An Empirical Study of the Effects of Context-Switch, Object Distance, and Focus Depth on Human Performance in Augmented Reality

Divya Gupta

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Industrial and Systems Engineering

Dr. Deborah Hix, Co-Chair

Dr. Robert J. Beaton, Co-Chair

Dr. Robert S. Schulman, Committee Member

Dr. Tonya L. Smith-Jackson, Committee Member

June 2004
Blacksburg, Virginia

Keywords: Augmented Reality, Context-Switch, Focus Depth, Focal Length
Copyright 2004, Divya Gupta

An Empirical Study of the Effects of Context-Switch, Object Distance, and Focus Depth on Human Performance in Augmented Reality

Divya Gupta

(ABSTRACT)

Augmented reality provides its user with additional information not available through the natural real-world environment. This additional information displayed to the user potentially poses a risk of perceptual and cognitive load and vision-based difficulties. The presence of real-world objects together with virtual augmenting information requires the user to repeatedly switch eye focus between the two in order to extract information from both environments. Switching eye focus may result in additional time on user tasks and lower task accuracy. Thus, one of the goals of this research was to understand the impact of switching eye focus between real-world and virtual information on user task performance.

Secondly, focus depth, which is an important parameter and a depth cue, may also affect the user's view of the augmented world. If focus depth is not adjusted properly, it may result in vision-based difficulties and reduce speed, accuracy, and comfort while using an augmented reality display. Thus, the second goal of this thesis was to study the effect of focus depth on task performance in augmented reality systems.

In augmented reality environments, real-world and virtual information are found at different distances from the user. To focus at different depths, the user's eye needs to accommodate and converge, which may strain the eye and degrade performance on tasks. However, no research in augmented reality has explored this issue. Hence, the third goal of this thesis was to determine if distance of virtual information from the user impacts task performance.

To accomplish these goals, a 3x3x3 within subjects design was used. The experimental task for the study required the user to repeatedly switch eye focus between the virtual text and real-world text. A monocular see-through head- mounted display was used for this research.

Results of this study revealed that switching between real-world and virtual information in augmented reality is extremely difficult when information is displayed at optical infinity. Virtual information displayed at optical infinity may be unsuitable for tasks of the nature used in this research. There was no impact of focus depth on user task performance and hence it is preliminarily recommended that manufacturers of head-mounted displays may only need to make fixed focus depth displays; this clearly merits additional intensive research. Further, user task performance was better when focus depth, virtual information, and real-world information were all at the same distance from the user as compared to conditions when they were mismatched. Based on this result we recommend presenting virtual information at the same distance as real-world information of interest.

ACKNOWLEDGEMENTS

First, I would like to extend my sincere gratitude to Dr. Deborah Hix, my committee co-chair, for being such a great advisor. She guided, encouraged, and helped me at every step in my thesis and never let me feel alone. Her enthusiasm made this a wonderful experience and she was truly inspirational and instrumental in my thesis work. I am also very grateful to Joseph Gabbard. He helped me develop an understanding of the field and was always there to provide support in everyway possible.

A very special thanks to John Lucas, for programming the software that was used for this thesis, helping me run participants, and being flexible in accommodating my research work with his student responsibilities. I would also like thank Dr. Edward Swan. His participation in our meetings and sharing with us his ideas and insight was very helpful in developing and refining research ideas. He also created the Perl script that was used for running the software.

Thanks to my committee co-chair, Dr. Beaton and committee members, Drs. Smith-Jackson and Schulman for providing valuable feedback that helped improve the quality of my research. Dr. Schulman's assistance in conducting the statistical analyses was of immense help. I consider myself very fortunate to have worked with all these outstanding researchers.

This thesis would not have been possible without the support of the Office of Naval Research. Dr. Larry Rosenblum, Program Manager, has generously sponsored augmented reality research at Virginia Tech for many years. I would also like to thank Concurrent Technologies, Mr. John Garoutte, and Mr. Upul Obeysekare for their support in conducting this research.

My friends in Blacksburg helped me get through this entire process. Nikhil had an open ear for my problems and ideas and always encouraged me. I am thankful to have had him around.

Last in this list, but first in ultimate importance, I would like to thank my parents for instilling in me the joy of learning and providing me the opportunity to succeed and my sister for being my friend and guide. I would have been nothing without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Augmented reality, which superimposes virtual graphics or text onto an actual real-world view, intends to enhance the user's understanding of the world. By supplying the user with virtual information about the environment, augmented reality aims to assist in user task performance. However, the virtual information, coupled with a complex physical world, creates the potential for excessive perceptual and cognitive load on the user. The user needs to perceive, read, and assimilate real-world and virtual information delivered through augmented reality. This requires a continuous shift in eye focus and mental attention between the real-world and virtual scenes. As a result, an augmented environment is unnatural compared to a real-world environment, as in the real world all information is exclusively in the real-world context and does not involve any switching of eye focus and attention between real-world information and virtual information.

Focus depth is the distance from the user at which virtual information appears to be displayed, and is "in focus". It is one of several depth cues used in augmented reality. Currently some augmented reality displays have user-adjustable focus depths while others have fixed-focus depths. In fixed-focus depth displays, the user's eye always focuses at a fixed distance specified by the display hardware, whereas in displays with an adjustable focus depth, the distance at which the user's eye focuses can be varied by the user. Although focus depth is an important parameter as it controls the user's view of the augmented world, it has received little attention in augmented reality systems.

In augmented reality, real-world objects and the virtual information are usually at different distances from the user. The user's eye needs to rapidly adjust to these different distances in order to perceive information in both contexts. A change in the eye's target distance causes the user to experience the phenomenon of eye accommodation and convergence (Charman, 2000) which may result in eye fatigue (Miller, Pigion, Wesner, & Patterson, 1983), incorrect distance estimation (Roscoe, 1984), missed targets (Edgar, Pope, & Craig, 1994), and can be slow and stressful over long periods of time. Researchers have not studied the impact of eye accommodation and convergence in augmented reality systems.

Much of the development in augmented reality has been technology driven (Stedmon, Hill, Kalawsky, & Cook, 1999). There has been little research studying the human as a component of augmented reality systems. While technology is still developing, it is important to address human factors issues at the earliest opportunity in the design process. The larger goal of this thesis was to examine specific human capabilities and limitations within augmented reality system and their impact on user task performance.

1.1 Problem statement

An objective behind all augmented reality systems is to provide a better understanding of the real world through additional information supplied via virtual graphics. The better understanding provided by augmented reality is then expected to aid in task performance (McGee, 1999). This thesis attempted to study three aspects of the impact of augmented environments on human information processing and task performance. These three issues are as follows.

First, as mentioned previously, augmented reality consists of the integration of two components: real-world objects/scene and virtual information (graphics and/or text) that augments the real-world objects/scene. The role of augmenting information is to assist the user without deteriorating the user's perception of the real world that contains primary task-related information. Several augmented reality applications, such as in manufacturing and repair, navigation, and annotation, require the user to continuously switch eye focus and mental attention between the real-world scene and virtual graphics. Context-switching for this study is defined as the switch in visual and mental attention between real-world and virtual information. The need to context-switch between real-world information and virtual-world information may deteriorate user performance by requiring the user to spend additional time and effort on visual tasks. However, empirical studies to determine if there is any effect of context-switching are lacking in augmented reality research.

Secondly, focus depth influences the user's view of the augmented environment. Changing the focus depth causes the virtual graphics to appear closer or farther away. As mentioned previously, some displays have an adjustable focus depth while in others it is fixed. The need for displays with adjustable focus depth for augmented reality systems has not been justified through empirical studies and we do not know if focus depth impacts user task performance. Moreover there is a complete lack of practical guidelines for augmented reality display manufacturers on the issue of focus depth.

Lastly, all augmented reality applications, including medical visualization, annotation, military operations, manufacturing and repair, and navigation, present the user with an augmented view in which real-world objects and virtual graphics are situated at several different distances from the user. In order to extract information from the augmented environment, real-world tasks require the user to shift eye focus back and forth to different distances. Frequent shifting of gaze to different depths may result in excessive strain on the accommodation and convergence mechanism of the eye (Neveu, Blackmon, & Stark, 1998). Currently augmented reality display manufacturers recommend displaying virtual information at the same depth as the real-world object of interest so that there is minimal eye accommodation. However, this recommendation has not been verified empirically. Moreover, in complex environments where there are several real-world objects situated at different depths, and several pieces of virtual information need to be displayed, the question of what distance is optimal for displaying virtual information remains unanswered. Thus there is a lack of research on the issue of optimal virtual graphics distance for maximizing performance on user tasks and minimizing visual fatigue.

1.2 Objective of proposed research

As mentioned in the previous section, this thesis investigated the impact of augmented environments on information processing capabilities of the user. The three goals of this research were to answer the following questions, by providing evidence through an empirical study:

1. What is the effect of changing eye focus between real-world objects and virtual objects on user task performance in augmented reality systems?
2. To what extent does focus depth impact user task performance in augmented reality?
3. To what extent do different real-world object distances, virtual object distances, and focus depths impact user task performance in augmented reality? In other words, is there any difference in user task performance when distance to the real-world object, distance to the virtual object/text, and focus depth are all at the same distance from the user as compared to when they are at different distances?

There is a dearth of literature to provide human-centered recommendations for the design of both hardware and software user interfaces for augmented reality systems. Most of the work in augmented reality addresses only technical aspects and fails to incorporate the user into the system. This thesis provided some initial empirical findings to support the development of an improved user-centered augmented reality system.

1.3 Approach to proposed research

The experimental design for this research consisted of a 3x3x3 within subjects design. Distance to real-world text, distance to virtual text, and focus depth were the three independent variables. Task completion, which was defined as the number of subtasks completed per task, and accuracy, defined as the number of correct responses per task, were the dependent variables. In order to empirically study the three research questions discussed in section 1.2, we devised an experimental task, which forced participants to switch eye focus and thus context-switch between real-world and virtual-world information. A monocular see-through head mounted display with adjustable focus depth was used to display virtual information. Participants performed a visual task by counting the number of times a virtual-world letter appeared on a real-world monitor, embedded in random character strings.

Two control conditions were used in the experiment. For the first control condition, the experimental task was performed in an exclusively real environment using two monitors, to serve as a baseline measure of performance. Comparison of the first control condition where all information was in the real world with the treatment condition where information was presented through the augmented environment helped to determine the cost of switching between real world and virtual world. In the second control condition, the three independent variables were at the same distance from the user. A comparison of the second control condition with the treatment condition helped to determine if user performance was better when the independent variables (focus depth, distance to virtual text, and distance to real-world text) were all at the same distance from the user. Based on the empirical study, we drew conclusions regarding the impact of focus depth and virtual object distance for augmented reality systems.

2. LITERATURE REVIEW

2.1 Augmented reality

Augmented reality is often regarded as a variation of virtual reality (Azuma, 1997). While virtual reality completely immerses the user in a synthetic environment and the user cannot see the real world, augmented reality allows the user to see the real world with virtual objects superimposed upon the real-world view. Thus augmented reality is used to refer to a combination of real-world objects and virtual objects. As such each term in the phrase "augmented reality" provides insight into its meaning (McGee, 1999). Augmented reality has been referred to by Feiner, MacIntyre, and Seligmann (1993) as knowledge enhancement of the world, as it truly augments the users' real-world view with information that cannot be directly detected with their own senses.

To understand the characteristics of augmented reality, it is important to consider its relationship with virtual reality and reality. It is convenient to view reality and virtual reality at two opposite ends of a continuum which is referred to as the reality-virtuality continuum (Milgram, Takemura, Utsumi, & Kishino, 1994). Augmented reality is described as one of many mixed reality systems. Mixed reality systems are those that have both real and virtual objects of varying extent presented together. A system with more real-world objects would be categorized as augmented reality whereas a system with more virtual objects would be labeled augmented virtuality. At the two ends of the continuum are exclusively real and exclusively virtual environments. The reality-virtuality continuum is shown in Figure 1.

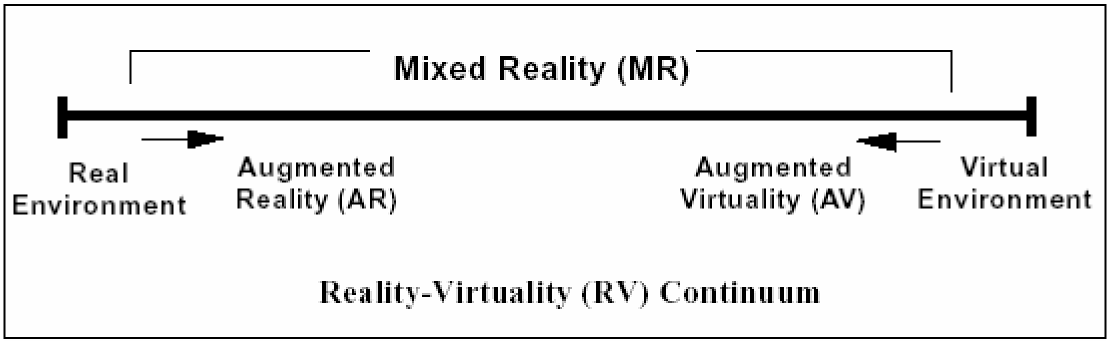

Figure 1. The reality-virtuality continuum

Note. From "Augmented reality: A class of displays on the reality virtuality continuum", by P.H. Milgram, H. Takemura, A. Utusumi, and F. Kishino, 1994, *SPIE: Telemanipulator and Telepresence Technologies*, *2351*, p. 283. Copyright 1994 by SPIE. Reprinted with permission of copyright holder.

2.2 Augmented reality vs. virtual reality

The reality-virtuality continuum of Milgram et al. reinforces that although augmented reality and virtual reality belong to a common class of mixed reality systems, they have fundamental differences which present distinct technical and human factors challenges. Some of the key differences between the two systems are discussed below.

One of the major challenges in augmented reality is accurately aligning the virtual world and real-world images. Many applications, such as medicine, require a very accurate registration of real and virtual objects. This is not such a problem in virtual environments as the user only sees virtual objects. Another major hurdle in augmented reality is the requirement for long range sensors and trackers that accurately report the location of the user and surrounding objects in the environment (Azuma, 1997). These sensors are, however, sufficient for virtual environments (Satoh, Anabuki, & Yamamoto, 2001). In augmented reality, virtual objects are superimposed on a complex real-world view, hence the display often becomes cluttered and unreadable, imposing significant cognitive and perceptual load on the user. There is also the issue of virtual objects occluding real-world objects and the problem of how to convey depth information about occluded objects. For outdoor augmented reality a major constraint is the limited brightness of the display, which makes the display unreadable in bright sunlight. Weight and size of the display, power requirements, limited field of view, and tethering continue to restrain outdoor use of augmented reality (Azuma, 2001). Some of these challenges are less severe in virtual reality, as the user is completely immersed in a synthetic environment.

Virtual environments, on the other hand, have been battling with the problem of vertigo and simulator sickness. This is not so much an issue in augmented reality as orientation cues are available to the user (Stedmon, 2001). Also virtual environments have a much higher requirement for realistic images as compared to augmented reality. In augmented reality, virtual objects merely supplement the real world, which does not necessarily have to be drawn realistically.

Virtual reality has been the subject of intensive research over the past few decades due to its applications in training systems, entertainment, medicine, and visualization. Although augmented reality came into existence in the 1960s, it is only in the past decade that it has received attention from the research community. As such, augmented reality can be regarded as behind its counterpart, virtual reality, in maturity.

2.3 Augmented reality systems

The ability of augmented reality to supply the user with additional information has found applications in medical visualization, maintenance and repair, annotation, entertainment, and military navigation and targeting (Azuma, 1997). To be effective, most augmented reality applications require the system to be portable and not restrict user movement. This has been made possible by wearable computers. Wearable computers are defined as portable, tetherless, light-weight computers which allow hands-free-usage and are literally worn by the user. They are usable at any time with minimal distraction and allow the user to roam the real world without being restricted to stationary machines (Feiner, MacIntyre, Hollerer, & Webster, 1997).

A mobile augmented reality system uses a wearable computer along with a tracking system and a head-mounted display. The tracking system which is mounted onto the display unit tracks the position and orientation of the user's head in the environment (Livingston et al., 2002). The wearable computer provides power and rendering of the 3D graphics and the head-mounted display allows the user to see the augmented view. Figure 2 shows a user wearing a head-mounted display and tracker with a wearable computer on his back.

Figure 2. Wearable computer

Note. From “A cost-effective usability evaluation progression for novel interactive systems,” by D. Hix, J.L. Gabbard, J.E. Swan II, M. Livingston, T. Hollerer, S.J. Julier, et al., 2004, *Hawaii International Conference on Systems Sciences*. Copyright 2004 by Deborah Hix. Reprinted with permission of the author.

In general, augmented reality displays can be divided into two main classes: optical-see through and video see-through systems. Optical see-through displays work by placing combiners in front of the user’s eyes. These combiners allow the user to see the view of the world directly. The display is generally used in a head-mounted device, although other devices such as hand-held displays have also been used (Rekimoto, 1997).

In video see-through systems, the real world is recorded with one or two cameras mounted on the helmet. The video cameras provide the user’s view of the real world. Video from these cameras is combined with graphic images created by the scene generator synthesizing the real and virtual images on a monitor before the user’s eyes (Azuma, 1997; Pasman, 2001; Sugihara & Miyasoto, 1999). The choice of display type is dependent on the augmented reality application for which it is intended. Despite certain limitations, optical see-through systems are preferred due to cost and safety issues (Azuma, 1997).

Head-mounted displays can be configured so that the outside real-world environment is always visible. Such displays are called see-through displays. On the other hand, opaque head-mounted displays are like using a computer. The user is unable to see objects in the outside real world through the display, just like a monitor obscures objects placed directly behind it (Yeh, Wickens, & Seagull, 1998). Further, head-mounted displays can also be classified as monocular, biocular, or binocular. In monocular displays, image is presented to only one eye with the other eye having an unaided view of the real-world environment. Biocular displays provide the same view to both the eyes whereas binocular displays provide a slightly different view to each eye allowing the user to perceive depth based on stereopsis (Yeh et al., 1998).

