2009

Direct Manipulation Of Virtual Objects

Long Nguyen

University of Central Florida

STARS Citation

Nguyen, Long, "Direct Manipulation Of Virtual Objects" (2009). *Electronic Theses and Dissertations, 2004-2019*. 3907.
https://stars.library.ucf.edu/etd/3907

DIRECT MANIPULATION OF VIRTUAL OBJECTS

by

LONG K. NGUYEN
B.S.E. University of Central Florida, 1988
M.S.E.E University of Central Florida, 1991

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Industrial Engineering & Management Systems
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Summer Term
2009

Major Professor: Linda C. Malone

ABSTRACT

Interacting with a Virtual Environment (VE) generally requires the user to correctly perceive the relative position and orientation of virtual objects. For applications requiring interaction in personal space, the user may also need to accurately judge the position of the virtual object relative to that of a real object, for example, a virtual button and the user's real hand. This is difficult since VEs generally only provide a subset of the cues experienced in the real world. Complicating matters further, VEs presented by currently available visual displays may be inaccurate or distorted due to technological limitations.

Fundamental physiological and psychological aspects of vision as they pertain to the task of object manipulation were thoroughly reviewed. Other sensory modalities – proprioception, haptics, and audition – and their cross-interactions with each other and with vision are briefly discussed. Visual display technologies, the primary component of any VE, were canvassed and compared. Current applications and research were gathered and categorized by different VE types and object interaction techniques. While object interaction research abounds in the literature, pockets of research gaps remain. Direct, dexterous, manual interaction with virtual objects in Mixed Reality (MR), where the real, seen hand accurately and effectively interacts with virtual objects, has not yet been fully quantified.

An experimental test bed was designed to provide the highest accuracy attainable for salient visual cues in personal space. Optical alignment and user calibration were carefully performed. The test bed accommodated the full continuum of VE types and sensory modalities for comprehensive comparison studies. Experimental designs included two sets, each measuring depth perception and object interaction. The first set addressed the extreme end points of the

Reality-Virtuality (R-V) continuum – Immersive Virtual Environment (IVE) and Reality Environment (RE). This validated, linked, and extended several previous research findings, using one common test bed and participant pool. The results provided a proven method and solid reference points for further research. The second set of experiments leveraged the first to explore the full R-V spectrum and included additional, relevant sensory modalities. It consisted of two full-factorial experiments providing for rich data and key insights into the effect of each type of environment and each modality on accuracy and timeliness of virtual object interaction.

The empirical results clearly showed that mean depth perception error in personal space was less than four millimeters whether the stimuli presented were real, virtual, or mixed. Likewise, mean error for the simple task of pushing a button was less than four millimeters whether the button was real or virtual. Mean task completion time was less than one second. Key to the high accuracy and quick task performance time observed was the correct presentation of the visual cues, including occlusion, stereoscopy, accommodation, and convergence. With performance results already near optimal level with accurate visual cues presented, adding proprioception, audio, and haptic cues did not significantly improve performance. Recommendations for future research include enhancement of the visual display and further experiments with more complex tasks and additional control variables.

In loving memory of my
bà ngoại, Mai Thị Anh,
who, through her jolly ways of living,
taught me compassion and life

ACKNOWLEDGEMENTS

This dissertation work has been long and arduous. I would not have embarked on such an expedition without the encouragement of loved ones. I could not have completed this journey without the help of many around me. Words cannot adequately express the gratitude I owe.

First and foremost, I thank the committee for their endless contribution. Dr. Linda Malone always put the student first and made sure I had the wherewithal to overcome every hurdle, whether it's research challenges or administrative obstacles. Dr. Jannick Rolland spent a great deal of resources and long hours to provide me with state-of-the-art visual display and continuous guidance. Dr. Bob Kennedy, whom I am so honored to have worked with even before I started this dissertation, never hesitated to spend the precious time to share his extensive knowledge and priceless research library. Dr. Kay Stanney, renowned in the field, never missed an opportunity to guide and encourage me. Dr. Kent Williams supported my studies long before this dissertation started and continued to guide me throughout the years. Dr. Karwowski selflessly came to the rescue, thankfully, when there was no other solution in sight. I am humbled and grateful to have learned from such well-known experts.

Completing a dissertation while employed full time would be impossible if it weren't for support from the work place. Dr. Robert Breaux encouraged me to apply for the doctorate program. Dr. Jim Patrey, Dr. Joseph Cohn, and Dr. Andy Mead helped me to get started on my first human factors research efforts. My supervisors, Mr. Gary Fraas, Mr. Maynard Zettler, Mr. Ed Meyer, and Dr. Dan Dwyer encouraged higher education and have all been completely supportive of my studies. From NAS Jacksonville, Mr. Mike Muehlbauer, CDR Kevin Rasch, CDR Allen Crisp, and LCDR Michael Sypnewski went far out of their ways to provide me the

ideal facility and participants for the experiments. My colleagues, Dr. Kelsey Henderson, Dr. Joan Johnston, and Mr. Curtis Conkey spent much time to help and more importantly, continuously urged me to push forward.

