. *Spatial Cognition and Computation, 1* (4), 3-7.

Werner, S., & Schmidt, K. (1999). Environmental reference systems for large-scale spaces. *Spatial Cognition and Computation, 1*, 447-473.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.)*, Attention and performance VIII , 8* (pp. 239-257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman and D. R. Davies (Eds.*), Varieties of attention* (pp. 63-102). New York, NY: Academic Press.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York, NY: Harper Collins.

Wickens, C. D. (1999). Frames of reference for navigation. In D. Gopher and A. Koriat (Eds.), *Attention and Performance XVII* (pp. 114-144). Cambridge, MC: The MIT Press.

Wickens, C. D., Gordon, S. E., & Liu, Y. (1997). *An Introduction to Human Factors Engineering*. Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., Liang, C., Prevett, T., & Olmos, O. (1996). Electronic maps for terminal area navigation: Effects of frame of reference and dimensionality. *The International Journal of Aviation Psychology, 6* (3), 241-271.

Witmer, B. G., & Sadowski, W. J. (1998). Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Human Factors, 40* (3), 478-488.

Witmer, B. G., Sadowski, W. J., & Finkelstein, N. M. (2000). *Training dismounted soldiers in virtual environments: Enhancing configuration learning* (United States Army Research Institute for Behavioral and Social Sciences Technical Report No. 1103). Alexandria, VA: Army Research Institute.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators & Virtual Environments, 7*, 225-240.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21* (3), 229-244.

Youngblutt, C., & Perrin, B. M. (2002). *Investigating the relationship between presence and performance in virtual environments*. Proceedings of the IMAGE 2002 Conference. Scottsdale, AZ.