Each of the display types: monocular, biocular, and binocular have their advantages and disadvantages. In terms of complexity, monocular display is the simplest as it requires only one image source and one set of optics and is generally lighter than either binocular or biocular displays. It has a wider field of view of the far domain due to one uncovered eye which is hypothesized to allow for better target detection performance in the periphery (CuQulock-Knopp, Sipes, Torgerson, Bender, & Merritt, 1996) and for greater safety when operating under low illuminations (Kooi, 1993; Lippert, 1990).

Disadvantages of the monocular display include lack of depth information, the small amount of space for information display, and the potential for binocular rivalry (National Research Council, 1997). Binocular rivalry is defined as the failure to fuse two dissimilar images. In monocular viewing, one eye views the real-world image and the other eye views the virtual image. When these two images are sufficiently different they may cause the visual system to suppress the image from one eye. Over time the dominant image may shift from eye to eye so that two monocular views appear as alternating images. In general, the dominant image will be the one with greater intensity, contour, contrast, and motion (Yeh et al., 1998).

Both biocular and binocular transparent head-mounted display configurations are believed to give the user greater visual comfort, improve detection, and recognition of obstacles, and require less training than the monocular display (Blake & Fox, 1981; Lippert, 1990). The biocular display is more complex than the monocular display as it requires a second set of optics, thus making the display slightly heavier than the monocular display, but eliminates the problem of binocular rivalry through two-eyed presentation of data.

The binocular head-mounted display is the most complex, requiring two sets of optics and two image sources. This display is the only one which allows for stereoscopic viewing and three-dimensional depth perception (Davis, 1997), and as a result is the heaviest and most difficult to adjust to the viewer (National Research Council, 1997). The motivation for stereo viewing includes improved spatial perception with better visual filtering of noise, enhanced image quality, better object recognition, less training time, and greater user satisfaction (Davis & Hodges, 1995; Drasic, 1991). Davis (1997) lists scenarios in which binocular viewing is more effective than biocular viewing.

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

An extensive literature review conducted by Yeh, Wickens, and Seagull (1998) on the three different types of display configurations: monocular, biocular, and binocular, revealed no clear advantage for any one type of display. Benefits of the addition of stereo for head-mounted displays were present in tasks which required maneuvering along a path. Comparisons of depth perception and monitoring tasks showed no advantage to the addition of stereo but a benefit for presentation to two eyes over only one eye. In target detection, there was no difference between the monocular, biocular, and binocular displays in dark illumination, but an advantage for monocular and binocular configurations over the biocular configuration in lit environments. There was no advantage for any of the three displays on wayfinding tasks. Yeh and colleagues found that the one vs. two eyed viewing condition caused only muted effects on performance in favor of the biocular condition. The literature review suggested that each configuration seems best suited for a specific situation, though no one display is optimal for all situations.

2.4 Augmented reality and the user

Technology has developed at an exponential rate over the last few centuries; however human perceptual and cognitive capacities have changed very little. Although the goal of technology is to assist humans in their work, technological advances also increase the volume and flux of available information, thus straining limited human perceptual and cognitive capabilities. Augmented reality is an example of technology that provides the user with additional information not available through the natural environment. Much of the development in augmented reality has been based on technology, and little research has looked at its impact on the user as part of the system. There are only a few efforts, presented below, that have begun to examine human capabilities and limitations in augmented environments.

A series of experiments was conducted to determine the effect of augmented reality on cognitive capabilities of the user. Stedmon and Stone (2001) summarized five experiments that evaluated comprehension and retention, perceptual interference, stimulus detection, response to alerts, and influence of clutter when additional information was delivered through augmented reality. Overall these experiments demonstrated that augmented reality provides information as effectively as a computer monitor. As a result of this research, it was also reinforced that information delivered through augmented reality should not place competing demands for attention on the user but should be complementary and contextually bound by the primary task.

Usability evaluation of ‘ARQuake’, an augmented reality based outdoor version of a desktop game was conducted by Thomas, Krul, Close, and Piekarski (2002). The user moved through the outdoor space and was able to view game features consisting of monsters, weapons, and other objects of interest in the physical world through the head-mounted display. The informal evaluation consisted of feedback from users who used the system for an hour. The evaluation addressed issues related to field of view, tracking, sense of presence, choice of screen colors, and appearance of the augmented scene. Input from users revealed the field of view of 24 degrees too narrow, tracking resulted in making graphics “jumpy”, and several colors used in the game were unnatural and jarring. Although informal and limited to only a few users, the evaluation did help to get valuable feedback about the usability of the game.

Ellis (1988) conducted rigorous experimental research on the depth perception of virtual and real objects in an augmented reality setting. He was more concerned with viewing difficulties at short distances (less than 2 meters) rather than larger distances. In Ellis’s study, participants moved a real pointer to match the depth of a virtual object under different viewing conditions in augmented environments. Interesting significant results involving depth perception and augmented environments were reported. An induced transparency occurs in users wearing a see-through display when a virtual object is placed behind a real, normally opaque, object. The placement of objects causes a perceptual transparency in the real object that creates a heightened sense of integral perception. This effect may enhance integral perception; however, the objects perceptually lose their correct depth.

Baird (1999) empirically demonstrated the benefits of using wearable augmented reality computing systems in manufacturing assembly tasks. Baird compared single-eye opaque and see-through augmented reality systems against traditional computer-display and paper-based instructional methods. Participants assembled computer boards using the four instructional techniques. Baird found that augmented instructional methods were superior to either the computer display or paper-based methods. Participants were not only faster at assembly with the augmented systems but also made fewer errors. Participants reported some usability problems such as poor image contrast and comfort issues with the head-mounted display. While no attention was paid to the effects of change in eye focus between the computer generated images and real objects on the user, Baird did show experimentally that augmented reality can aid in manufacturing assembly.

Furmanski, Azuma, and Daily (2002) focused on the complex problem of visualization of occluded objects by using guidelines from perception, psychology, and cognitive science literature. They were concerned with conveying the difference between what is normally perceptible and what is extra-sensory to the user in ways that are easy to visualize in a cluttered and complex environment. Based on the survey of literature they developed their own guidelines to support visualization in augmented reality. A preliminary experiment was conducted by the authors to determine the cues that are most effective in conveying depth information and to test their own system. The results of this experiment revealed that occlusion is the dominant perceptual cue for depth judgments. The other cues that could affect the localization of augmented information are size, transparency, color, and motion. Cues from motion parallax failed to convey depth in this study.

McGee (1999) studied the perceptual, cognitive, and human factors implications of combining real-world scenes with computer-based images. He conducted three experiments to assess the perceived integrality of the augmented scenes. In the first experiment participants subjectively assessed the integrality of different scenes at two levels of transparency (opaque and transparent). He reported that augmented reality scenes where the computer-generated graphics have a high level of transparency are perceived more integrally. From the other two experiments, he found that increasing the number of computer-generated images from two to eight decreased integral perception, whereas high integral graphics aided performance on learning and assembly tasks. Guidelines for presenting information via augmented reality were also developed based on the results of the experiments.

The Touring Machine that was developed by Feiner et al. (1997) at Columbia University is a prototype mobile augmented reality system. To evaluate the system the authors themselves walked around Columbia University campus using the augmented reality system. Augmented information consisted of textual labels with information about surrounding buildings. They wore a see-through display to view real and synthetic visual information, used a handheld computer the size of a personal data assistant to interact with the system, and carried supplementary equipment in a backpack. The informal evaluation helped the authors to arrive at several implications for users of their system. The most critical one was that even a very small number of augmented graphics can create a good deal of confusion depending on the task. Graphics made it very difficult to multitask both walking and extracting information from the augmented reality system. Although not addressed in this research, one of the important issues that might affect performance in such augmented reality applications is the change in eye focus to different depths between the real-world view and the virtual text.

The discussion above highlights the scarcity of a user-centered approach in augmented reality research and development. The user has been incorporated only to test new systems. Although several researchers have called for a greater emphasis on human factors in augmented reality (e.g., Azuma, 1997; Caudell, 1994; Helander, 1998; Kalawsky, Stedmon, Hill, & Cook, 2000), currently there are only a few scattered efforts. Researchers have directed efforts at improving the state of the technology and have paid inadequate attention to the study of the impact of augmented reality on the user. It is important to determine how humans perceive information in augmented environments before building new and improved technologies. With this in mind, the goal of this thesis is to help fill the void of information about human perception and cognition in augmented reality systems.

2.5 Human vision and accommodation

The optical media for the human eye consist of the cornea, retina, acqueous humor, crystalline lens, and the vitreous humor (Goss & West, 2002). The crystalline lens shown in Figure 3 is a biconvex transparent body of viscoelastic collagen-filled cells contained in a lens capsule behind the iris. The lens capsule is suspended from ciliary muscles by strands of tissue called zonules. The function of the lens is to focus light on the retina. In order to focus on objects that are at different distances from the eye, the lens needs to change its focal length. The ability of the eye to adjust its focus length is known as *accommodation*. According to the Helmholtz theory, when the eye accommodates on a nearby object, the ciliary muscles contract, causing the lens to assume a more convex shape (Goss & West, 2002). This is illustrated in Figure 3, which shows the accommodated lens expanding longitudinally. On the other hand, when a distant object is viewed, the ciliary muscles relax, increasing tension on the zonules which pull on the capsule and crystalline lens, causing the lens to become thinner and flatter. Figure 3 compares the lens shape of an unaccommodated lens with an accommodated lens.

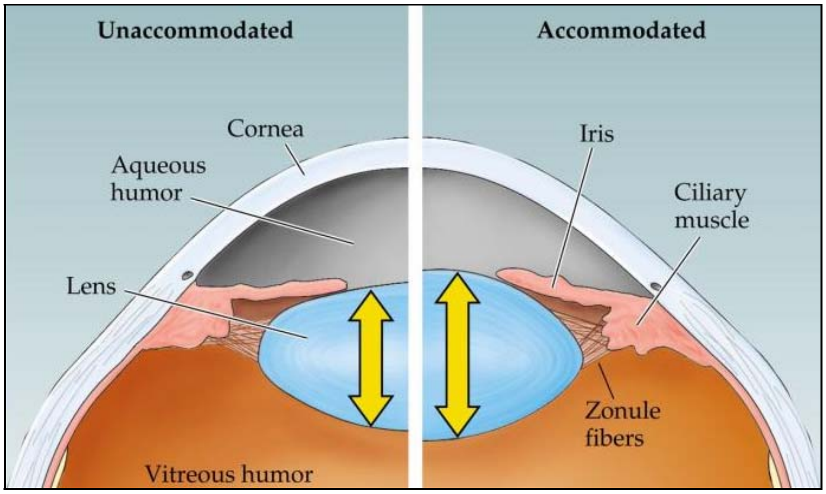

Figure 3. Cross-section of the human eye

Note. From *Neuroscience*, by D. Purves, G.J. Augustine, D. Fitzpatrick, L.C. Katz, S. Lamantia, and J.O. McNamara, et al. (Eds.), 2001, New York: Sinauer Associates. Copyright 2001 by Sinauer Associates. Reprinted with permission of copyright holder.

Four conditions are identified as effective stimuli for accommodation. These are blur, proximity, binocular disparity, and empty field (Goss & West, 2002). In the absence of any stimulus, accommodation does not go to a zero level but instead comes to rest at an intermediate level called the dark focus or the resting point of accommodation. Thus the dark focus is the distance at which the eyes are focused when there is nothing to focus on (Tufano, 1997). Accommodation of the eye is measured in diopters (Moses & Hart, 1987), which is the reciprocal of the distance at which the eye is focused in meters. The ability of an individual to accommodate at different distances decreases with age. People become slower at accommodating as they grow older (Moses & Hart, 1987) and the ability to focus at different distances is severely deteriorated by the time they are forty (Fakuda, Kanada, & Saito, 1990).

Until the 1970s researchers believed that the dark focus for the eyes was optical infinity, however investigations following invention of the optometer reported that accommodation was in fact at an intermediate distance of about an arm's distance (Iavecchia, Iavecchia, & Illiana, 1988; Jaschinski-Kruza, 1991; Norman & Ehrlich, 1986; Tufano, 1987). Measurements have revealed huge variability in the dark focus of individuals (Miller et al., 1983; Norman & Ehrlich, 1986). Leibowitz & Owens (1978) reported a dark focus mean value of 1.5 diopters for a sample of 220 students and Simonelli reported a mean of 2.67 diopters for a sample of 114 students versus a value of 1.19 diopters for 154 U.S. Air Force recruits (cited in Norman & Ehrlich, 1986). Standard deviations of 0.77 diopters and a range of 4.0 diopters have been found in the Leibowitz and Owens sample, whereas Simonelli found standard deviations of 2.57 diopters for students and 1.5 diopters for the Air Force recruits. These huge variations in dark focus were attributed to differences in individuals.

Accommodation of the eye in either direction causes eyestrain (Miller et al., 1983). In other words, a stimulus placed closer than the dark focus position or placed farther than the dark focus will result in eyestrain. Accommodation requires relatively large amounts of muscle constriction and is most closely related to eye fatigue (Östeberg, 1980). When the eyes are focused at a distance corresponding to the dark focus, ciliary muscles are relaxed and there are only small fluctuations in accommodation (Jaschinski-Kruza, 1988). Based on functional properties of the accommodation system, several researchers concluded that accommodative strain on the visual system will be low if work is performed at a viewing distance that is same as the dark focus (Johnson, 1976; Östberg, 1980; Roscoe, 1985).

Raymond (1986) measured reaction time to a subtle transformation of one letter to another (e.g., E to B) for a monocular condition. The stimulus changed randomly between two and twelve seconds and the participant was expected to press a button as soon as the stimulus changed. The experiment was conducted using a monitor at different distances from the user. She found that reaction time was dependent on viewing distance and was shortest at an intermediate viewing distance that was different for each individual. Faster reaction at an intermediate distance was attributed to the image being at the individual's dark focus, although in this experiment no physiological measurements of dark focus were taken.

Research on virtual image displays attempted to determine the influence of an individual's accommodative mechanisms on detection of distant targets. Participants performed a complex task of detecting and recognizing targets presented at infinity, while monitoring a virtual image display at different distances from the user. Some interesting significant results revealed an effect of distance on detection and recognition performance. The authors recommended adjustment of the virtual image display to match each individual's accommodative mechanism (Norman & Ehrlich, 1986). This study not only accounted for the dark focus but also range of accommodation. Based on the results of two experiments, Iavecchia et al. (1988) also concluded that the eye focus for any stimulus is dependent on an individual's dark focus. The eye tends to focus within a range around the dark focus depth, and how far the eye moves away from the dark focus is determined by ambient conditions, acuity demands of the task, and existence and nature of a textural gradient.

A study on visual display units was conducted to determine if optimal viewing distance was determined by an individual's dark focus. Jaschinski-Kruza (1988) tested two groups of people having dark focus of 1 diopter and 2 diopters (corresponding to a distance of 50 centimeters and 100 centimeters, respectively). Participants performed a search and comparison task for two hours with the visual display unit placed at a distance of 50 centimeters and 100 centimeters. In both groups, visual strain was higher for the 50 centimeter distance and 100 centimeter distance was preferred. Jaschinski-Kruza concluded that favorable viewing distance is not necessarily the distance that agrees with dark focus. However he also agreed that dark focus does play a role in determining optimal viewing distance, as individuals with a dark focus of 2 diopters experienced less strain compared to the 1 diopter dark focus group when viewing at a distance of 50 centimeters.

The practical utility of dark focus came under serious doubt as a result of Jaschinski-Kruza's study where people with the dark focus of 50 centimeters preferred the 100 centimeter viewing distance. Moreover, different values of dark focus obtained by laser and infrared measurement techniques led researchers to look for other answers (Andre & Owens, 1999).

2.6 Role of vergence in human vision

While accommodation helps to focus the eye at different distances by changing the focal length, another mechanism called *vergence* helps to actually view objects that are closer to the eye (Davson, 1990). *Convergence*, which is a type of vergence movement, is the inward rotation of the eyes in opposite directions and is used to view objects that are very close to the user. Convergence allows the image of objects to be projected at the same relative place on each retina and prevents double vision. Viewing objects at close distances exerts a greater strain on the muscles that converge the eyes. Similar to the dark focus or the resting point of accommodation, there is also a resting point of vergence or dark convergence which is the distance at which the eyes converge when there is no object to converge on. According to Owens, individual values of dark convergence lie between 0 to 2.5 meter angles (as cited in Jaschinski-Kruza, 1991). Both accommodation and convergence movements work together in order to maintain clear vision.

Jebaraj, Tyrrell, and Gramopadhye (1999) addressed the issue of the combined influence of dark focus and dark convergence through a visual inspection task. Participants had to search for defects in enlarged images of contact lenses which took 40 minutes each for distances of 20 centimeters and 60 centimeters. Results revealed a significant difference between the two distances, with the 20 centimeter condition taking almost twice the amount of time. Reported visual fatigue was also greater at the 20 centimeter distance. Physiological measures of dark focus found that neither inspection performance nor fatigue was related to the dark focus of the participant. Though the authors were unable to explain the relationship among dark focus, dark convergence, viewing distance, and their effect on visual performance, they concluded that it would be better to fit the job to the individual than to devise guidelines which may not be effective for everyone due to huge individual differences.

It has been demonstrated through previous studies that participants with a far dark convergence experience greater fatigue and decrement in performance when viewing at close distances (Owens & Wolf-Kelly, 1987). However Heuer, Hollendiek, Kröger & Römer reported no relationship between visual fatigue and dark convergence (as cited in Jaschinski-Kruza, 1991). Also dark accommodation does not correspond with optimal viewing distance as discussed before. Jaschinski-Kruza (1991) attempted to determine the viewing distance that produced least strain on both accommodation and convergence mechanisms. In a search and comparison task, it was found that dark convergence may be more important than dark focus in determining optimal viewing distance for continuous visual display unit work. The more distant an individual's dark convergence, the greater visual fatigue the individual is likely to experience when focusing at near distances. Thus optimal viewing distance should match an individual's dark convergence. But dark convergence ranges upto infinity and measurement of dark convergence is not simple. This limits the practical utility of dark convergence as a measure for viewing distance.