Lastly, and most importantly, my family understands too well my insatiable thirst for academic knowledge. This has been a most selfish undertaking, indulging myself in scholarly studies while loved ones toiled without me day in and day out, year after year. Kids – Anthony, Thomas, Britney, and Victoria – you are my hope and sunshine. Anna, you are my everything. Thank you, Anna, for your unwavering support, sacrifice, and encouragement for my lavish hobby. The hardest part of this journey has been glimpses back at your hardship and my neglect for our family. This has been loads of fun for me and mighty tough for you. Yet, you hung on, continuously pushed, and climbed each step with me. We made it. *You* did it.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABREVIATIONS

AR	Augmented Reality
ARPAB	Augmented Reality Performance Assessment Battery
BOOM	Binocular Omni-Orientation Monitor
CAVE	Cave Automatic Virtual Environment
FOV	Field of View
HCI	Human Computer Interaction
IG	Image Generation
IVE	Immersive Virtual Environment
IPD	Interpupil Distance
HMD	Head Mounted Display
HMPD	Head Mounted Projection Display
HTD	Head Tracked Display
MHP	Model Human Processor
MR	Mixed Reality
NVG	Night Vision Goggle
RE	Real Environment
OLED	Organic Light Emitting Diode
OSTHMD	Optical See-Through Head Mounted Display
PreCogPAB	Pre-Cognition Performance Assessment Battery
RSD	Retinal Scanning Display
R/R	Real object next to Real object or Real hand manipulating Real object

R/V	Real object next to Virtual object or Real hand manipulating Virtual object
VEPAB	Virtual Environment Performance Assessment Battery
VRD	Virtual Retinal Display
VSTHMD	Video See-Through Head Mounted Display
V/V	Virtual object next to Virtual object or Virtual hand manipulating Virtual object

CHAPTER ONE: INTRODUCTION

The Navy training community, as with other industries, over the past decade has explored the potential of Virtual Environments (VEs) as solutions for many applications. Allard (1997, p. 2-1), then VE research program manager at the Office of Naval Research (ONR), stated at a NATO conference, “VE training systems have the great advantages of compactness, deployability, software reconfigurability, and affordability.” The Navy trainers have been getting more and more compact and reconfigurable ever since.

The Need for Research in Direct Interaction with Virtual Objects

Examples of the enormity of conventional live training systems and exercises include Naval Strike and Air Warfare Center (NSAWC), Capability Exercises (CAPEX), and the Joint Task Force Exercise (JTFEX). Each of these events requires a large area in the dessert or at a seaport and a number of ships or airplanes (Clancy, 1999).

In the 90s the Navy fielded the F/A-18 Tactical Operation Flight Trainer (TOFT), the Air Force, their F16 Distributed Mission Trainer, and the Marines, their Landing Craft, Air Cushioned (LCAC) Full Mission Trainer (Schaffer, Cullen, Cohn, & Stanney, 2003). These trainers do not require hundreds of acres of land to conduct training. However, they each do require a high bay area or an entire building.

More recently, as a result of VE research Allard (1997) directed earlier, the Navy fielded a virtual reality training system, Virtual Environment Submarine – VESUB2000. The entire training system is not much larger than a mockup of the bridge the conning officer stands in when commanding the submarine. ONR's most current VE research program, Virtual Technology and Environment (VIRTE), leveraged similar technology to reduce the footprint of trainers and produce portable and deployable microsimulator systems (White, Arena, Newton, & Hopper, 2003).

Shrinking full mission trainers to deployable size, however, remains a challenge mainly due to the necessity for a physical cockpit. VESUB was an exception since conning officers do not manipulate controls within a physical cockpit, but rather, control the vessel's simulation through verbal commands to a synthetic crew (Munro, Breaux, Patrey, & Sheldon, 2002; Hays, Vincenzi, & Bradley, 1998). The need for affordable, portable, deployable, and reconfigurable trainers still remains.

Shrinking the cockpit to computer bits and bytes means having to interact with virtual objects. Herein, lies the research challenge. Visual displays presenting two-dimensional images, even if correctly designed, cannot provide perfect three-dimensional views (Wann, Rushton, & MonWilliams, 1995). Furthermore, haptic displays (or haptics), for the sense of touch, are difficult to implement due to mechanical force feedback requirements and can be even bulkier than a cockpit mockup. Recent development of new head mounted displays (Cakmakci & Rolland, 2007; Cakmakci, Vo, Thompson, & Rolland, 2008; Hua, Girardot, Gao, & Rolland, 2000; Hua, Gao, Biocca, & Rolland, 2001; Martins, Shaoulov, Ha, & Rolland, 2007; Rolland, 2000; Rolland, Biocca, Hamza-Lup, Ha, & Martins, 2005; Rolland, Krueger, & Goon, 2000; Rolland, Parsons, Poizat, & Hancock, 1998; Rolland, Wright, & Kancherla, 1997; Rolland, Yoshida, Davis & Reif, 1998) or techniques for simple touch feedback (Schiefele, 2000), may provide the needed edge to tackle problems associated with direct interaction with virtual objects.