In another study, Jaschinski-Kruza (1990) evaluated performance when gaze was shifted every two seconds between a computer screen and document under two conditions: 1) both screen and document were at 50 centimeters from the user; 2) screen was at 70 centimeters from the user whereas the document was positioned 50 centimeters away. Contrary to expectations, visual strain was not greater when shifts in gaze were required to different distances as compared to when they were at the same distance. People preferred a viewing distance between 50 and 84 centimeters with a mean of 64 centimeters. This study provides evidence against the widespread use of ergonomic guidelines that recommend a monitor distance of 50 centimeters from the user.

To date we do not know the exact oculomotor mechanism responsible for viewing images at different distances nor its effect on human visual performance. Various theories have been put forward but none have resolved the controversy surrounding accommodation, convergence, viewing distance, and visual performance. Due to wide individual differences in dark focus and dark convergence, it is next to impossible to establish standards to optimize performance (Jebaraj et al., 1999). Although there is no conclusive evidence in literature on optimal viewing distance or the underlying mechanism for providing such recommendations, ergonomic guidelines for visual display workstations continue to recommend viewing distances in the range of 30 to70 centimeters (Human Factors Society, 1984). The range of 30 to 70 centimeters may in fact be too close, as most empirical evidence has shown better performance and lesser visual fatigue at distances greater than 50 centimeters. This thesis also attempted to determine if suggestions from visual display unit literature can be applied to augmented reality for displaying virtual information to the user with the objective of minimizing visual fatigue and maximizing performance.

2.7 Relation of augmented reality to head-up displays

Head-up displays are closely related to augmented reality displays. According to Azuma (1997) head-up displays commonly used in aviation are similar to augmented reality displays. A head-up display is a virtual-image display in which instrument symbology is superimposed on the forward view of the operator (Weintraub & Ensing, 1992). Head-up displays replace or augment the instrument display in a car or an airplane cockpit. The graphics are projected onto a plane near or infront of the windshield

Differences between head-up displays and augmented reality lie in their application of technologies rather than their technical definitions. The purpose of head-up displays is to provide the driver or pilot with instrument panel information superimposed on the forward view and eliminate the need to look down on the instrument panel. Head-up displays have been extensively researched in human factors to demonstrate their benefit relative to traditional head-down instrument panels. Augmented reality, on the other hand, augments the view of the user with many different sorts of additional information. Despite differences in application areas, both augmented reality and head-up displays superimpose virtual graphics onto the real-world view of the user to enhance understanding.

A major issue in head-up displays literature is the optical distance from the user at which symbology should be displayed. The distance at which symbology is displayed should reduce the need for eye focus adjustment between the real-world view and the symbology. Inuzuka, Osumi, & Shinkai (1991) performed an experiment to determine the in-plane location, optical distance, brightness, and color of head-up display symbology. To determine the optimal distance at which to display symbology, the authors measured recognition time of five speedometers, each displayed at different distances from the user. They found that recognition time for older subjects was less at distances greater than 2.5 meters and became constant at a farther distance. Similar to this, Kato, Ito, Shima, Imaizumi, and Shibata (1992) reported no improvement in recognition time and subjective evaluation beyond a distance of 2.0 meters when participants were required to look at a road 10 meters ahead and then focus on a display placed at varied distances.

Research by Okabayashi and Sakata (1991) required participants to read Snellen's figures on an automotive head-up display by varying the fineness of the image along with the distance of Snellen's figures from 0.7 meters to 5 meters. Snellen's figures consist of characters (letters or forms) of varying sizes that are used to assess an individual's visual acuity. Correct response rate was used as the evaluation parameter to determine information displayed by the head-up display while maintaining correct recognition of the forward view. Authors found that when the head-up display image was moved away from the eye, performance improved more for older adults than for the younger group. They concluded that virtual symbology should be closer to the forward view to ensure correct recognition of information as it reduces eye accommodation. Weintraub (1984) had similar findings from his experiment where he varied the distance of the display symbology, luminance, and position of the display. A measurement of time taken by pilots to make a decision about landing revealed that as display symbology was moved closer to the eye, decisions about landing the plane took longer. Both these studies concluded that display symbology should be closer to the forward view and away from the eye.

The findings from the study by Weintraub (1984) and Okabayashi and Sakata (1991) are different from those discussed earlier by Inuzuka et al. (1991) and Kato et al. (1992). While Inuzuka et al. and Kato et al. recommend that the symbology be displayed at approximately 2.0 meters from the user's eye, Okabayashi and Sakata and Weintraub recommend a distance closer to optical infinity.

A study by Wolffsohn, Edgar, and McBrien (1998) examined the effect of ocular accommodation and response time on automotive head-up display images by varying display distances from 0.8 meters to optical infinity. The task required participants to maintain attention to head-up display images as well as attend to distant real-world objects at a distance of 6 meters from the observer. They found no difference in accommodative levels of participants over the range of distances. The average eye focus distance was found to be between 3 and 4 meters despite the primary task of viewing distant objects at 6 meters. This was attributed to the cognitive demands placed on the user to detect and respond to changes in the outside world scene and the head-up display image. Also response time to detecting changes in stimulus did not vary with the changing distance of the head-up display. However, participants took longer to respond to changes in the head-up display image when they had to refixate from the distant real-world image as compared to when they only had to focus on the head-up display image. This finding indicates the negative impact of accommodation on task performance arising as a result of objects or graphics being situated at different distances from the user. Also, accuracy was lower at distances closer than 1 meter as compared to distances equal to or greater than 1 meter.

There seems to be no consensus amongst researchers on the optical distance at which to display head-up display symbology (Weintraub & Ensing, 1992; Wolffsohn et al., 1998). Despite this it is accepted in the surface transportation community that automotive head-up displays should be focused at between 2.0 and 2.5 meters (Tufano, 1997). Tufano argues that the justification for displaying head-up display symbology at approximately 2.0 meters is based on portions of studies by Inuzuka et al. (1991) and Kato et al. (1992) that addressed only the effect of image distance on extraction of information from the head-up display and not its effect on perception of real objects in the driver's forward view. This is an important observation since the primary task of the user is to attend to objects in the real world. However, effect of head-up display image distance on perception of real-world objects has not been addressed.

Research to determine the distance for head-up display images is very relevant to augmented reality. As mentioned earlier, augmented reality is similar to head-up displays as they both superimpose virtual images on the real-world view of the observer. Although the issue of optimal image distance for head-up displays has received a lot of attention from the research community, there is no research that examined this problem in the context of augmented reality.

2.8 Depth cues

In virtual and augmented reality systems, visual impression of depth for virtual information can be created by employing different depth cues. Cutting (1997) has discussed nine different depth cues along with their relative efficacy based on just noticeable differences (JND) in depth of two objects at different distances from each other. These nine sources for creating depth impression are occlusion, height in visual field, relative size, relative density, aerial perspective, binocular disparity, accommodation, convergence, and motion parallax. Depending on the distance of objects from the user, relative effectiveness of the cues change. While occlusion, relative size, and relative density are strong depth cues at all distances, accommodation, convergence, and binocular disparity are strongest at distances less than 2 meters (Cutting, 1997). Aerial perspective, on the other hand, is a stronger cue at distances greater than 1000 meters. Motion parallax is a very good cue for judging absolute distance up to 5 meters and relative distance at greater than 5 meters (Cutting, 1997).

Depth cues to be used in this thesis study include focus depth, motion parallax, and relative size. These depth cues help the user to perceive the distance of virtual information from the real-world information as well as from oneself. Focus depth is a depth cue which allows the user to perceive the distance to virtual text or objects. It represents the distance from the user's eye to the point at which the eyes focuses on virtual graphics and the graphics appear "in focus". Focus depth is controlled by accommodation and convergence of the eye. As discussed earlier, accommodation is the mechanism of the eye by which it changes the shape of its lens to focus on objects that are closer or farther away. Objects at other distances in the field appear blurred and out of focus. In order to accommodate on an object at a certain distance, our brain tells the muscles of the eye to contract or expand to bring the object into focus. Convergence is the mechanism by which the eyes rotate inwards in opposite direction in order to focus on an object that is close. This inward rotation of the eyes helps the user to infer the distance of the object. The amount of accommodation and convergence that the visual system undergoes can be used as cues to depth but they are effective only at shorter viewing distances of 2 to 3 meters (Flannagan, Sivak, & Simpson, 2001).

Motion parallax refers to the relative motion of objects across the retina where near objects move faster across the retina than far objects. When users move their heads, by virtue of motion parallax they will be able to infer depth based on the relative motion of objects in the field. Near objects will move faster than far objects. Since in this thesis study users will be stationary and may only occasionally move their head, motion parallax will be a weak depth cue.

Relative size, as a depth cue for this thesis, is defined as the relative size of virtual and real-world text. At different distances from the user, the size of real-world text on the monitor and virtual text will appear to vary. Thus at farther distances the text will appear smaller whereas at closer distances it will appear larger. This will enable the user to perceive relative distance of the text.

2.9 Focal length of display and focus depth

Focal length of a display in physical terms is defined as the distance from the center of the lens to a point where all light rays passing through the lens converge. It is a property of the hardware of the display. Focal length is shown in Figure 4.

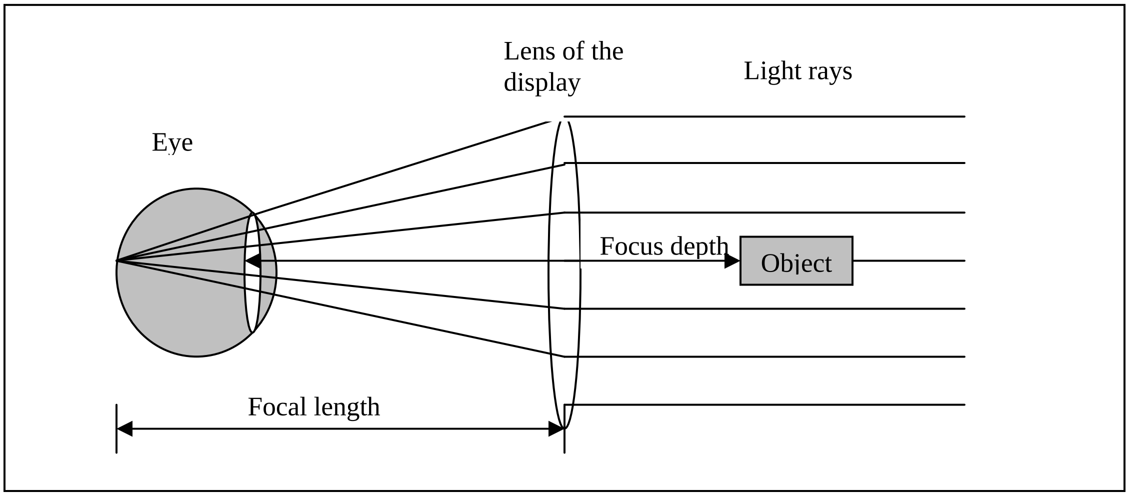

Figure 4. Focal length of a lens

A shorter focal length may cause objects to appear farther away, whereas longer focal lengths cause objects to appear closer (Glumm, Kilduff, & Masley, 1992). Changing the focal length also affects the field of view. The longer the focal length, the shorter the field of view, and visa versa (Glumm et al., 1992). Thus focal length is an important parameter that controls other display properties. The focal length of a lens is measured in diopters. It can also be measured in meters, which is the reciprocal of the value in diopters.

Focus depth is distinct from focal length although the two are mathematically related. Focus depth is the distance at which the eye is focused in order to clearly look at an object. If there is an object 1 meter away, the brain tells the lens muscles of the eyes to contract or expand so that the object is “in focus”. Focus depth is one of several depth cues which are used to perceive depth of an object in the field. Some head-mounted displays have an adjustable focus depth ranging from one foot to optical infinity. Optical infinity is defined as a distance so far removed from the lens that rays of light reflected to the lens from a point at that distance may be regarded as parallel. A distance of 6 meters can be considered as optical infinity (Simonelli, 1979).

A study was conducted to determine the effect of three focal lengths with the objective of identifying the focal length that maximizes performance on remote driver performance (Glumm et al., 1992). The results revealed a significant difference among the three focal lengths of 12 millimeters, 6 millimeters, and 3.5 millimeters, with the 6 millimeter distance being superior in speed and accuracy. The shorter focal length of 3.5 millimeters resulted in a bird’s eye view and most participants were uncomfortable with it. The intermediate 6 millimeters focal length offered an acceptable field of view and depicted object distance most accurately.

Ijsselsteijn, Ridder, and Hamberg (1997) assessed the influence of image disparity, convergence distance, and focal length on subjective assessment of depth, naturalness of depth, and quality of depth for a stereoscopic display. They found that people prefer stereo presentation of images with an optimal disparity of 4 centimeters. The effect of focal length was most pronounced at higher image disparity levels with the shorter focal length of 10 millimeters preferred at all the three dependent variables. The scarce literature on the topic of focal length may indicate that focal length does have an effect on comfort and quality of viewing an image. However, in the context of augmented reality, no research has systematically examined the impact of focal length and focus depth on human performance.

2.10 Summary of literature review

While technological development has been given high priority in augmented reality, the simultaneous corresponding emphasis on human factors issues is missing. A detailed survey of literature in augmented reality was conducted to elicit issues that researchers have been studying. We found that most of the studies are technology driven and very little work has been done to understand the impact of augmented reality on human perception, which should be accorded high priority in order to advance the development of usable technology. One such piece of missing research is the effect of change in eye focus between real-world objects and virtual graphics on user task performance. Also, there is virtually no research that has examined the impact of focus depth on human task performance in augmented reality systems.

From literature on the physiology of the human eye, it was found that the eye experiences accommodation and convergence when focusing on objects at different depths. Accommodation and convergence can result in visual fatigue, blurriness, double vision, and hence decreased performance on tasks. Due to the need to accommodate and convergence to focus on objects located at different distances from the user, augmented reality may not be able to reach its potential. Various theories have been put forward to better understand the functioning of the eye. Studies have demonstrated lesser visual fatigue when the eye is focused at an arm's length. Literature in the field of head-up displays, which is similar to augmented reality, was also been surveyed to seek answers to these complex questions. Researchers have been pondering over the issue of optimal head-up display image distance for over two decades. While earlier it was believed that optical infinity was better for displaying images, some later studies showed that 2.0 to 2.5 meters is superior for displaying head-up display images. Despite attempts to determine the optimal display distance, there is no consensus amongst researchers on findings and as such the question still remains unanswered.

Literature review also revealed that focus depth is a completely unexplored issue in augmented reality. This thesis attempted to contribute to the field of augmented reality by answering questions related to human perception and cognition in augmented environments and laying a foundation for future work.

3. METHODS

3.1 Goals and hypotheses

This thesis consisted of three goals which are stated below:

1. To determine if there is any difference in user task performance when information is presented in an exclusively real-world environment as compared to an augmented environment.
2. To determine the impact of focus depth on user task performance.
3. To determine if there is any difference in user task performance when distance to virtual text, focus depth, and distance to real-world text match (i.e., are all the same value), as compared to conditions when they are not matched.

Based on the three goals mentioned above, three research hypotheses were formulated which are stated as follows:

1. User task performance will be better when all information is presented in an exclusively real-world environment as compared to conditions when information is presented in an augmented environment.
2. Focus depth will have a significant effect on user task performance.
3. User task performance will be better when distance to virtual text, focus depth, and distance to real-world text match (i.e., are all the same value), as compared to conditions when they are not matched.

The hypotheses were tested using an empirical study.

3.2 Experimental design

This study consisted of a 3x3x3 within subjects design. The independent variables were distance to real-world text, distance to virtual text, and focus depth. Figure 5 diagrammatically shows the three independent variables and their relationship with each other.

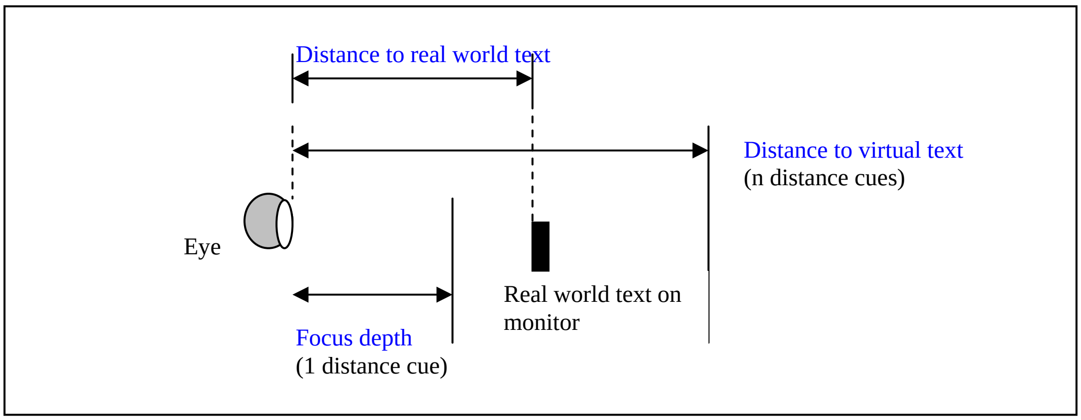

Figure 5. Relationship among independent variables

Description of the three independent variables is as follows:

1. **Distance to real-world text** is the distance from the user's eye to the real-world text on the monitor, measured in meters.
2. **Distance to virtual text** is the perceptual distance of the virtual text from the user's eye. It is a geometric property of the software of an augmented reality system and is encoded by three *depth cues* for this study: relative size of text, motion parallax, and focus depth. It represents the distance at which the brain interprets the virtual text to be positioned in the field. Distance to virtual text is manipulated by the software that supports the augmented reality display. While focus depth was manipulated separately as an independent variable, the other two

depth cues (relative size of text and motion parallax) were manipulated together so that they were congruent with each other and consistent across all conditions.