Therefore, a study was conducted herein that reviewed current potential issues with direct manipulation of virtual objects. Experiments were carried out using novel, promising techniques not yet considered in the literature for such application. A virtual environment that can be developed to effectively support such interaction would be a significant step towards eliminating bulky, physical cockpit mockups. The ability to directly interact with virtual objects could have a significant impact not only for navy training, but also for other training communities, such as commercial air, for other application domains, such as virtual prototyping, or for Human Computer Interaction (HCI) techniques in general, such as 3-D Graphical User Interfaces (GUIs).

CHAPTER TWO: LITERATURE REVIEW

This chapter reviews some basic definitions and how they are used in this document. The second section of the chapter reviews fundamental physiological and psychological aspects of vision, audition, haptic, and proprioception. These sensory modalities and their cross-interaction are highly relevant to the task of directly manipulating virtual objects. A number of pitfalls related to visual stress and adaptation and vestibular side effects are highlighted. The third section reviews visual display technologies, the primary component of any VE, relevant to this work. Head Mounted Projection Display (HMPD) is one that has a set of features that combine key advantages of head mounted displays and projection systems. The last section in this chapter reviews current applications and research in object manipulation. While manipulation of virtual objects is not a research gap, it has only been demonstrated in an indirect manner using a virtual representation of the hand. Direct, dexterous, manual manipulation of virtual objects in Mixed Reality (MR) has not yet been fully experimented within the literature.

Definitions

Reality-Virtuality continuum, optically real and virtual images, real and computer-generated images, real and virtual objects, Cutting's spaces, and Shneiderman's definition of direct manipulation are discussed below.

Milgram's Reality-Virtuality Continuum

Milgram (Milgram & Kishino, 1994; Milgram & Colquhoun, 1999) introduced the Reality-Virtuality (R-V) continuum, which encompasses all environments from Reality Environment (RE) to purely Virtual Environment (VE, or Virtuality). Within this continuum, Augmented Reality (AR) refers to the real or physical world enhanced by computers. Augmented Virtuality (AV), on the other hand, refers to the computer-generated world enhanced with real world images. AR and AV overlap extensively. AR begins at but excludes the RE end point. Similarly AV begins at but excludes the VE end point. Additionally, mixed-reality (MR) is a term Milgram used to encompass the entire continuum except for the end points (Figure 1).

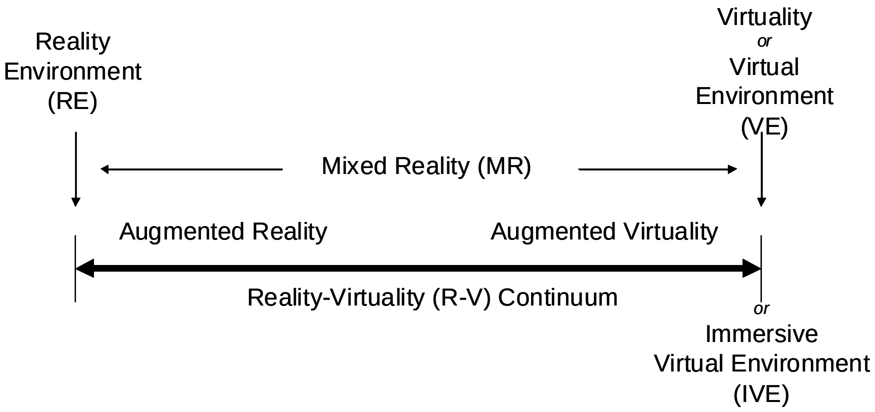


Figure 1. Milgram's Reality-Virtuality Continuum.

This continuum is referred to numerous times in this dissertation in extensive comparisons and discussions of RE, pure VE, and MR. Additionally, since VE is often used as a broad term that can encompass MR, another widely used term, Immersive Virtual Environment (IVE), is used in this document to refer to pure VE or Virtuality.

Optically Real Image and Optically Virtual Image

The NATO RTG (NATO RTO, 2003) suggested that it is probably in optics alone, that the word "virtual" has a definitive meaning. In this optics community the term virtual image has an authoritative, concise, textbook definition that is accepted internationally and contrasts sharply with the definition of real image. Virtual image pertains to the perceived image, but one without light rays physically impinging on or passing through that image. An example is the image behind a mirror. A real image, on the other hand, requires that light rays forming that image actually focus on it and pass through it or reflect from it (Figure 2). For clarity, this dissertation consistently refers to this concept as "optically virtual image" and "optically real image" to distinguish it from any other uses of real and virtual images that may be discussed elsewhere. These terms are also used later to distinguish between optically virtual images, afforded by the Head Mounted Display (HMD), and optically real images, afforded by the HMPD.

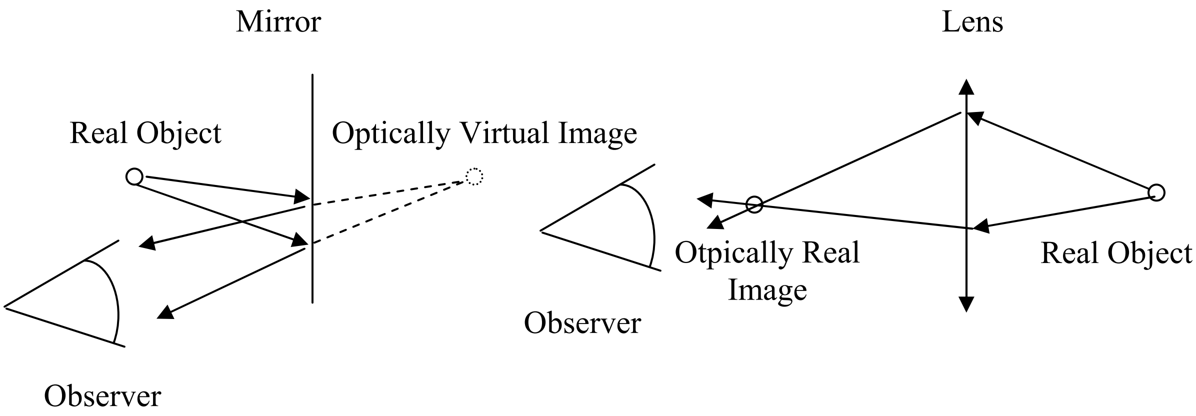


Figure 2. “Optically Real Image” and “Optically Virtual Image”.

Real Image and Computer-Generated Image

According to Milgram’s definition (Milgram et al., 1999), virtual images are derived from computer models and real images are derived from physical objects. Examples of purely virtual images are models created on and displayed by the computer. Examples of purely real images, by this definition, include unaltered photographs and video that are not created nor processed by the computer. There is some subtle confusion with this definition. Since many camcorders and cameras today automatically, digitally, and necessarily compress images, a strict application of Milgram’s definition would find the resulting images to be virtual. However, they are, intuitively, real images, similar to those produced by conventional cameras and camcorders. Milgram’s definitions are useful for differentiating between environments, e.g., RE, MR, and IVE, but applying them to images can cause confusion.

This dissertation consistently describes images captured by cameras or camcorders as "photographs" or "video images" rather than "real" or "virtual" images as defined by Milgram (1999). Likewise images created from computer models are referred to as "computer-generated” images instead of "virtual” images. The use of the term "real image" (not “optically real image”) in this dissertation is, therefore, strictly reserved for direct view of physical objects. The term “virtual image” is avoided as much as possible and is replaced with computer-generated images to eliminate any confusion.

Real Object and Virtual Object

There are real objects and there are also virtual objects. These terms are carefully reserved and distinguished from real images and virtual images as well as from real environments and virtual environments. It is widely accepted in the VE community that virtual objects are artifacts that are generated from computer models, and although they have no physical existence, they can be perceived. This distinguishes virtual objects from real objects that truly can be seen and touched. This dissertation focuses on experiments measuring the accuracy and speed of the real hand (real object) manipulating virtual cockpit controls (virtual objects). Appendix B expands upon the subtle distinctions between the ways the terms “real” and “virtual” are applied in this dissertation.

Cutting’s Spaces

Visual cues have been historically categorized as static vs. dynamic or binocular vs. monocular (Ellis, Bucher, and Menges 1995). For example, NASA (1981) groups the cues of image disparity (stereopsis) and convergence into the binocular category, and groups accommodation, motion parallax, size of familiar objects, linear perspective, interposition (occlusion), aerial perspective, shadows, and light intensity into the monocular category. Ellis et al. (1995) also noted that a new way of categorizing visual cues has been introduced by Cutting and Vishton (1995) based on distance, which places more emphasis on behavioral affordances.

Cutting and Vishton (1995) separated the space around an observer into three categories: personal, action, and vista. Personal space is immediately around the user and generally within arm's reach, i.e., out to about 2 meters. Action space is just beyond personal space and goes out to 30 meters. Vista space is from 30 meters out. This dissertation concentrates on personal space, which prioritizes visual cues differently than those for action or vista space. Personal space emphasizes different characteristics for visual displays and places more stringent requirements on other VE apparatus.

Direct Manipulation

Shneiderman and Maes (1997) refer to direct manipulation as a HCI technique where the user is afforded the ability to see and manipulate virtual objects directly with full control and predictability. This definition is widely used in the literature and includes the use of tools such as menus in GUIs (Eberts, 1999). In this dissertation, however, a narrower definition of direct manipulation is used: one that refers to natural, dextral interaction where the hand is touching the real object or is directly adjacent to the virtual object.

Senses

Object manipulation is an egocentric task that requires multiple sensory modalities. These modalities generally include vision, audition, touch, and proprioception. Of these, vision is important and often required for localization, spatial acuity, and perception of shape, size, texture, and body orientation (Boff & Lincoln, 1986). Since object manipulation is primarily a spatial task where vision is generally dominant this study concentrates more on this modality and side effects associated with it. Other relevant modalities are discussed more briefly.