3. **Focus depth** is the distance from the user's eye to the point at which the user's eye focuses on the virtual text and the text appears "in focus". It is an optical property of the hardware of the display and is one of several depth cues. Focus depth can be matched with all the other depth cues or can be mismatched as shown, for example, in Figure 7. Focus depth in this study was manipulated by moving a slider on the hardware display, discussed in Section 3.6, to cause the eye to focus at different depths in the augmented world.

To further clarify the concept of depth cues and perception of distance, consider the case of a user in a virtual reality CAVE environment. Graphics are projected on the CAVE wall. In the case of a CAVE, focus depth is the distance between the user's eye and the CAVE wall. However, users may not perceive focus depth to be the actual distance between themselves and the virtual graphics. The CAVE can give the user a feeling of immersion in an environment with different depths (e.g. users may perceive themselves to be immersed in a 100 meter room though the CAVE wall may only be 6 meters away). This is due to the influence of other depth cues, such as binocular disparity, motion parallax, relative size, shading, occlusion etc., which can provide the user with perception of depth. Depending upon which cue or combination of cues is more powerful, the user may perceive the distance to virtual graphics as being encoded by those depth cue(s) and not necessarily the distance represented by focus depth.

Each of the three independent variables was presented at three levels corresponding to distances from the user of near, medium, and far. The three levels are represented in Table 1. The justification for selection of these three levels of the independent variables is included in Section 3.4.

Table 1. Levels of independent variables

Distance to real-world text	*Distance to virtual text*	*Focus depth*
Near – 0.7 m	Near – 0.7 m	Near – 0.7 m
Medium – 2 m	Medium – 2 m	Medium – 2 m
Far – 6 m	Far – 6 m	Far – 6 m

The experimental design consisted of a treatment condition with 24 unique cells and two control conditions. In control condition 1, the experiment was exclusively performed in a real-world environment. In control condition 2, all three independent variables were matched to the same distance. The treatment condition and control conditions are shown in Tables 2 and 3 respectively.

Table 2. Treatment condition and control condition 2 data matrix (m = meters)

Distance to real-world text	*Focus depth*	*Distance to virtual text*		
		Near (0.7 m)	Medium (2 m)	Far (6 m)
Near (0.7 m)	Near (0.7 m)	Control 2		
	Medium (2 m)			
	Far (6 m)			
Medium (2 m)	Near (0.7 m)			
	Medium (2 m)		Control 2	
	Far (6 m)			
Far (6 m)	Near (0.7 m)			
	Medium (2 m)			
	Far (6 m)			Control 2

Thus the 24 unlabeled cells in Table 2 represented the treatment condition when the levels of the independent variables were mismatched. The shaded cells labeled "control 2" represent control condition 2 when the levels of all three independent variables were matched.

Control condition 1 is represented in Table 3. In control condition 1, the experiment was performed exclusively in the real world using two monitors, A and B. Since control condition 1 was in the context of the real world, focus depth no longer existed as an independent variable. Only the cells labeled "control 1" represent control condition 1 for this study, as both the monitors were at the same physical distance from the user for these three cells.

Table 3. Control condition 1 data matrix (m = meters)

Distance to text A	*Distance to text B*		
	Near (0.7 m)	Medium (2 m)	Far (6 m)
Near (0.7 m)	Control 1		
Medium (2 m)		Control 1	
Far (6 m)			Control 1

3.3 Dependent variables

The dependent variables for this experiment were:

1. Accuracy, which was defined as the number of correct responses for each task.
2. Task completion, which was defined as the number of subtasks completed for each task in a time period of 25 seconds. The time limit of 25 seconds per task was based on an informal pilot study which revealed that participants took a mean time of 30 seconds to complete one task consisting of five subtasks. For the experiment, time limit for each task was reduced to 25 seconds in order to impose some pressure on participants and prevent their completing all subtasks. The decision to impose a time limit for tasks was based on the results of a similar augmented reality study conducted at Virginia Tech that revealed only a 1.5% error in response (Gabbard, 2003). It was hypothesized that the extremely low error rate was due to the time versus accuracy tradeoff. In order to maintain high accuracy, participants increased the time for task completion, because they were under no time constraints. To protect against this, a time limit was imposed to

force participants to perform tasks/subtasks as fast as possible while maintaining high accuracy which in turn would eliminate the low error rate.

3.4 Rationale for levels of independent variables

Selection of the three levels for distance to virtual text was based on inconclusive findings in literature that claim better performance when the eye is focused to these distances. Researchers have found that when the eye focuses at approximately an arm's distance, accommodation is minimal and the eye is in a relaxed position (Iavecchia, et al., 1988; Jaschinski-Kruza, 1991; Norman & Ehrlich, 1986; Tufano, 1987). An arm's distance has been approximated to about 0.7 meter, which is the first level chosen for the virtual text distance. The rationale for the second level of the independent variable is based on evidence found in head-up displays literature that states that a distance of approximately 2 meters leads to faster performance on visual detection tasks (Inuzuka et al., 1991; Kato et al., 1992). Other research on head-up displays has also reported that display symbology at optical infinity is better for visual tasks (Okabayashi & Sakata 1991; Weintrub, 1984). According to Simonelli (1979), optical infinity is roughly 6 meters. Hence 6 meters was selected as the third level of the virtual text distance.

Levels of the other two independent variables were the same as the distance to virtual text. Real-world object distances of 0.7 meter, 2 meters, and 6 meters encompass a wide range found in indoor augmented reality applications. Working distance for personal space applications like surgery and maintenance and repair is roughly an arm's length, whereas navigation in an indoor environment, which is constrained by the size of a room, can be approximated to 6 meters. The farthest object distance of 6 meters is likely to be frequently found in outdoor augmented reality applications as well. Manipulation of focus depth from 0.7 meter to optical infinity included a wide range supported by the Nomad display used in this study. The Nomad display is described in Section 3.6.

3.5 Participants

Twenty-four graduate and undergraduate students (14 male and 10 female) from Virginia Tech participated in the study. The mean age of participants in the study was 22.58 years with a standard deviation of 2.48. Twenty-four participants were sufficient to satisfy a power requirement of 0.8, which represents a realistic and reasonable value for research (Cohen, 1977). A "large" effect size approximated as 0.4 by Cohen at a significance level of 0.05 was used to determine the sample size from Cohen's sample size tables (Cohen, 1977, p. 384). Also, multiples of 12 participants were required for counterbalancing the levels of distance to real-world text and order of presentation of control condition 1 and treatment condition. Counterbalancing is discussed in section 3.8.

Participants were recruited using flyers posted on campus bulletin boards, newsgroups, email, and word-of-mouth. They were compensated at the rate of $10 per hour or secured course credit for their participation. Participants were screened according to the following criteria:

1. Should be less than 35 years in age.
2. Should have at least normal or corrected 20/25 far vision. This was tested using Snellen's Test for visual acuity described in Appendix A.

3. Should have at least normal or corrected 20/25 near vision. This was tested using the Runge Near Point Card Test described in Appendix B.
4. Should not be diagnosed with attention deficit disorder (ADD). This requirement was included in the call for participation flyers and emails, and was simply self-selected. Questions related to ADD symptoms were also included in the pre-experiment questionnaire to examine any outliers from speed and accuracy measures.

Selection of the above-mentioned criteria is based on input from augmented reality display manufacturers who claim that users may suffer from distortion in vision if they do not have perfect vision (Obeysekare, U.R., personal communication, January 22, 2004). Furthermore, after approximately 40 years of age, the ability of the human eye to accommodate decreases (Moses & Hart, 1987) and we wanted participants who would be able to accommodate. A conservative approach was adopted and participation was limited to individuals less than 35 years in age. The restriction on participants suffering from attention deficit disorder was due to the possibility that such participants might experience additional difficulty in switching between real world and virtual world. To prevent any confounding effect on task performance, individuals not satisfying all the above criteria were not eligible to participate in the study.

3.6 Experimental apparatus

The original goal of this thesis was to use a binocular see-through head-mounted display that would provide stereoscopic vision as a depth cue. However, the display did not support proper functioning of stereo. To produce functional stereoscopic images, the presentation device must be able to present left and right eye images that are exactly equivalent in image quality (resolution, brightness, contrast, aspect ratio, etc.). After significant effort, we determined that it was impossible to adjust settings of each control box so that the images were precisely equivalent in terms of optical adjustments. This, in itself, was a severe limitation of the two-eyepiece design, and for our purposes and rendered display in-operable for stereoscopic augmented reality research and development. In addition to the above-mentioned problem we could not get the left and right eye image planes to be the same shape and size, and as a result could not get the two images positioned into a single, fused left/right image plane. To create functional stereoscopic images, the left and right eye viewing planes must be the same shape and size and must also perfectly overlap to create a single left/right image plane. These findings have been documented to improve the development of stereoscopic head-mounted displays for augmented reality.

Due to the limitations of the binocular head-mounted display, we used a monocular see-through head-mounted display called Nomad, made by Microvision. The unique feature of this head-mounted display was its adjustable focus depth. The Nomad projects very low intensity laser beams onto the retina of the eye which allows the user to view virtual images. It has a resolution of 800x600 pixels. The field of view is 23x17 degrees for each eye, which is equivalent to a 17 inch monitor at an arm’s length distance. Luminance of the Nomad ranges from 1 to 800 foot lambert. It supports monochrome red only for the display color. The total head-worn weight of the Nomad is approximately 226.8 grams. Figure 6 shows the monocular Nomad used in this research.

Figure 6. Nomad

A movable slider on the Nomad allowed adjustment of focus depth from 1 foot to optical infinity. As the focus depth slider of the Nomad was not calibrated, an initial calibration was performed using a dioptermeter with a -5/+5 range. A dioptermeter is an ophthalmologic instrument. The dioptermeter and eye pieces of the Nomad were mounted on an optical rail with a separation of 30 millimeters between the two. The set-up is shown in Figure 7. The Nomad was connected to a laptop which sent a test pattern consisting of text and fine lines to the Nomad display. An initial adjustment of the dioptermeter was performed by turning the eye piece dial to put the cross-hair in focus. Then the focus range of the dioptermeter was set to the desired level (e.g., for 2 meters distance the focus range was set to 0.5 diopters). While viewing the Nomad display through the dioptermeter, the slider on the Nomad was adjusted till a sharp test pattern image was obtained. This slider position was marked on the display as this corresponded to one of the focus depth levels. Similarly, focus depths of 0.7 meter and 6 meters were marked on the Nomad display based on measurements using the dioptermeter.

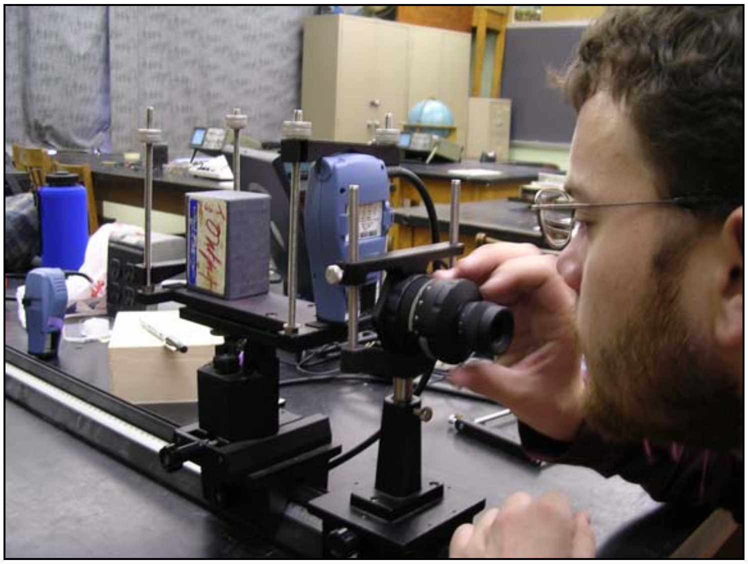

Figure 7. Calibration of Nomad using a dioptermeter

The testbed for the augmented reality system used in this study was a custom application developed at Virginia Tech using the Diverse toolkit also developed at Virginia Tech. Graphics rendering was done through Diverse using Open GL Performer. The code was written in C++ and ran on Linux OS. Head tracking for the system was done using the Intersense IS 900.

Two 19 inch LCD monitors set to a resolution of 800x600 were also used in the experiment. The monitors were placed on two portable carts so that they could be moved easily to the different levels of distance to real-world text. External lighting in the room was controlled by using black curtains and lights in the room were switched off for better visibility. Also, virtual text was projected on the black background.

Real-world text on the monitor and virtual text were positioned side by side so that they did not occlude each other. To ensure legibility of text at all distances, the ANSI standard of 22 arc minutes for minimum height of characters was followed (Human Factors Society, 1984).

Positioning of the monitors relative to the user is depicted in Figure 8. At the near distance of 0.7 meters from the user, there was no separation between the two monitors. At 2 meters, the centers of the monitors were set apart by a distance of 81.5 centimeters whereas for 6 meters, they were separated by a distance of 196.5 centimeters. This was based on geometrical calculations to standardize the horizontal eye scanning at the three distances. The same criteria applied for the treatment condition where one of the monitors was replaced with virtual text through the Nomad.

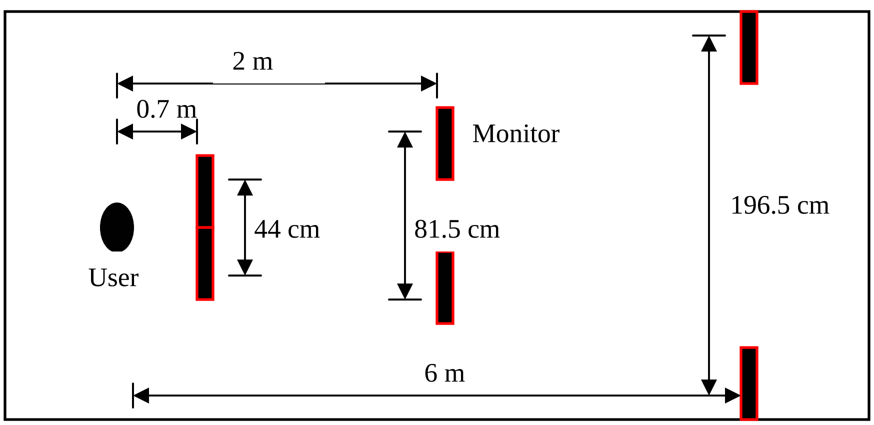

Figure 8. Top view of experiment set-up (not drawn to scale)

3.7 Experimental task

The experimental task was designed to force participants to switch focus between real world and virtual world and hence repeatedly context-switch. The task was a relatively low-level cognitive task consisting of perception of characters, scanning, recognition, memory, decision making, and motor response.

For treatment condition and control condition 2, a single monitor was used to display real-world text. Virtual text was displayed using the Nomad on the real-world background, which consisted of a black cloth hung from the ceiling in the back of the experimental space. Software for the Nomad only supports red color for the virtual text, hence for consistency, the color of font on the monitor was also red with a black background.

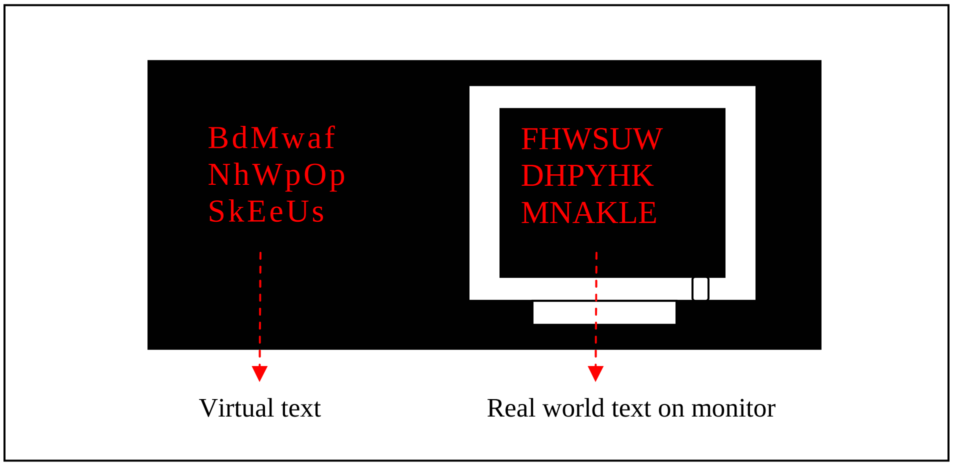

Figure 9. Treatment condition and control condition 2

As shown in Figure 9, the participant saw three strings of upper case random characters from the English alphabet on a monitor (i.e., real-world text) alongside three different strings of alternating upper and lower case random characters through the Nomad (i.e., virtual text). The length of each string for both monitor and virtual text was fixed to six characters for a total of 18 characters. The participant was instructed to locate a pair of identical letters, one of which was upper case and the other was lower case, from the virtual text, called the *target letter* (e.g., eE in Figure 9). Placement of the target letter in the virtual text was randomized. Alphabets “i” ,“l”, and”j” were not included in the list of possible target letters due to their similarity in appearance of upper case and lower case. On locating the target letter, the participant was instructed to look at the real-world text on the monitor and count how many times the target letter appeared in the real-world text. Placement of the target letter in the real-world text was also randomized. The participant responded by pressing the number keys on a keyboard to indicate the number of times the letter appeared in the real-world text as follows:

- 0 – target letter does not appear in real-world text
- 1 – target letter appears once in real-world text
- 2 – target letter appears twice in real-world text
- 3 – target letter appears thrice in real-world text

The participant's response constituted a subtask. After completion of a subtask, the real-world text remained the same, whereas the virtual text changed to a new set and a new target letter was indicated by repetition of a letter in upper and lower case. The participant had 25 seconds to complete a maximum of five subtasks, which constituted one task. Any input made after 25 seconds was not considered. On completion of the 25 seconds or five subtasks, whichever came first, both screens blanked for 2 seconds and a new set of virtual text and real-world text were presented to the participant. This constituted the second task. The participant performed four task repetitions for each cell of the design. After completion of one cell, participants were asked to subjectively rate eye fatigue by answering the following question which was displayed to them as virtual text: "Please rate the condition of your eyes". The scale which is shown below was also available on paper for participants. Participants responded by pressing the appropriate number key on the keyboard.