Visual Sensory Modality

Numerous cues affect visual perception. A non-exhaustive list of vision cues and characteristics generally considered in current image generation and display systems includes acuity, unmatched by display device resolution, instantaneous field of vision, again unmatched by display device field of view, aerial and linear perspective, size and height, texture and gradient, lighting, shade, shadow, and motion parallax. Of these, the more advanced display systems consider additional cues including occlusion, stereopsis, vergence, accommodation, and depth of field. Basic vision principles for each of these latter cues and for motion parallax are briefly reviewed below. More importantly, visual aftereffects related to these are also discussed.

Occlusion refers to interposition of one object on another and is readily provided in IVE applications by software and Image Generation (IG) systems. MR, however, requires precise, real-time tracking of real objects and extensive computation in order to generate a completely correct occlusion. Occlusion is an important and sometimes the only depth cue for certain visual conditions, for example, for objects in the far-field (vista space) where there are no other effective cues (Cutting &Vishton, 1995).

Stereopsis refers to depth perception arising from lateral retinal image disparity (Boff & Lincoln, 1986). Lateral separation of the eyes provides each with a different perspective and therefore disparate images that are fused together psychologically providing for depth perception.

Convergence refers to inward rotation of the eyes. Divergence refers to outward rotation. Vergence is the general term for both. Conjugate refers to horizontal eye rotation in the same direction (Burdea & Coiffet, 1994; Popescu, Burdea, & Trefftz, 2002; Boff & Lincoln, 1986). Vergence, in addition to stereopsis, has also been shown to provide depth cues primarily in personal and action space (Ellis & Menges, 1997, 1998).

Accommodation refers to focusing of the eye. Reflex accommodation refers to focusing of a blurred image by changing the lens until the image is sharp. Tonic accommodation refers to the focus point while the eye is "resting", generally 0.5 to 2.0 m, or while observing bright large surfaces. Proximal accommodation refers to a focus point based on a priori knowledge about object distance. Dark focus refers to the focus point in the absence of light, about 1.7 diopters, corresponding to a focal length of about 0.6 m. Accommodation has also been shown to provide depth cues primarily in personal and action space (Ellis & Menges, 1997; Popescu et al., 2002).

Vergence and accommodation are neurally cross-linked, i.e., each affects the other. Accommodation producing vergence eye movements is termed accommodation vergence. Similarly vergence producing accommodation is termed vergence accommodation (Mon-Williams, Plooy, Burgess-Limerick, & Wann, 1998; Wann & Mon-Williams, 2002).

Closely related to accommodation is depth of field. Depth of field is the range where objects are in focus instantaneously. For a nominal 4mm eye pupil, the depth of field is +/-0.07 diopters (Valyus, 1966). This equates, for example, to depth of field of 0.93 to 1.07 m for a scene about 1m away. However, tolerance for blur considerably extends this apparent depth of field (Boff et al., 1986).

Visual Depth Cues

Separating the spaces into categories allowed Cutting and Vishton (1995) to group and rank nine visual cues relevant to depth perception. They made arguments for each cue's effective range, and graphed its effectiveness over distance. These cues include occlusion, binocular disparities, motion perspective, relative size, convergence, accommodation, relative density, aerial perspective, and height in visual field. Table 1 ranks the order of importance of each cue in each category of space.

Table 1. Cutting's Ranking of Nine Visual Depth Cues

Order of Importance	Personal Space	Action Space	Vista Space
1	Occlusion	Occlusion	Occlusion
2	Stereopsis	Height in Visual Field	Relative Size
3	Motion Parallax	Motion Parallax / Relative Size	Height in Visual Field
4	Relative Size	Motion Parallax / Relative Size	Relative Density / Aerial Perspective
5	Convergence / Accommodation	Stereopsis	Relative Density / Aerial Perspective
6	Convergence / Accommodation	Relative Density	Motion Parallax
7	Relative Density	Aerial Perspective	Stereopsis
8	Aerial Perspective	Convergence / Accommodation	Convergence / Accommodation
9	----------	Convergence / Accommodation	Convergence / Accommodation

Of the nine cues identified by Cutting and Vishton (1995), occlusion, binocular disparities (stereopsis), convergence, and accommodation, discussed previously, are the cues most difficult to provide accurately and are, consequently, rarely provided correctly by visual display systems. The remaining five cues, motion parallax, relative size, relative density, aerial perspective, and relative height in visual field, are generally more easily provided and usually included by the visual display devices in Virtual Environments (NASA, 1981). Of these, motion parallax and relative size are among the top four strongest visual cues for depth perception in personal space. Also, Cutting and Vishton did not include height in visual field in the personal space category and a few cues beyond the nine. These are relevant to this dissertation and warrant a brief discussion.

Motion parallax is the appearance of movement of stationary objects in the foreground relative to other stationary objects in the background caused by an observer's motion. For example, when looking out to the side of a moving vehicle, distant objects on the horizon appear to move with the observer whereas nearby objects such as the lampposts appear to move in the opposite direction. Spatial and temporal accuracy for motion parallax in VE depend heavily on tracking systems that may be far from perfect. In VE systems, if head tracking is too far off or its latency too high, motion parallax cues suffer and performance in the VE degrades as a result, especially in personal space where feedback is immediate and physiological fidelity requirements are high.

Relative size cues require two objects with some spatial relationship to each other, one object placed at different distances, or an object that has a familiar or assumed size (Cutting & Vishton, 1995). It is worth noting that while the cue of occlusion is the only one stronger than relative size in all three spaces (personal, action, and vista), it only provides ordinality information, that is, only data about which object is closer. The relative size cue can provide ordinality and scaled information, that is, data about which is closer and by how much.

Cutting and Vishton (1995) excluded height in visual field from the list of effective cues in personal space (Table 1). Height in the visual field refers to the apparent rise of the ground in the far-field relative to that of the near-field. Because Cutting and Vishton considered the observer to be one who is standing, this cue is not relevant in personal space. The standing observer does not see much of the ground immediately around the body. In this dissertation, however, the observer could be sitting and could be surrounded by cockpit panels around chest height. Therefore, apparent rise of a point farther on the surface of a cockpit panel relative to a closer point on the same panel can potentially provide a significant cue. Consequently, in such situations, height of field can be an effective cue in personal space.

Vishton and Cutting (1995) also intentionally left out other typical visual cues in their model and made rationale for the omission. These include linear perspective, brightness, lighting and shading, texture gradients, kinetic depth, kinetic occlusion and dis-occlusion, and gravity. They noted that these other cues are either covered within combinations of the nine (texture gradient, linear perspective, kinetic occlusion/dis-occlusion), provide more object shape cue rather than depth cue (texture gradient, shading, and kinetic depth), or else are inconsistent or applicable for only very specific situations (brightness, light, gravity). Regardless, some of these cues are relevant for this dissertation and the task of manipulating cockpit instrument controls.

For example, Cutting and Vishton (1995) noted that linear perspective is only the combination of size, density, and compression expressed in the form of parallel lines. However, these parallel lines can represent many surfaces within a cockpit environment, e.g., front, overhead, and side instrument panels, rectangular displays, and rectangular instrument control and indicators sections. Therefore, linear perspective is an important cue for such applications. Similarly, Cutting and Vishton noted that brightness and light offer no depth cues in situations of uniform lighting. However, it may be relevant in a cockpit environment where lighting can be non-uniform such as that coming from a local source. For shades, Cutting and Vishton noted that this cue provides information on object shapes rather than depth. However, shape information is relevant for the task of reaching and manipulating virtual buttons, dials, or switches. Finally, kinetic occlusion is a combination of occlusion and motion parallax, both of which are significant in the task of manipulating cockpit controls.

Visual Stress

Muscles lowering the eyes also aid in convergence and those raising the eyes also aid in divergence. Therefore, inappropriate gaze angle, for example, convergence at high angle, produces visual stress. Vergence adaptation can also be induced if the optics is misaligned in a non-collimating system. Strabismus (cross-eyed) could result from pressure placed on a suboptimal binocular system (Wann et al., 2002). Ehrlich (1999) found that dark vergence is a good measure for identifying visual stress and suggests the same be used to assess whether participants have achieved readaptation.

HMDs use refractive lenses that affect how the user focuses to see a sharp image. If improperly designed, HMDs can induce accommodative adaptation (Wann et al., 2002; Wann et al., 1995; Kawara, Ohmi, & Yoshizawa, 1996). Mon-Williams, Tresilian, Strang, Kochhar, and Wann (1998) found evidence of neural compensation associated with prolonged exposure to defocused conditions.

Cross-link adaptation can be induced when conflict between accommodation and vergence exists (Wann et al., 1995; Azuma, 1997). This is a concern especially for stereoscopic displays because the system portrays three-dimensional space on a two-dimensional image plane (Wann et al., 1995). Accommodation remains fixed on the image plane but vergence angle changes with the depth of the perceived 3-D object. This disassociation of accommodation and vergence is more pronounced when larger depth intervals are displayed. Southard (1997) noted that accommodation vergence conflict is probably the only source of visual stress that cannot be eliminated completely in current HMD designs.

A number of experiments have been conducted to quantify aftereffects and to compare binocular (stereoscopic) with biocular (non-stereoscopic) displays. Mon-Williams et al. (Mon-Williams & Wann, 1998; MonWilliams, Wann, & Rushton, 1993) found in one specific setting that stereoscopy did not cause visual stress over short viewing periods. However, using a stereoscopic display that required change in vergence for a period of ten minutes did cause visual stress. Rushton, MonWilliams, and Wann (1994) also found evidence of adverse effects from binocular display usage after 10 minutes. They ran the same experiment again using a biocular (instead of binocular) display for 30 minutes and found no adverse effects.

Valyus (1966, p. 371) found that a change in convergence angle of up to 1.6 degrees with fixed accommodation is acceptable. Exceeding this tolerance leads to excessive accommodation vergence conflict. Hua, Gao, Brown, Ahuja, and Rolland (2002) used this guideline to successfully design an AR using binocular display and did not report any visual stress problem.