1	2	3	4	5	6	7
very rested	rested	somewhat rested	neither rested nor fatigued	somewhat fatigued	fatigued	very fatigued

To minimize carryover effect of fatigue, a rest break of 45 seconds was provided between cells where participants were instructed to close their eyes and relax. The rest break also allowed experimenters to adjust position of monitors and/or focus depth of the display. After the rest break, the next task was presented to the participant in a similar manner. The entire experiment consisted of 120 tasks for each participant.

The presence or absence of the virtual target letter in the real-world text was randomized by the following criteria:

- Virtual target letter not present in the real-world string – 25% of the total number of subtasks
- Virtual target letter present once in the real-world string – 25% of the total number of subtasks
- Virtual target letter present twice in the real-world string – 25% of the total number of subtasks
- Virtual target letter present thrice in the real-world string – 25% of the total number of subtasks.

For control condition 1, which was exclusively in the real world, two monitors were used. Strings consisting of random characters were displayed on both monitors as shown in Figure 10.

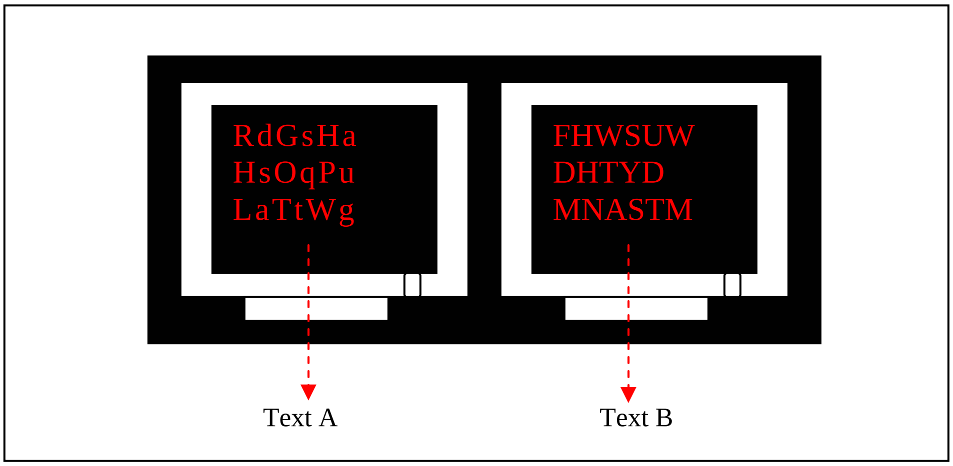

Figure 10. Control condition 1

The procedure was exactly the same as just described except that the virtual text was replaced with text A on a real-world monitor and real-world text was called text B. The participant located the repeating upper and lower case target letter in text A and counted and reported the number of times the target letter appeared in text B. Participants wore the head-mounted display for control condition 1 though it was turned off.

This task was designed to minimize pre-attentive processing. The target letter, which constituted a pair of identical upper and lower case letters, forced participants to scan through the virtual text each time the subtask was presented. Several other visual cues such as underlining, larger font size, and bold text were considered for highlighting the target letter; however we realized that this would result in pop-out phenomenon wherein the participant would locate the target by merely looking for the visual cue. Random variation of cues for tasks/subtasks was also considered but some cues being stronger than others would have a confounding effect on the experiment. The pair of identical upper and lower case alphabets as a cue for the target letter had the advantage of forcing participants to scan through the virtual text. Scanning through virtual and real-world text resulted in participants spending time interacting with both the pieces of information which is representative of real-world augmented reality tasks.

3.8 Presentation order

As mentioned earlier, the experiment consisted of 24 treatment cells and two separate control conditions each consisting of three cells. All participants received all treatment and control conditions. It was decided not to counterbalance order of presentation of the 30 cells as this would require a very large number of participants. A completely randomized design was also not feasible as this would necessitate too-frequent moving and adjusting of monitors and equipment. Hence, it was essential to counterbalance the three levels of distance to real-world text using a 6x3 Latin Square so that they varied the least and minimum moving of monitors and equipment was involved. The Latin Square is depicted in Figure 11. The presentation order of control conditions and treatment condition also needed to be counterbalanced for the same reason, using a 2x2 Latin Square which is shown in Figure 12. The three levels of focus depth and distance to virtual text were randomized. Perl's random script generator was used to instantiate the counterbalancing. This ensured that all random numbers were well-generated from a statistics package which the Perl script used. Using a Perl script made the final experimental control program stateless as it could be started and stopped at any point in the experimental control script, and was thus robust to crashes and other problems that inevitably occur while running a participant.

		Participants					
		1	2	3	4	5	6
	1	N	M	F	F	N	M
Order	2	M	F	N	M	F	N
	3	F	N	M	N	M	F

Figure 11. 6x3 Latin Square Design

Note: N represents near distance, M represents medium distance, F represents far distance for distance to real-world text. The counterbalancing was repeated for the remaining 24 participants.

		Participants	
		1	2
Order	1	T	C2
	2	C2	T

Figure 12. 2x2 Latin Square

Note: T represents treatment condition and C2 represents control condition 2. The counterbalancing was repeated for the remaining 24 subjects

3.9 Experimental procedure

Each participant was asked to read and sign the informed consent form (Appendix C). A test of near and far visual acuity was administered using the Runge Near Point Card and Snellen Chart respectively. Participants who did not satisfy 20/25 near and far vision criteria were not allowed to participate in the study. Procedure for the two acuity tests is included in Appendix A and Appendix B. Eligible participants completed a pre-experiment questionnaire (Appendix D) to collect information regarding age, vision, depth perception, and attention deficit disorder symptoms. Finally, they were given instructions about the experiment (Appendix E).

Figure 13. Participant performing the experiment

Participants were seated on an adjustable-height chair and assisted in wearing the Nomad. The Nomad eyepiece was always positioned in front of the left eye to minimize adjustment of the experimental set-up. Initial adjustment of chair height was required to eliminate the effect of differences in participant height on their view through the Nomad. Figure 13 shows a picture of a participant performing the experiment. Practice trials were performed at each of the three distances, for a total of 12 tasks, to allow participants to familiarize themselves with the task and the Nomad display.

After the practice trials, the experiment began.

Participants had 25 seconds to complete a maximum of five subtasks. They responded by pressing the appropriate number key on the keyboard. After 25 seconds, both screens blanked for 2 seconds and then the next task was presented. The following task also consisted of five subtasks which had to be completed in a time limit of 25 seconds. Different tasks were performed four times for each treatment cell and control cell, for a total of 120 tasks per participant. Following completion of tasks for one cell, participants were asked to subjectively rate the condition of their eyes on a scale ranging from 1 to 7 where 1 was very rested, 4 was neither rested nor fatigued, and 7 was very fatigued by pressing the appropriate key on the keyboard. The graphical representation of the scale which was mentioned in section 3.7 was made available to participants. After rating, participants were asked to close their eyes and relax for 45 seconds. Between treatment cells and control cells, the levels of the independent variables were also manipulated as required, by manually changing the setting of the slider on the display, moving the monitors to required distances, or changing the virtual text distance through the software which was done automatically. After the 45 seconds break period, participants proceeded with tasks for the next cell.

The same procedure was followed for treatment and control condition 1. Figure 14 shows a picture of the participant's view of control condition 1.

Figure 14. Participant's view of control condition 1

On completion of the experiment, participants were asked to answer the post-experiment questionnaire (Appendix F) to assess any eye fatigue or physical discomfort arising as a result of the experiment. An informal interview was also conducted to gather information about the experiment from participants. The interview questions are included in Appendix G. Finally they were debriefed, compensated, and thanked for participating in the study. The experiment on average lasted two hours.

4. RESULTS

As mentioned in section 3.1, this thesis consisted of three goals. Parallel to the three research goals, three hypotheses were formulated. Results of the empirical study were analyzed by inferential statistics which are presented below.

4.1 Goal 1. Context-switching

The first goal of this thesis was to determine if there is any difference in user task performance when information is presented in an exclusively real-world environment as compared to an augmented environment. It was hypothesized that user task performance would be better when all information is presented in an exclusively real-world environment as compared to an augmented environment. The hypothesis was tested by comparing user performance for control condition 1, which consisted of text on two monitors and hence required no switching between real-world and virtual information with the treatment condition which consisted of real-world text on a monitor and virtual text. The hypothesis was tested on the two dependent measures: task completion and accuracy. In addition to the objective measures of performance, subjective ratings of fatigue were also analyzed to assess the impact of context-switching on eye fatigue.

One-tailed, paired t-tests were performed to test hypothesis 1. Accuracy, task completion, and fatigue rating of cell A_1B_1 in Table 4 were compared with accuracy, completion rate, and fatigue of cell $R_1F_1V_1$ in Table 5. Similarly, cell A_2B_2 in Table 4 was compared with $R_2F_2V_2$ in Table 5, and A_3B_3 with $R_3F_3V_3$ using separate t-tests at a significance of 0.05.

Table 4. Real-world data matrix

Distance to text A (A)	*Distance to text B (B)*		
	Near (0.7 m)	Medium (2 m)	Far (6 m)
Near (0.7 m)	A_1B_1		
Medium (2 m)		A_2B_2	
Far (6 m)			A_3B_3

Table 5. Augmented world data matrix

Distance to real-world text (R)	*Focus depth (F)*	*Distance to virtual text (V)*		
		Near (0.7 m)	Medium (2 m)	Far (6 m)
Near (0.7 m)	Near (0.7 m)	$R_1F_1V_1$	$R_1F_1V_2$	$R_1F_1V_3$
	Medium (2 m)	$R_1F_2V_1$	$R_1F_2V_2$	$R_1F_2V_3$
	Far (6 m)	$R_1F_3V_1$	$R_1F_3V_2$	$R_1F_3V_3$
Medium (2 m)	Near (0.7 m)	$R_2F_1V_1$	$R_2F_1V_2$	$R_2F_1V_3$
	Medium (2 m)	$R_2F_2V_1$	$R_2F_2V_2$	$R_2F_2V_3$
	Far (6 m)	$R_2F_3V_1$	$R_2F_3V_2$	$R_2F_3V_3$
Far (6 m)	Near (0.7 m)	$R_3F_1V_1$	$R_3F_1V_2$	$R_3F_1V_3$
	Medium (2 m)	$R_3F_2V_1$	$R_3F_2V_2$	$R_3F_2V_3$
	Far (6)	$R_3F_3V_1$	$R_3F_3V_2$	$R_3F_3V_3$

4.1.1 Task completion and accuracy

The t-tests revealed that there was a significant effect of context-switching on task completion at the far distance of 6 meters, t (23) = 5.83, $p < 0.001$, one-tailed. Participants completed a greater number of subtasks/tasks when information was presented exclusively in the real-world control conditions, $M = 3.93$ ($SD = 0.67$) as compared to the augmented reality treatment conditions, $M = 2.85$ ($SD = 0.90$). Also, there was a significant effect of context-switching on accuracy at the 6 meters distance, t (23) = 5.322, $p < 0.001$, one-tailed. When the task was performed exclusively in the context of the real world at a distance of 6 meters, $M = 3.41$ ($SD = 0.68$), participants had better accuracy as compared to accuracy for augmented reality conditions, $M = 2.27$ ($SD = 0.84$). In other words, context-switching had a negative impact on performance when information was presented at the far distance of 6 meters. Participants had better task completion and were also more accurate at control condition 1 consisting of two monitors as compared to the treatment condition when they had to context-switch between virtual and real-world information at the 6 meter distance. There was no significant effect of context-switching at near or medium distances on task completion or accuracy.

4.1.2 Fatigue rating

One-tailed, paired t-tests revealed that context-switching had a significant effect on fatigue rating for each of the three distances. Thus, there was a significant difference between cell A_1B_1 in Table 4 and $R_1F_1V_1$ in table 5, A_2B_2 and $R_2F_2V_2$, and A_3B_3 and $R_3F_3V_3$ for fatigue rating. Participants found task performance in the real-world context (control condition 1) less fatiguing for all distances as compared to the augmented reality context (treatment condition). Summary results for fatigue rating are presented in Table 6.

Table 6. t-tests on fatigue rating for context-switching hypothesis

Distance	*Mean*		*Standard Deviation*		*t value*	*p value*
	Real world context	Augmented world context	Real world context	Augmented world context		
Near	2.333	3.583	1.606	2.263	3.315	0.001
Medium	2.750	3.875	2.026	2.027	2.689	0.006
Far	2.333	3.333	1.809	1.880	2.477	0.011

4.2 Goal 2. Focus depth

The second goal of this research was to determine if there was any impact of focus depth on user task performance. It was hypothesized that focus depth would have a significant effect on the speed and accuracy of user task performance. This was tested using a 3x3x3 within subjects, analysis of variance (ANOVA) for each of the dependent measures at a significance level of 0.05. In addition to speed and accuracy measures, the ANOVA was also performed on subjective measures of fatigue.

4.2.1 Task completion and accuracy

Results revealed that there was no significant main effect of focus depth on either task completion, F (2, 46) = 0.1, p = 0.9084 or accuracy, F (2, 46) = 0.23, p = 0.7974. However, there was a strong main effect of distance to virtual text on task completion, F (2, 46) = 80.20, p < 0.001 and accuracy, F (2, 46) = 139.68, p <0.001. The interaction between distance to real-world text and distance to virtual text was significant for task completion, F (4, 92) = 2.64, p = 0.0384 and approached significance for accuracy, F (4,

92) = 1.99, p = 0.102. All other effects and interactions were non-significant. The ANOVA summary for task completion and accuracy is represented in Tables 7 and 8 respectively.

Table 7. ANOVA summary table for focus depth hypothesis on task completion

Source	*df*	*SS*	*MS*	*F value*	*p value*
Between subjects					
Subjects (S)	23	159.025	6.914		
Within subjects					
Distance to real-world text (R)	2	0.083	0.041	0.06	0.9459
RxS	46	34.49	0.749		
Distance to virtual text (V)	2	264.204	132.102	80.20	<0.001*
VxS	46	75.773	1.647		
Focus depth (F)	2	0.0812	0.041	0.1	0.9084
FxS	46	19.396	0.421		
RxV	4	3.948	0.987	2.65	0.0384*
RxVxS	92	34.324	0.373		
RxF	4	1.453	0.363	1.36	0.2547
RxFxS	92	24.612	0.268		
VxF	4	0.944	0.236	0.91	0.4605
VxFxS	92	23.802	0.259		
RxVxF	8	0.560	0.070	0.43	0.8995
RxVxFxS	184	29.694	0.259		
Total	647	672.389			

Table 8. ANOVA summary table for focus depth hypothesis on accuracy

Source	*df*	*SS*	*MS*	*F value*	*p value*
Between subjects					
Subjects (S)	23	135.238	5.879		
Within subjects					
Distance to real-world text (R)	2	1.192	0.5961	1.04	0.3609
RxS	46	26.316	0.572		
Distance to virtual text (V)	2	333.247	166.623	139.68	<0.001*
VxS	46	54.873	1.193		
Focus depth (F)	2	0.183	0.091	0.23	0.7974
FxS	46	18.506	0.402		
RxV	4	3.149	0.787	1.99	0.1019
RxVxS	92	36.327	0.395		
RxF	4	1.657	0.414	1.26	0.2899
RxFxS	92	30.166	0.328		
VxF	4	1.152	0.288	0.82	0.5170
VxFxS	92	32.394	0.352		
RxVxF	8	0.698	0.0872	0.36	0.9411
RxVxFxS	184	44.825	0.243		
Total	647	719.923			

The significant distance to virtual text factor was analyzed using Tukey's Honestly Significant Difference Test at a significance of 0.05. Tukey's Test revealed that task completion for virtual text at 6 meters was significantly poorer than task completion at the near or medium distances. Also, accuracy at a virtual text distance of 6 meters was significantly poorer than accuracy at near or medium distances. Table 9 depicts the means and standard deviations for task completion and accuracy for the three levels of virtual text distance. There was no significant difference between the near and medium distances of virtual text.

Table 9. Means and standard deviation for distance to virtual text

Level	*Task completion*	*Accuracy*
Near	4.000 (0.658)	3.623 (0.684)
Medium	4.074 (0.652)	3.693 (0.645)
Far	2.684 (1.019)	2.141 (0.956)

The interaction between distance to real-world text and distance to virtual text for task completion is represented in Figure 15. From this figure it can be inferred that the interaction between distance to virtual text and distance to real-world text was very weak. As seen in the graph, both the near and medium virtual text distances were nearly identical in the number of tasks completed for the three real-world text distances. The difference between these two distances (i.e., near and medium) and the far virtual text distance decreased as the distance to real-world text increased from near to far. This is indicated by the non-parallel far and near/medium virtual text lines. In other words, when distance to virtual text was at the far level and distance to real-world text increased from near to far level, there was a slight increase in the number of tasks completed. However, when distance to virtual text was at the near or medium levels and distance to real-world text increased from near to far levels, there was a slight decrease in the number of tasks completed.

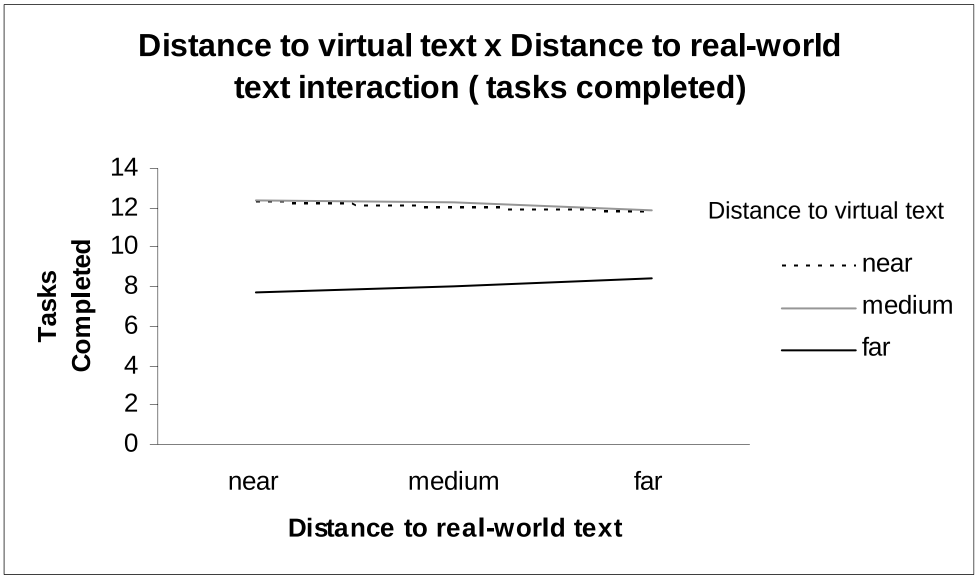

Figure 15. Interaction effect (task completion)

From Figure 16 we observed that the interaction between distance to virtual text and distance to real-world text for accuracy was very similar to the interaction for task completion. Here too, the near and medium virtual text distances were almost identical in accuracy, as indicated by the parallel lines. The difference between the near/medium virtual text distance and the far virtual text distance decreased as the distance to real-world text increased from near to far.

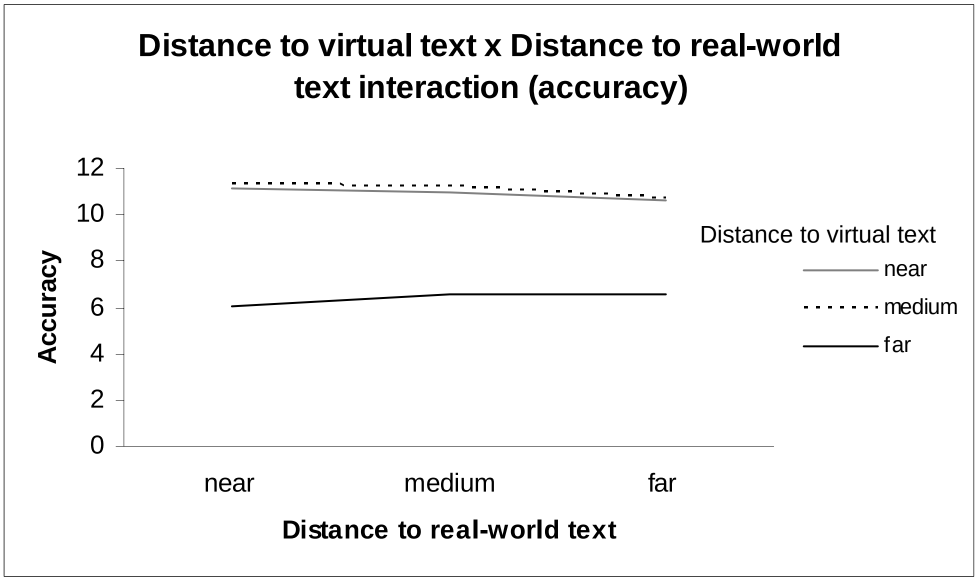

Figure 16. Interaction effect (accuracy)

4.2.2 Fatigue rating

The analysis of variance was also conducted on fatigue rating. Table 10 shows the ANOVA summary for fatigue rating. Similar to findings from objective measures of speed and accuracy, it was found that distance to virtual text had a significant effect on fatigue, F (2, 46) = 0.29, p = 0.005. All other effects and interactions were non-significant.

Table 10. ANOVA summary table for focus depth hypothesis on fatigue rating

Source	*df*	*SS*	*MS*	*F value*	*p value*
Between subjects					
Subjects (S)	23	1488.641	64.723		
Within subjects					
Distance to real-world text (R)	2	3.567	1.783	0.29	0.7473
RxS	46	279.913	6.085		
Distance to virtual text (V)	2	10.132	5.066	5.96	0.005*
VxS	46	39.126	0.851		
Focus depth (F)	2	0.595	0.297	0.28	0.7588
FxS	46	49.330	1.072		
RxV	4	1.061	0.265	0.46	0.7616
RxVxS	92	52.568	0.571		
RxF	4	1.932	0.483	0.26	0.8997
RxFxS	92	167.698	1.822		
VxF	4	3.951	0.988	1.23	0.3038
VxFxS	92	73.901	0.803		
RxVxF	8	3.021	0.377	0.49	0.8633
RxVxFxS	184	142.238	0.773		
Total	647	2317.674			

Tukey's Honestly Significant Difference Test on the significant distance to virtual text factor revealed that the far distance of 6 meters was rated more fatiguing than the medium distance. However, there was no significant difference between the near and far conditions or the near and medium conditions on fatigue rating. The means and standard deviations for the three levels are depicted in Table 11.

Table 11. Means and standard deviation for fatigue rating

Level	*Fatigue rating*
Near	3.611 (1.872)
Medium	3.745 (1.935)
Far	3.439 (1.867)

4.3 Goal 3. Matched vs. mismatched distances

The third goal of this thesis was to examine if there is any difference on user task performance when distance to virtual text, focus depth, and distance to real-world text are at the same distance from the user (matched conditions) as compared to conditions when they are at different distances (mismatched conditions). According to our hypothesis, performance would be better when distance to real-world text, distance to virtual text, and focus depth are at the same distance from the user as compared to when they are at different distances. To test the hypothesis, the mean of the three control condition 2 cells for task completion and accuracy depicted in Table 5 ($R_1F_1V_1$, $R_2F_2V_2$, and $R_3F_3V_3$) which represent matched conditions, was compared with the mean of the remaining mismatched 24 treatment cells in Table 5, using a paired, one-tailed, t-test at an alpha level of 0.05.

4.3.1 Task completion and accuracy

Results of the paired t-test revealed a significant effect of matched vs. mismatched condition on task completion, t (23) = 2.02, p = 0.028, one tailed. When distance to real-world text, distance to virtual text, and focus depth were matched, M = 3.72 (SD = 0.54), participants completed more tasks than conditions where the independent variables were at mismatched distances, M = 3.57 (SD = 0.52).

Paired t-test on accuracy also resulted in a significant effect of matched vs. mismatched distances, t (23) = 1.783, p = 0.043, one tailed. When distance to real-world text, distance to virtual text, and focus depth were at the same distance from the user, M = 3.29

(SD = 0.57), participants had higher accuracy than in the mismatched conditions, M = 3.14 (SD = 0.47).

Further analyses were also conducted to compare matched and mismatched conditions by excluding the far level for distance to virtual text. Since the ANOVA revealed a hugely significant effect of distance to virtual text, we were interested in determining if the above-mentioned difference between matched and mismatched conditions would still hold true if we excluded distance to virtual text from analysis. Thus, the mean of cells $R_1F_1V_1$ and $R_2V_2F_2$, which are blue in Table 12 and represent matched near and medium conditions, were compared with the 16 yellow cells which represent mismatched conditions. The white cells representing virtual text presented at far distances were excluded from this analysis.

Table 12. Data matrix for matched vs. mismatched distance hypothesis

Distance to real-world text (R)	*Focus depth (F)*	*Distance to virtual text (V)*		
		Near (0.7 m)	Medium (2 m)	Far (6 m)
Near (0.7 m)	Near (0.7 m)	$R_1F_1V_1$	$R_1F_1V_2$	$R_1F_1V_3$
	Medium (2 m)	$R_1F_2V_1$	$R_1F_2V_2$	$R_1F_2V_3$
	Far (6 m)	$R_1F_3V_1$	$R_1F_3V_2$	$R_1F_3V_3$
Medium (2 m)	Near (0.7 m)	$R_2F_1V_1$	$R_2F_1V_2$	$R_2F_1V_3$
	Medium (2 m)	$R_2F_2V_1$	$R_2F_2V_2$	$R_2F_2V_3$
	Far (6 m)	$R_2F_3V_1$	$R_2F_3V_2$	$R_2F_3V_3$
Far (6 m)	Near (0.7 m)	$R_3F_1V_1$	$R_3F_1V_2$	$R_3F_1V_3$
	Medium (2 m)	$R_3F_2V_1$	$R_3F_2V_2$	$R_3F_2V_3$
	Far (6)	$R_3F_3V_1$	$R_3F_3V_2$	$R_3F_3V_3$

The paired t-test revealed a significant effect of matched vs. mismatched conditions on task completion, t (23) = 1.730, p = 0.048. Participants completed a greater number of tasks in the matched condition, M = 4.15 (SD = 0.53) than in the mismatched condition, M = 4.023 (SD = 0.53). However, matched vs. mismatched conditions had less of an effect on accuracy and was only close to significance, t (23) = 1.79, p = 0.076.

4.3.2 Fatigue rating

Matched vs. mismatched distances had no significant effect on ratings of fatigue. Participants rated both conditions as being equally fatiguing, t (23) = 0.012, p = 0.9905.

4.4 Correlation analyses

Correlations between trial and task completion, accuracy, and fatigue were conducted using Pearson's coefficient at a significance of 0.05. A weak but significant positive correlation between trial and fatigue, r (23) = 0.266, $p < 0.001$, two-tailed was found. Thus as the experiment progressed, participants rated eye fatigue as being slightly higher. There was a small positive correlation between trial and task completion, r (23) = 0.095, $p = 0.015$, two-tailed. Thus, participants showed a very weak practice effect and completed more subtasks/tasks as the trials progressed. There was no correlation between trial and accuracy, r (23) = 0.056, $p = 0.1509$, two-tailed.

4.5 Summary of hypothesis testing results

This section summarizes the results of hypothesis testing presented in sections 4.1- 4.4.

Hypothesis 1: Context-switching will have a significant affect on task performance.

1. Context-switching had a significant effect on task completion at the far distance of 6 meters. Task completion was lower in the augmented condition as compared to the real-world condition at the far distance of 6 meters.
2. Context-switching had a significant effect on accuracy of tasks at the far distance of 6 meters. Accuracy was lower in the augmented condition as compared to the real-world condition at the far distance of 6 meters.
3. Context-switching had no effect on task completion or accuracy at the near or medium distances.

4. Context-switching had a significant effect on fatigue at all three distances. Context-switching in the augmented conditions resulted in higher ratings of fatigue as compared to real-world conditions when context-switching was absent.

Hypothesis 2: Focus depth will significantly affect task performance

1. Focus depth had no significant effect on task completion or accuracy.
2. Distance to virtual text had a significant effect on task completion at the far distance of 6 meters as compared to near or medium distances. There was no difference between the near and medium levels of distance to virtual text on task completion.
3. Distance to virtual text had a significant effect on accuracy of tasks at the far distance of 6 meters as compared to the near and medium distances. There was no difference between the near and medium levels of distance to virtual text on accuracy.
4. There was no effect of distance to real-world text on task completion or accuracy.
5. There was a significant, but weak, interaction effect on task completion between distance to real-world text and distance to virtual text.
6. Distance to virtual text had a significant impact on fatigue rating with the far distance of 6 meters being rated worse than the medium distance. There was no significant difference in fatigue rating between the near and medium distances or the near and far distances of virtual text.

Hypothesis 3: Matched conditions will have better performance than mismatched conditions

1. Matched vs. mismatched conditions had a significant effect on task completion. Matched conditions resulted in higher task completion as compared to mismatched conditions.
2. Matched vs. mismatched conditions also had a significant effect on task accuracy. Matched conditions resulted in higher accuracy as compared to mismatched conditions.
3. Excluding the far level of distance to virtual text from analysis resulted in a significant effect of matched vs. mismatched conditions on task completion (considering near and medium distances for virtual text only). Matched conditions had higher task completion as compared to mismatched conditions.
4. Excluding the far level for distance to virtual text from analysis resulted in a close to significant effect of matched vs. mismatched conditions on accuracy of tasks (considering near and medium distances for virtual text only). Matched conditions had higher accuracy as compared to mismatched conditions.
5. Matched vs. mismatched conditions had no impact on fatigue rating.

Additional results:

1. There was a weak but significant positive correlation between trial and fatigue
2. There was a weak but significant positive correlation between trial and task completion.
3. There was no correlation between trial and accuracy.

4.6 Post-experiment questionnaire and interviews

Feedback was gathered from participants through a post-experiment questionnaire and informal interviews after completion of the experiment. Participants were uncomfortable wearing the display and 87.5% of the participants complained of some form of eye discomfort including blurry vision, eye strain and fatigue, dizziness, and scratchy eyes. They felt that their left eye which wore the monocular display got more fatigued than their right eye. Participants also complained of physical discomfort such as headaches from wearing the head-mounted display and 41.6% of participants found the display too heavy to be worn on the head for extended periods of time. Other problems included discomfort due to the display lens touching the nose, neck ache, difficulty adjusting the display to fit properly on the head, and awkwardness while using the display.

Participants also complained of problems focusing on the far distance virtual text. They experienced the far distance text getting blurred after sometime and had a lot of trouble reading the far distance text. While all participants could see the far distance text properly in the beginning, they perceived it getting blurry sometime after the start of the experiment. Three participants said that they struggled to even discern the different letters. All participants did consider the far distance to virtual text as more fatiguing to see than the near or medium text distances.

When asked about preference for the virtual text distance, 17.5% preferred the near distance of 0.7 meters, 70% preferred the medium distance of 2 meters, and none liked the far distance of 6 meters. Further, 12.5% of participants were undecided between medium and near distances. Participants also noted difficulties in performing the task when the real-world text distance was mismatched from the virtual text distance. They said that the mismatched distances slowed them down as they had to spend additional time switching between the two worlds.

5. DISCUSSION

This section provides interpretations of the results presented in section 4. Based on these interpretations, conclusions and recommendations have been provided.

5.1 Interpretation of results

5.1.1 Goal 1. Context-switching

The first goal of this thesis was to determine if context-switching between real-world information and virtual information impacts user task performance. This impact was assessed by comparing performance in a real-world environment with an augmented environment, using dependent measures of speed and accuracy. In addition to objective measures of speed and accuracy, subjective ratings of fatigue on a bipolar seven point scale were also evaluated. For validity of the comparison, augmented reality conditions were similar to the real-world conditions (e.g. near-matched augmented condition was compared with the near real-world condition, medium-matched augmented condition was compared with medium real-world condition, and far-matched augmented condition was compared with far real-world condition).

The test on hypothesis 1 revealed that the far level of distance to virtual text had a significant impact on task completion and accuracy as a consequence of context-switching between virtual and real-world information. Thus participants performed better when the task was in an exclusively real-world environment at 6 meters as compared performing the task in an augmented environment at 6 meters. Participants completed fewer tasks and were also less accurate when virtual information was presented at the far distance. Analysis of fatigue ratings supported findings from objective measures. Participants found the 6 meter distance to virtual text more fatiguing than the near or medium distances. There was no difference in task performance between real world and augmented conditions when information was displayed at 0.7 meter or 2 meters distance.

Comments from participants after completion of the experiment revealed that the far distance virtual text got blurry after completion of a part of the experiment. However, this blurring effect did not occur in the beginning of the experiment but sometime later during the course of the experiment. This anecdotal evidence was supported by the positive correlation between trial number and fatigue wherein participants rated eye fatigue as higher as the experiment progressed and by the main effect of the distance to virtual text factor from the analysis of variance.

It is conjectured that the blurry vision was the result of repeatedly changing eye focus to accommodate to the far distance of 6 meters. As discussed in the literature review in section 2.5, the eye has a natural resting point of accommodation which is the distance to which the eyes are adjusted when there is nothing on which to focus and this resting point represents the most relaxed and natural eye focus distance. Similar to the resting point of accommodation is the resting point of convergence, which represents the distance at which the eyes naturally converge when there is nothing to converge to. Resting point of accommodation is said to be about an arm’s length (Tufano, 1987). The eye lapses back to its resting point of accommodation and convergence in the absence of effective stimuli. The 6 meters distance was very different from the resting point of accommodation and convergence and hence may have resulted in greatest strain of the eye's oculomotor mechanism. In the beginning of the experiment, participants' eyes were rested and they did not feel the strain from accommodation and convergence; however as the experiment progressed the repeated need to accommodate and converge added to eye fatigue resulting in blurry vision for the far level of distance to virtual text.

Another factor that may have caused the blurring effect was resolution of the display. The head-mounted display supports a resolution of 800x600. This resolution is sufficient to easily read text at closer distances; however at far distances such as 6 meters it resulted in poor legibility, despite our best efforts to make the text large enough to be easily readable at all distances. The text size at 6 meters, although smaller than the text size at near and medium distances, was well above the standard of 22 arc minutes for legibility (Human Factors Society, 1984) and hence was not a cause by itself for the degradation in performance. However, the two above-mentioned factors (poor resolution and smaller text size), along with the far virtual text distance that necessitated a greater need for accommodation and convergence, may have jointly exacerbated eye fatigue as the trials progressed, resulting in blurry vision, fatigue, and hence deteriorated performance.

There was no difference in performance between augmented conditions and real-world conditions when text was displayed at near and medium conditions. Thus context-switching between real-world and virtual information in the augmented environment at distances of 0.7 meter and 2 meters did not deteriorate performance or cause eye fatigue. At distances closer to the user, the text size was large and hence the resolution of 800x600 did not have an impact on performance. Moreover, since the text was closer to the resting point of accommodation and convergence, the eye underwent minimal focus adjustments.

Context-switching had a significant effect on ratings of fatigue at all three distances. Thus participants rated the augmented condition as more fatiguing than the real-world condition. We see that objective measures of performance were not in agreement with subjective measures. Thus, for speed and accuracy measures of performance, context-switching had an impact *only* at the far distance, while for fatigue rating it had an impact on *all* three distances. It is hypothesized that the greater level of fatigue for the augmented condition was due to the need to switch repeatedly between virtual and real-world text as well the relatively poor resolution of virtual text which does not duplicate real-world text.