Other Modalities and Cross-Modal Interactions

Other sensory modalities besides vision are salient to interaction within a Virtual Environment. Among these are audition, haptics, and proprioception. These and their cross-modal interactions are briefly discussed below.

Audition is required or optimal for perceiving temporal patterns and time intervals (Boff et al., 1986). Auditory cues enhance awareness of a VE and are especially useful for collision or tactile cueing (Shilling & Shinn-Cunningham, 2002). Auditory displays can be monaural, stereo, or spatial. Monaural, or diotic, displays provide identical sounds to both ears. Stereo, or dichotic, displays employ delay and intensity differences on the left and right channel to simulate some directionality. Spatial audio utilizes models of the head (Head Related Transfer Function, HRTF) or of the room to generate sounds with rich spatial cues.

Haptic cue refers to the sense of touch and includes sensations arising from stimulation of receptors in the skin and associated tissues (Vince, 1995). While vision dominates spatial tasks and audition dominates temporal tasks, haptics can provide substitute or redundant information that improves overall perception or task performance (Popescu et al., 2002; Boff et al., 1986). In dark environments, haptics is also used for spatial discrimination (Popescu et al., 2002). Mon-Williams et al. (1998) found that when visual background is lacking, haptics can be more dominant in spatial tasks than vision. Force resolution also increases spatial resolution, which makes haptics critical for complex direct manipulation tasks in virtual environments (Popescu et al., 2002).

Proprioception is a sense of body position and movement (Boff et al., 1986). Proprioception uses receptors within joints, muscles and deep tissues (Vince, 1995). Proprioception can dominate in spatial tasks, when vision is lacking (Boff et al., 1986). The vestibular system also contributes to proprioception and is tied to head motion affecting eye-hand coordination.

Both vision and audition represent spatiotemporal information. The overlap results in redundancy if synchronized and augments perception in both senses (Popescu et al., 2002). Similarly for the haptics-visual pair, Biocca Kim, and Choi (2001) demonstrated that haptic sensation is perceived when there is no such stimulation if the visual sensation is convincing enough. This perceptual illusion is termed synesthesia. Gross (2004) also demonstrated instances where substituting cues from one modality for another that is absent improves perception of affordances and, therefore, task performance.

Vestibular System and Associated Side Effects

The vestibular system includes Semicircular Canals (SCCs) and otolith organs that primarily sense the head's rotational and translational movements. The three approximately orthogonal SCCs detect angular accelerations. The two otolith organs, the utricle and the saccule, detect linear acceleration in the horizontal and vertical directions, respectively (DiZio & Lackner, 2002; Stoffregen, Draper, Kennedy, & Compton, 2002). Locomotion or navigation tasks in VEs can induce vection. This sense of perceived self-motion is usually derived from visual cues often times with corresponding but poorly matched physical motion cues and other times without any motion cues at all. The result is discrepancy between visual system and vestibular system, which can cause negative side effects, such as disorientation and nausea.

Besides navigation tasks, proprioception is also directly related to motor control (DiZio et al., 2002). Some perceptual-motor tasks utilize vestibular input (Stoffregen et al., 2002). In particular, tasks that require head rotation, for example locating controls on surrounding panels within a cockpit necessarily rely on vestibular inputs to provide information on head movements. In conjunction with occulomotor response, this helps to stabilize gaze and is commonly referred to as vesibulo-ocular reflex (VOR). Both SCCs and the otolith organs affect rotational VOR, with SCC inputs more dominant than otolith. Any error in stabilization of the eyes brought about by incorrect visual motion cues in such a VE can cause adaptation of proprioception and the vestibular system. This perceptual adaptation, compensating for conflicting cues between visual, haptic, and vestibular systems in the VE, can produce aftereffects that degrade subsequent performance in the normal environment (DiZio et al., 2002; Stoffregen et al., 2002).

The accuracy of simple eye-hand coordination tasks, such as pointing, is affected by changes in visual, haptic, or vestibular systems. The pointing errors stemming from undesirable vestibular adaptation are known as past-pointing (Stoffregen et al., 2002). Manual tasks following VE exposure, specifically, from see-through HMDs, can be significantly more inaccurate and may take longer to complete. This error stems mainly from the visual scene generally not being perfectly matched in gain (rotational speed) or phase (lag) with vestibular inputs. Therefore, the VOR necessarily adapts in order to maintain stable vision in the VE, at the expense of a subsequent mismatch in the physical environment after VE exposure (Stoffregen et al., 2002).

Visual Display Technology

Vision is the dominant modality for spatial tasks (Boff et al., 1986) such as object manipulation. Visual display technology relevant to this task is broken down into three categories consistent with Milgram’s Reality-Virtuality (R-V) continuum and discussed below.

Reality (RE) Display Technology

Using Milgram's definition, RE, or reality, which is on the far left side of the Reality-Virtuality (R-V) continuum, includes video displays. These provide unaltered video images of the real world and have a number of applications, such as collaborative VEs or video teleconferencing. One application that is somewhat relevant to this dissertation's topic of object manipulation is laparoscopy, in which surgeons use video displays in conjunction with probes to visualize the work area as they perform an operation (Birkfellner et al., 2002).