5.1.2 Goal 2. Focus depth

The second goal of this thesis was to study the impact of focus depth, which is a hardware parameter of the display, on user task performance. Focus depth was not found to impact user task performance or subjective eye fatigue ratings.

The lack of effect of focus depth on user task performance was unexpected. As mentioned earlier, focus depth controls the user's view of the augmented world. Based on limited information on the topic of focus depth, we know that focus depth causes the eye to focus at certain distances. As a result, if there is a mismatch between focus depth and stimuli distance, it was expected that the stimuli (real-world and virtual text) would appear blurred and out-of-focus as the eye would focus at the distance corresponding to focus depth and would not accommodate to the stimuli distance. We predicted that for conditions when focus depth was at different distances than the virtual text or real-world text, a conflict between the stimuli distance and the position where the eye focused would cause a negative impact on task performance. Thus, the lack of effect of focus depth is contrary to our hypothesis.

Comments from participants about the manipulation of focus depth to different levels revealed that changing the focus depth setting caused the augmented view to move closer or farther away from them. However, besides the change in perceived distance, participants could not identify any effect of changing focus depth on their view of the augmented environment. Manipulation of focus depth to distances different from the stimuli distances did not prevent participants from accommodating and converging their eyes to see the stimuli clearly. Thus we infer that focus depth did not cause participants to focus their eyes at a given distance.

We speculated that the lack of effect of focus depth on task performance may be in part due to the task scenario. The user’s view consisted of two very similar pieces of real world and virtual information with no other objects to focus at or to distract the user. Thus the stimuli were very strong and focus depth was unable to cause the eye to focus at its distance. This represents an unrealistic situation as in reality the environment is complex and cluttered and users often need to divide attention between different stimuli. It is possible that in the absence of a single strong stimulus focus depth may have a greater effect on performance. Thus further research is required to study focus depth using realistic scenarios in a complex and dynamic environment.

5.1.3 Goal 3. Matched vs. mismatched distances

The final goal was to determine if mismatched distances for real-world text, focus depth, and virtual text negatively impact user task performance as compared to matched distances. To assess the associated hypothesis, a comparison was made between matched and mismatched conditions.

We found that task completion was lower for mismatched conditions. Also, participants had lower accuracy when completing tasks that belonged to mismatched conditions as compared to matched conditions. A further comparison between matched and mismatched conditions for only near and medium distances (excluding far distance) also produced similar results. Participants had poorer task completion as well as accuracy for mismatched conditions. Matched vs. mismatched conditions had no effect on fatigue.

From the discussion presented in section 5.1.2, we know that focus depth did not have an impact on task performance. Thus, we are able to conclude that the mismatched conditions had poorer performance than matched conditions due to a mismatch in distances of real-world text and virtual text and not because focus depth was mismatched.

When participants performed tasks in conditions where distance to virtual text and distance to real-world text were at the same distance, minimal effort and time were spent in accommodating and converging. On the other hand, when virtual text and real-world text were at different distances, participants had to refocus their eyes repeatedly to different distances to perform the task. This may have slowed participants down, resulting in lower task completion. Accuracy was also affected due to the need to accommodate and converge the eye to different depths.

During post-experiment interviews, participants stated that switching between real-world text and virtual text when distances were mismatched was difficult for them. They felt that mismatched distances slowed them down and they took additional time making the switch between real-world text and virtual text.

The analysis of variance for the focus depth hypothesis revealed an interaction effect between distance to real-world text and distance to virtual text. This interaction supports the matched vs. mismatched distance hypothesis. Participants had higher task completion when virtual text was at the same distance as real-world text from the user. When both virtual text and real-world text were presented at the far level, participants completed a greater number of tasks as compared to when virtual text was at the far level and real-world text at near or medium levels. Further, when virtual text was at near or medium level, completion was better when real-world text was also at the near or medium distance as compared to far level. Thus from the interaction we observe that performance was better when real-world text distance was matched to the virtual text distance.

5.1.4 Additional observations

Although from the ANOVA we found that focus depth had no significant impact on performance, we observed a significant effect of distance to virtual text on speed and accuracy of user performance. Further analysis showed that the far distance of 6 meters was significantly worse than the near or medium distances. Thus participants were able to complete fewer tasks and were less accurate when virtual text was displayed at 6 meters. The closer distance of 0.7 meters and 2 meters for virtual text did not have an impact on task performance.

Reasons behind the poor performance at the 6 meter distance are the same as discussed earlier for context-switching. The combined effect of relatively poor display resolution, smaller text size as compared to the near and medium distances, and excessive need to accommodate may have been the reasons for blurry vision and thus poor performance at the 6 meter virtual text distance.

Ratings of fatigue showed that the far distance of 6 meters was perceived as more fatiguing than the medium distance by participants. There was no difference between the near and medium distance to virtual text or the near and far distance to virtual text. This is not unexpected as participants preferred the medium virtual text distance and disliked the far virtual text distance the most. Thus, the fatigue ratings for distance to virtual text followed the pattern found through subjective ratings of the three virtual text distances during post-experiment interviews.

There was also no significant impact of distance to real-world text on performance. This reinforces the fact that the combined effect of resolution of the head-mounted display, size of text, and the need to repeatedly accommodate/converge the eye played a role in making the 6 meter distance to virtual text lower for the task and not just size of text and accommodation/convergence. Both the virtual text and real-world text had the same size of text and participants needed to accommodate/converge the eye by an equal amount. So while the real-world text did not result in poor performance, the virtual text at 6 meters may have resulted in lower accuracy and completion of tasks due to the combined effect of resolution of the display, text size, and the need to accommodate and converge.

5.2 Conclusions and recommendations

Based on the results of this study the following conclusions have been drawn.

1. Virtual text should not be displayed at a distance greater or equal to 6 meters. With the Nomad's display resolution of 800x600, 6 meters (which corresponds to optical infinity) may be too far to ensure good legibility.
2. Performance measures indicate that there was no difference in task performance between the near (0.7 meter) and medium (2 meters) virtual text distance. However, 70% preferred the medium virtual text distance of 2 meters. Despite this, we cannot recommend displaying virtual text at 2 meters as another finding of this thesis was that virtual information should be displayed as close as possible to the real-world object of interest. So virtual information should be positioned depending upon the real-world object distance.

3. Context-switching between virtual and real-world information causes greater eye fatigue. Efforts should be made to minimize eye fatigue by developing displays with higher resolution. This is likely to relieve the eye of some fatigue.
4. Since focus depth does not have an impact on user task performance or fatigue, an initial recommendation is that head-mounted displays should be manufactured as fixed focus depth displays. Fixed focus depth displays are cheaper and easier to manufacture. However, this finding merits additional intensive research before such a recommendation can be definitively made.
5. Virtual text should be displayed at a distance that is as close as possible to the real-world object of interest. This will put the least amount of strain on the eye's oculomotor mechanism as it will minimize the need to accommodate and converge.
6. As mentioned earlier, the Nomad resolution of 800x600 may be unsuitable for displaying text at far distances. Hence displays with higher resolution may be required if text is to be displayed at far distances such as in outdoor augmented reality where real-world objects of interest may be far from the user.
7. The ergonomics of head-mounted displays needs to be improved. The weight of 226.8 grams for head-mounted displays is too heavy to be worn for long periods of time. In applications such as warfighting and maintenance and repair, the additional weight of the head mounted display can pose serious problems for the user. Also, head-mounted displays do not fit properly for all people and result in physical discomfort and fatigue. Better adjustments are required so that the head-mounted display can accommodate different head sizes.

5.3 Limitations and future research

This research was a first step to the development of human-centered augmented reality systems. An overarching aim of this thesis was to understand the impact of augmented reality on human perception and cognition. It attempted to determine the optimal distance of virtual text from the user in augmented reality, study the impact of focus depth, and comprehend the effect of context-switching on user task performance and fatigue.

This study used a visual scanning task consisting of random character strings. The nature of the task, though carefully designed to study the research hypotheses, was not particularly realistic. Future studies should employ tasks with greater ecological validity. They should be based on real-world tasks that a user is likely to perform with the aid of augmented reality. Verification of results of this thesis using practical task scenarios is required.

Further, in this study, real-world information and virtual information were very similar (red text on a black background). Such homogeneity of conditions is unlikely in reality. The real world consists of complex scenes with different colors, shapes, and lighting conditions. Complex and dynamic real-world backgrounds are likely to aggravate problems in context-switching. Users may suffer from attentional tunneling where they concentrate only on one world due to its dominance in stimulus conditions. This would cause users to neglect information in the non-dominant world and render augmented reality useless. Thus, future research should explore the questions of this thesis in more realistic (i.e., complex and dynamic) settings.

This study found no effect of focus depth on task performance. However, this may have been due to the task scenario which consisted of only two pieces of information (real-world text and virtual text) as very strong as stimuli. Thus the eyes did not focus at the distance corresponding to focus depth but were able to accommodate and converge to the level of the stimuli distance. In reality, stimuli may not be so strong visually and we may be able to see an impact of focus depth in such situations. Future studies should investigate the impact of focus depth in cluttered and complex environments. Also, studies should look at the impact of focus depth at distances closer than 0.7 meters as accommodation and convergence is a powerful cue at distances closer to the user.

This study was limited to three distances of 0.7 meter, 2 meters, and 6 meters. While we were able to establish that 6 meters distance is not optimal for displaying virtual text, there may be distances besides 0.7 meter and 2 meters that result in better task performance. Hence future studies should include a smaller step interval for testing the various distances. Also, this study used relative size of text as a depth cue. Thus text at the 0.7 meter was larger than text at 6 meters. This might have played a role in poor performance at the 6 meter distance. Future research should study performance at different distances using a constant text size. Thus, irrespective of distance, text sizes should appear to be the same size from the user's perspective.

An interesting extension of this study pertains to outdoor augmented reality. In outdoor augmented reality, real-world information tends to be more scattered and at greater distances from the user as compared to indoor augmented reality. Moreover, the outdoor environment is complex and dynamic. The question of optimal distance for virtual text is even more critical for the outdoor realm. Should virtual information still be displayed at the same distance as the real-world objects of interest which may be very far from the user? Such questions need to be answered in future research.

A finding of this research was that virtual information should be presented as close as possible to the real-world object of interest. However, in a cluttered environment where there are several real-world objects of interest situated at different depths from the user, this recommendation cannot be used. Thus further research is required to solve this problem.

Results of this research had two conflicting findings. On one hand we found that virtual information displayed at optical infinity is not suitable for augmented reality and thus virtual information should be displayed closer to the user. However, based on results of this study, we also found that performance is better when virtual information and real-world information are at the same distance from the user. This presents an interesting problem if real-world information is presented several meters from the user. In such cases, should virtual information be displayed at that far-off distance (optical infinity) to match the real-world object or should virtual information be displayed closer to the user and thus result in mismatched distances for real-world information and virtual information? Further studies are required to explore and resolve this issue.

5.4 Summary

The first motivation for this study was to determine if user task performance deteriorates in an augmented environment as compared to a real-world environment. Based on the empirical study, we found that task performance deteriorates in an augmented environment only if virtual text is displayed at a distance as far as optical infinity. We found no difference between real world and augmented world task performance at distances of 0.7 meter and 2 meters. Thus, optical infinity may be unsuitable for displaying virtual text at the current 800x600 resolution of head-mounted displays such as the Nomad.

Secondly, we attempted to assess the impact of focus depth on user task performance. We found that focus depth had absolutely no impact on task performance or fatigue and hence preliminarily recommend that head-mounted display manufacturers may need to make only displays with fixed focus depth which are not only less expensive but also simpler to manufacture. However, more research is required to further study focus depth using realistic tasks in a complex environment.

We also found that when virtual text and real-world text are at different distance from the user, performance is lower and fatigue is greater than in conditions where real-world text and virtual text are at the same distance from the user. Thus we recommend that virtual information and real-world information be displayed at the same distance from the user. Lastly, participants found the head-mounted display to be too heavy for use during extended periods of time and thus the weight, fit, and other ergonomic issues of these displays needs to be improved.

This research is one of several starting points to advance the development of usable augmented reality systems. In order to develop usable augmented reality systems, it is very essential to incorporate the user into the system and study the impact of augmented reality on human perception and cognition. This research should be considered one of the first steps in stimulating numerous future human-centered research studies in the field of augmented reality.

6. REFERENCES

Andre, J.T., & Owens, D.A. (1999). Predicting optimal accommodative performance from measurement of the dark focus of accommodation. *Human Factors, 41*, 139-145.

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

Azuma, R., Balliot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent Advances in augmented reality. *IEEE Computer and Graphics, 21*, 34-47.

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

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

Caudell, T.P. (1994). Introduction to augmented and virtual reality. *Telemanipulator and Telepresence Technologies, 2351*, 272-281.

Charman, W.N. (2000). Accommodation and the through-focus changes in the retinal image. In O. Franzen, H. Richter, & L. Stark (Eds.), *Accommodation and vergence responses*. Switzerland: Berkhäuser Verlag.

Cohen, J. (1977). *Statistical power for the behavioral sciences* (Rev. ed.). Florida: Academic Press.

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

Cutting, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, and Computers*, *29,* 27-36.

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

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

Davson, H. (1990). *Davson's physiology of the eye* (5th ed.). New York: Pergamon Press.

Drasic, D. (1991). Skill acquisition and task performance in teleoperation using monoscopic and stereoscopic video remote viewing. *Proceedings of the Human Factors and Ergonomics Society*, 1367 - 1371.

Edgar, G.K., Pope, J.C., & Craig, I.R. (1994). Visual accommodation problems with head-up and helmet mounted displays. *Displays, 15*, 68-75.

Ellis, S. R. (1998). Localization of virtual objects in the near visual field. *Human Factors, 40,* 415-433.

Feinerberg, E.B. (2004). Visual acuity test. Retrieved January 24, 2004, from http://www.nlm.nih.gov/medlineplus/ency/article/003396.htm.

Fukuda, T., Kanada, K., & Saito, S. (1990). An ergonomic evaluation of lens accommodation related to visual circumstances. *Ergonomics*, *33*, 811-831.

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

Feiner, S., MacIntyre, B., & Seligmann, D. (1993). Knowledge based augmented reality. *Communications of the ACM, 36*, 52-62.

Flannagan, M.J., Sivak, M., & Simpson, J.K. (2001). The relative importance of pictorial and nonpictorial distance cues for driver vision. *Proceedings of the International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Aspen, CO,* 214-218.

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

Gabbard, J.L. (2003). *Real-world Scenes, Objects, and Textures as a GUI Background: Designing Effective Text Styles for Outdoor Augmented Reality*. Unpublished doctoral dissertation proposal, Virginia Tech, Virginia.

Glumm, M.A., Kilduff, P.W., & Masley, A.S. (1992). *A study of the effects of lens focus depth on remote driver performance*. Ohio: Army Research Laboratory.

Goss, D.A., & West, R.A. (2002). *Introduction to optics of the eye*. Massachusetts: Butterworth Heinemann.

Helander, M.G. (1998). Human factors opportunities in virtual and augmented environments. *Keynote Address - Global Ergonomics Conference, Cape Town, South Africa.*

Human Factors Society. (1984). *American national standard for human factors engineering of visual display terminal workstations* (ANSI/HFS 100-88), Santa Monica, CA.

Iavecchia, J.H., Iavecchia, H.P., & Illiana, R. (1988). Eye accommodation to head-up virtual images. *Human Factors, 30*, 689-702.

Ijsselsteijn, W.A., Ridder, H., & Hamburg, R. (1997). The effect of image disparity, convergence distance and focus depth on perceived quality in stereoscopic displays*. IPO Annual progress Report*, 32, 51-59.

Inuzuka, Y., Osumi, Y., & Shinkai, H. (1991). Visibility of head up displays (HUD) for automobiles. *Proceedings of the Human Factors and Ergonomics Society, 38*, 1574-1578.

Jaschinski-Kruza, W. (1988). Visual strain during VDU work: the effect of viewing distance and dark focus. *Ergonomics, 31,* 1449-1465.

Jaschinski-Kruza, W. (1990). On the preferred viewing distances to screen and document at VDU workplaces. *Ergonomics, 33*, 1055-1063.

Jaschinski-Kruza, W. (1991). Eyestrain in VDU users: viewing distance and the resting position of ocular muscles. *Human Factors, 34*, 69-83.

Jebaraj, D., Tyrrell, R.A., & Gramopadhye, A.K. (1999). Industrial inspection performance depends on both viewing distance and occulomotor characteristics. *Applied Ergonomics, 30*, 223-228.

Johnson, C.A. (1976). Effects of luminance and stimulus distance on accommodation and visual resolution, *Journal of the Optical Society of America*. *66*, 138-142.

Kalawsky, R.S., Stedmon, A.W., Hill, K., & Cook, C.A. (2000). A taxonomy of technology: Defining augmented reality. *Proceedings of IEA/HFES 2000 Congress*, 507-510.

Kato, H., Ito, H., Shima, J., Imaizumi, M., & Shibata, H. (1992). *Development of a hologram head up display*. Society of Automotive Engineers Technical Paper Series, Warrendale, PA.

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

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

Livingston, M.A., Rosenblum, L.J., Julier, S., Brown, D., Balliot, Y., Swan, E.S. et., al. (2002). An augmented reality system for military operations in urban terrain [Abstract]. *Proceedings of Interservice / Industry Training, Simulation & Education Conference, Orlando, FL.*

McGee, M.K. (1999). *Integral perception in augmented reality*. Unpublished doctoral dissertation, Virginia Tech, Blacksburg.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: A class of displays on the reality virtuality continuum. *SPIE: Telemanipulator and Telepresence, 2351*, 282-292.

Miller, R.J., Pigion, R.G., Wesner, M.F., & Patterson, J.G. (1983). Accommodation fatigue and dark focus: the effects of accommodation free visual work as by two psychophysical methods. *Perception and Psychophysics*, *34*, 532-540.

Moses, R.A., & Hart, W.M. (1987). *Adler's physiology of the eye* (8th ed.). St. Louis, MO: Mosby Publications.

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

Neveu, C., Blackmon, T., & Stark, L. (1998). Evaluation of the effects of a head-mounted display on ocular accommodation. *Presence, 7,* 278-289.

Norman, J., & Ehrlich, S. (1986). Visual accommodation and virtual image displays: target detection and recognition. *Human Factors*, *28*, 135-151.

Okabayashi, S., & Sakata, M. (1991). How an automotive head-up display affects driver's ability to recognize the forward view. *Proceedings of the Society for Information Display*, *32*, 31-37.

Östberg, O. (1980). Accommodation and visual fatigue in display work. In E. Grandjean & and E. Vigliani (Eds.), *Ergonomic Aspects of visual display terminals* (pp 41-52). London: Taylor and Francis.

Owens, D.A., & Wolf-Kelly, K. (1987). Near work, visual fatigue, and variations of occulomotor tonus. *Investigative Ophthalmology and Visual Science, 57,* 540-550.

Pasman, W., & Jansen, F.W. (2001). Distributed low latency rendering for mobile AR. *Proceedings of IEEE and ACM International Symposium on Augmented Reality, New York*, 107-113.

Raymond, J.E. (1986). Viewing distance affects reaction time to discriminate letters of a constant small size. *Perception and Psychophysics, 40,* 281-286.

Rekimoto, J. (1997). NaviCam: A magnifying glass approach to augmented reality. *Presence, 6*, 399-412.

Richmond Products. (n.d.). Retrieved March 2, 2004, from, http://www.richmondproducts.com/Stereo%20test%20Index.htm

Roscoe, S.N. (1984). Judgments of size and distance with imaging displays, *Human Factors*. *26*, 617-629.

Roscoe, S.N. (1985). Training good eyes to see better. *Displays, 6,* 218-220.

Satoh, K., Anabuki, M., Yamamoto, H., & Tamura, H. (2001). A hybrid registration method for outdoor augmented reality. *Proceedings of the IEEE and ACM International Symposium on Augmented Reality, New York,* 67-76.

Simonelli, N.M. (1979). Apparent size and visual accommodation under day and night conditions. *Proceedings of the Human Factors and Ergonomics Society Conference, TX, 23,* 374-378.

Stedmon, A.W. (2001). Re-viewing reality: human factors of synthetic training environments. *International Journal of Human Computer Interaction, 55,* 675-698.

Stedmon, A.W., Hill, K., Kalawsky, R.S., & Cook, C.A. (1999). Old theories, new technologies: comprehension and retention issues in augmented reality systems. *Proceedings of the Human Factors and Ergonomics Society TX, 43*, 1358-1362.

Stedmon, A.W., & Stone, R.J. (2001). Re-viewing reality: human factors of synthetic training environment. *International journal of Human Computer Studies, 55*, 675-698.

Sugihara, T., & Miyasoto, T. (1999). An adaptive brightness control for an augmented reality display. *Society of Information Display Technical Digest*, 812-815.

Thomas, B., Krul, N., Close, B., & Piekarski, W. (2002). Usability and playability issues for ARQuake. *Proceedings of the International Workshop on Entertainment Computing*, 455-462.

Tufano, D.R. (1997). Automotive HUDs: The overlooked safety issues. *Human Factors, 39*, 303-311.

Watt, W.S. (2003). How visual acuity is measured. Retrieved January 24, 2004, from http://www.mdsupport.org/library/acuity.html

Weintraub, D.J. (1984). The Utility of head-up displays: Eye focus vs. decision times. *Proceedings of the Human Factors and Ergonomic Society, 28*, 529-533.

Weintraub D.J., & Ensing, M.J. (1992). *Human factors issues in heads-up display design: the book of HUD*. Crew System Ergonomic Analysis Center.

Wolffsohn, J.S., Edgar, G.K., & McBrien, N.A. (1998). The effect of viewing a car head up display on ocular accommodation and response times. In A.G. Gale, I.D. Brown, C.M. Haselgrave, & S.P. Taylor, *Vision in Vehicles – VI* (pp.143-151).

Yeh, M., Wickens, C.D, & Seagull, J.F. (1998). Effects of frame of reference, viewing condition on attentional issues with helmet mounted displays. *Army Research Laboratory Technical Report Technical Report ARL-98-1/ARMY-FED-LAB-98-1*.

APPENDIX A. Snellen test

Purpose of Snellen test

A Snellen test is simply a way to evaluate visual acuity using a chart with different sizes of letters or forms. The test determines how accurately a person can see from a distance.

Procedure

Stand behind a line 20 feet from the eye chart. Cover one eye with the palm of your hand and read out loud starting from the topmost line of letters. Read the smallest line of letters that you see. If you are not sure of the letter, you may guess. Repeat this with the other eye (Feinerberg, 2004).

Snellen rating

A rating is assigned based on the results of the test. This is expressed as a fraction where the numerator is the distance from which the chart is being read (i.e. 20 feet) and the denominator represents the distance at which a normal eye can read the alphabet. Normal vision is denoted as 20/20 which means that from a distance of 20 feet a person can read what a normal eye can read from a distance of 20 feet. Similarly 20/40 denotes that from a distance of 20 feet a person can read what a normal eye can read at a distance of 40 feet (Watt, 2003).

APPENDIX B. Runge near point vision test

Purpose of Runge Near Point Card Test

The Runge Near Point Card Test is a test of near visual acuity. It consists of a card with letters of different sizes. The test determines how accurately a person can see near objects.

Testing procedure

The test is to be conducted in a well lit room. The card should be fixed at a distance of 16 inches from the participant. If the participant wears glasses or contacts for near vision (reading) then he/she should wear them while performing this test.

The participant should read the columns of letters on the card. The smallest column of letters correctly read by the participant is to be noted. Visual acuity corresponding to the column correctly read should be determined from the card. Normal vision for the Runge Near Point Card is 20/20.

Runge Near Point Card Rating

A rating is assigned based on the results of the test. This is expressed as a fraction where the numerator is the distance from which the chart is being read (i.e. 20 inches) and the denominator represents the distance at which a normal eye can read the alphabet. Normal vision is denoted as 20/20 which means that from a distance of 20 inches a person can read what a normal eye can read from a distance of 20 inches. Similarly 20/40 denotes that from a distance of 20 inches a person can read what a normal eye can read at a distance of 40 inches (Powell, L., personal communication, March 22, 2004)

APPENDIX C. Informed consent form

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
Informed Consent for Participants in Research Projects Involving Human Subjects

<u>Title of Study</u>**:** Empirical Study of the Effects of Context-Switch, Object Distance, and Focus depth on Human Performance in Augmented Reality

<u>Principal Investigators</u>**:** Divya Gupta and Deborah Hix

I. PURPOSE OF STUDY

This study involves research on augmented reality (AR) displays. An augmented reality system is one in which a user wears a see-through head-mounted display, and that display superimposes graphics and/or text on the real world that is visible through the display. We are evaluating the effect of AR display factors such as focal length, real world image distance, and augmenting graphics distance on human task performance, with the goal of making AR systems as effective and usable as possible. We are collecting research information to enhance basic scientific knowledge.

II. PROCEDURES

1. First, a test of far and near visual acuity will be administered to check for normal vision. Please ensure that you are wearing your lenses or glasses if you use them. You will be asked to stand at a certain distance and read alphabets displayed to you on the wall. Next, a card will be displayed and you will be asked to read alphabets printed on the card from a certain distance.
2. Then, you will complete a pre-experiment questionnaire to gather information regarding vision, age, depth perception etc.
3. We will introduce you to the equipment used in augmented reality systems and assist you in putting on the display. This display is a unique head-worn display called NOMAD made by Microvision.
4. You will be able to see the real world through the display. You will sit on a chair. The height of the chair will be adjusted so that you can see the monitor in your field of view. Please ensure that you are wearing your lenses or contacts if you use them. Several practice trials will be performed to familiarize you with the display and the task.

When you look through the display, you will see nonsense words on a monitor and virtual text hanging in the air. Your task is to first locate a pair of similar letters one of which is upper case and the other lower case (e.g. Aa, Ee, Gg) from the virtual text and then determine how many times that particular letter appears in the real world text on the monitor. You will respond by pressing the zero, one, two, or three number keys on the keyboard indicating that the letter is not present, present once, present twice, or present thrice in the real world text. This procedure will be repeated several times and you will be given 25 seconds to try and complete as many trials of it as possible. At the end of 25 seconds the screen will be blanked and you will be asked to rate the condition of your eyes according to the following scale by pressing the appropriate number on the keyboard. The experimenter will provide you with a description of the scale at that time. You will be asked to rest with your eyes closed for a period of 45 seconds after the rating. During this time we will be moving or adjusting some of the equipment. When the 15 second rest period is over a sound will alert you to open your eyes. **Please do not start with the next task until we till you to proceed.**

1	2	3	4	5	6	7
very rested	rested	somewhat rested	neither rested nor fatigued	somewhat fatigued	fatigued	very fatigued

5. You will repeat the same procedure with different nonsense words several times. You should work as quickly and accurately as possible; once you determine the number of times a letter appears in the text, you should press the corresponding key as soon as you can.
6. We expect the entire session to take about two hours. There will be plenty of breaks for you to rest.
7. When you have completed all tasks, we will ask you to fill out another questionnaire, to obtain your feedback on the experiment.

Please remember that we are evaluating the augmented reality display; we are not evaluating you in any way.

III. RISKS

The risks of this study are minimal and are not much more than using a computer. You might experience eyestrain which is not different from what one experiences when using a computer for long hours. If you experience any form of discomfort please inform the experimenters and they will assist you in removing the head mounted display.

IV. BENEFITS

As a participant, you will learn about state-of-the-art augmented reality systems, and you will have the knowledge that you are helping to make systems like this easier and more effective. You will also be told your far and near vision.

V. EXTENT OF ANONYMITY AND CONFIDENTIALITY

The data and all related information will be held in confidence and will not be associated with your name or identity outside the context of the study or in any published results. In particular, data will be indexed by number; at no time will your name be associated at any way with this number on any computer system. Only one hand-written list linking names and numbers will be kept, and this list will not be put in computerized format, nor will it be revealed to anybody outside the study. The data and related information will remain confidential even after the study has ended.

VI. COMPENSATION

You will be compensated for your participation at the rate of $10 per hour or will be given class credit for participation.

VII. FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time for any reason.

VIII. SUBJECT'S RESPONSIBILITES

I voluntarily agree to participate in this study. I have the responsibilities to abide by the rules of this project and will be truthful in answering all questions.

IX. SUBJECT'S PERMISSION

I voluntarily agree to participate in this study, and I know of no reason I cannot participate. I have read and understand the informed consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty.

______________________________ ________________

Signature Date

______________________________ ________________

Name (please print) Email

____

Contact phone

Should I have any pertinent questions about this research or its conduct, and research subjects' rights, and whom to contact in the event of a research-related injury to the subject, I may contact:

<u>Principal Investigator</u>
Divya Gupta
Email: divya@vt.edu
Tel: 540-953-5080

<u>Faculty Advisor</u>
Deborah Hix
Email: hix@cs.vt.edu
Tel: 540-231-6199

<u>Chair, Virginia Tech Institutional Review Board for the Protection of Human Subjects</u>
David M. Moore
Office of Research Compliance – CVM Phase II (0442)
Research Division
Email: moored@vt.edu
Tel: 540-231-4991

This Informed Consent is valid from March 2004 to March 2005.

APPENDIX D. Pre-experiment questionnaire

Augmented Reality Study Questionnaire

Please answer the following questions to the best of your knowledge by clearly **checking** the boxes. If you do not understand any question please feel free to ask the experimenter.

1. Gender:
   ☐ Female
   ☐ Male

2. Age:
   ______ Years

3. Please rate the condition of your eyes as they feel right now. Circle the appropriate response.

1	2	3	4	5	6	7
very rested	rested	somewhat rested	neither rested nor fatigued	somewhat fatigued	fatigued	very fatigued

4. Are you currently suffering from any eye diseases (e.g. conjunctivitis, sty etc.)?
   ☐ Yes
   ☐ No

5. Do you wear lenses or glasses?
   ☐ Yes
   ☐ No

6. Please indicate your corrected **far** vision (vision with glasses or lenses). If you do not wear lenses or glasses for far vision please write down your uncorrected far vision.
   ______ (e.g. 20/20).
   ☐ I don't know my far vision number

7. Please indicate your corrected **near** vision (vision with glasses or lenses). If you do not wear lenses or glasses for near vision please write down your uncorrected near vision.
   ______ (e.g. 14/14).
   ☐ I don't know my near vision number

8. Please rate your vision. If you wear lenses or glasses then rate your vision with respect to the lenses or glasses that you wear.
   ☐ Worse than normal vision
   ☐ Same as normal vision
   ☐ Better than normal vision

9. If you wear glasses or contacts are you wearing them now? Answer this question only if you wear lenses or contacts.
   - ☐ Yes
   - ☐ No

10. Do you have difficulty focusing on tasks?
    - ☐ Yes
    - ☐ No

11. Do you consider yourself to be an impulsive person
    - ☐ Yes
    - ☐ No

12. Are you easily distracted by irrelevant sights or sounds?
    - ☐ Yes
    - ☐ No

13. Are you very forgetful?
    - ☐ Yes
    - ☐ No

14. Do you find yourself making careless mistakes often?
    - ☐ Yes
    - ☐ No

15. Do you often feel restless or fidget with hands or feet?
    - ☐ Yes
    - ☐ No

16. Do you find it difficult waiting in line for your turn?
    - ☐ Yes
    - ☐ No

17. To the best of your knowledge, do you have any problems in depth perception?
    - ☐ Yes
    - ☐ No

If yes, briefly describe what problems you experience in depth perception

____

____

APPENDIX E. Instructions for participants

Instructions for participants

1. Please read and sign the informed consent form
2. First, a test of far and near visual acuity will be administered to check for normal vision. Please ensure that you are wearing your lenses or glasses if you use them. You will be asked to stand at a certain distance and read alphabets displayed to you on the wall. Next, a card will be displayed and you will be asked to read alphabets printed on the card.
3. Then, you will complete a pre-experiment questionnaire to gather information regarding vision, age, depth perception etc.
4. We will introduce you to the equipment used in augmented reality systems and assist you in putting on the display. This display is a unique head-worn display called NOMAD made by Microvision.
5. You will be able to see the real world through the display. You will sit on a chair. The height of the chair will be adjusted so that you can see the monitor in your field of view. Please ensure that you are wearing your lenses or contacts if you use them. Several practice trials will be performed to familiarize you with the display and the task.
6. When you look through the display, you will see nonsense words on a monitor and virtual text hanging in the air. Your task is to first locate a pair of similar letters one of which is upper case and the other lower case (e.g. aA, eE, Gg) from the virtual text and then determine how many times that particular letter (upper or lower case) appears in the real world text on the monitor. You will respond by pressing the zero, one, two, or three number keys on the keyboard indicating that the letter is not present, present once, present twice, or present thrice in the real world text. This procedure will be repeated several times and you will be given 25 seconds to try and complete as many trials of it as possible. After 25 seconds the screen will be blanked and to proceed with the next task you will need to press the space bar. You will repeat the same procedure with different nonsense words several times. You should work as quickly and accurately as possible; once you determine the number of times a letter appears in the text, you should press the corresponding key as soon as you can.
7. At regular intervals you will be asked to rate the condition of your eyes according to the following scale by pressing the appropriate number on the keyboard. This scale will also be made available to you on paper. After that please close your eyes and relax for 45 seconds. We will tell you when to start the next task.

1	2	3	4	5	6	7
very rested	rested	somewhat rested	neither rested nor fatigued	somewhat fatigued	fatigued	very fatigued

8. We expect the entire session to take about two hours.
9. When you have completed all tasks, we will ask you to fill out another questionnaire, to obtain your feedback on the experiment.

Please remember that we are evaluating the augmented reality display; we are not evaluating you in any way.

APPENDIX F. Post-experiment questionnaire

Post-Experiment Questionnaire

Based on the experiment that you just performed please answer the following questions.

1. Did you experience any kind of eye strain/eye discomfort during the experiment? If yes, briefly describe it (e.g. blurry vision, watering of eyes, tiredness etc.)

____

____

____

____

2. Did you feel any difference in your ability to complete the task when the monitor or virtual text position changed? If so, please explain why you thought it was easier or more difficult to complete the task at certain viewing distances as compared to other distances.

____

____

____

____

3. Please rate your physical comfort/discomfort level while using the display. Circle the appropriate response.

1	2	3	4	5	6	7
very comfortable	comfortable	somewhat comfortable	neither comfortable nor uncomfortable	somewhat uncomfortable	uncomfortable	very uncomfortable

If you experienced any kind of physical discomfort, briefly describe it.

____

____

____

4. Any comments about the experiment

Thank you for participation!

APPENDIX G. Post-experiment interview questions

1. Where you able to perceive virtual text as being situated at different distances from you? If yes, how many different distances could you perceive?
2. Which virtual text distance did you prefer (near, medium, or far) and why?
3. Which virtual text distance did you think was the hardest for you to read and why?
4. How would you compare the situations where the real-world monitor and virtual text were at matched distances as compared to when they were at different distances?
5. Did you feel any difference when the focus depth setting changed?

VITA

Divya Gupta received her Master of Science degree in Industrial and Systems Engineering (ISE) from Virginia Tech in June 2004. She was enrolled in the Human Factors program in the ISE Department. Divya holds a Bachelor of Engineering degree in Chemical Engineering from Bombay University, India. From May 2003 to May 2004 she worked as a Graduate Research Assistant for the Systems Research Center under Dr. Deborah Hix. Her thesis and research which was sponsored by the Office of Naval Research dealt with investigating the impact of augmented reality on human task performance. Her research interests lie in the user-centered design of interfaces. She will be working for Nokia Mobile Phones in San Diego, California as a Usability Scientist from June 2004.