Immersive Virtual Environment (IVE) Display Technology

Perhaps the most popular VE display is the opaque HMD, generally referred to as a HMD, with "opaque" omitted and assumed. HMDs provide total immersion in a virtual environment, visually. Typical characteristics that define the quality of HMDs include resolution, field of view (FOV), exit pupil size, brightness, contrast, color, head adjustment, weight, and eye relief distance. Higher-end HMDs also support two visual channels for displaying a different image to each eye, focus adjustment for myopic or hypermetropic users, and interpupil distance (IPD) adjustment to match a user's IPD.

A binocular Omni-Orientation Monitor (BOOM) is similar to a HMD but is attached to a mechanical arm (Blade & Padgett, 2002). The arm provides counter balance making the BOOM almost weightless to users. It also provides for six degrees of freedom tracking of viewing position and attitude. Tracking information measured by the mechanical arm is very precise. BOOM devices have been used in visualization applications such as virtual prototyping where the user analyzes whether or not one can touch virtual controls on a panel design.

Mixed-Reality (MR) Display Technology

Optical See-Through HMDs (OSTHMDs) provide an unhindered view of the real world. This assures that real world visual information is absolutely correct and instantaneous. OSTHMD, therefore, provides for perfectly synchronized information between visual and proprioception information. This supports a user’s manipulation of real cockpit controls well and is used in vehicle simulation. One drawback of the OSTHMD is that the computer generated image is generally simply superimposed onto the real image, not fused correctly to provide occlusion cues. One way to circumvent this limitation is to choose applications where the real and VEs are distinctly separated. For example, when looking up out of a cockpit window the pilots see a bright computer generated image. When looking down, the lower portion of the cockpit screen is dark and absent of computer images so that the pilots could see and interact with the brightly lit, physical cockpit (Rolland & Fuchs, 2000).

Video See-Through HMDs (VSTHMDs) can guarantee registration of real and virtual scenes to provide correct occlusion cues. However, this is done at the expense of a mismatch between vision and proprioception. The real scene is captured by a CCD camera and fed to a computer, which fuses virtual objects correctly into a video image before it is displayed to the user. Required computational power is significant and therefore a finite delay exists. High-end prototype systems also attempt to provide two cameras aligned with mirrors so that the images recorded are close to what each eye would see (Rolland et al., 2000).

Cave Automatic Virtual Environments (CAVEs) fully immerse a user with images projected on surrounding walls and have proven to be useful in 3-D visualization applications (Vince, 1995). Users generally wear stereoscopic glasses called shutters, which are synchronized with projectors to alternately open and close when the proper left or right image is displayed and shut off. Generally, the CAVE facility is large enough to accommodate multiple users; however, the image is geometrically correct for only one user. Hand tracking devices can also be added for gesture recognition for purposes of manipulating virtual objects. On the negative side, keystoning, a visual distortion as a consequence of off-axis projection is frequently observed. This also generally causes contradiction in accommodation, vergence, and perceived depth because the image is on the projection screen, where the eye is accommodating, but perceived depth and convergence could be at different points (Kakeya, Isogai, Suzuki, & Arakawa, 1999).

Virtual Workbenches, Virtual Tables, or Immersive Workbenches also use projection displays, but towards table surfaces instead of the walls. Like the CAVE, these can also employ shutter glasses for depth perception and tracked gloves for simple hand gesture recognition for interacting with a VE. These displays can also be grouped together in the category of head tracked displays (HTDs) used with fixed screen or projection systems (von Wiegand, Schloerb, & Sachtler, 1999; Leibe et al., 2000; Hinckley, Pausch, Proffitt, & Kassell, 1998).

Autostereoscopic displays can produce a stereo effect without the need for glasses. One example is Dimension Technologies' Virtual Window, which allows each eye to see only its corresponding columns and thus produces stereo effect (Burdea et al., 1994). Another simpler system uses color or polarization glasses to separate images for the two eyes. Kakeya (1999) demonstrated a more complex type of autostereoscopic display. This interesting design uses a large Fresnel lens to create an image plane towards the user. This novel technique allows the image plane to be strategically adjusted for optimal accommodation, thereby, minimizing accommodation/vergence conflict while improving depth perception. For these displays the user's head is not tracked so geometric accuracy of the scenes are not guaranteed.

Rolland, Krueger and Goon (2000) described a conceptual multiplanar volumetric display that can mitigate accommodation/convergence conflicts. Based on human acuity, an engineering analysis showed that such a device is within the capability of current technology. This has the potential for applications requiring visualization of the nearfield and farfield (personal, action, and vista space) simultaneously. Also, recently, Murali, Lee, and Rolland (2007) and Murali, Thompson, and Rolland (2009) demonstrated the embedding of liquid crystal and liquid lenses in optical system design carving the path to successful inclusion in Head Mounted Displays as well

Virtual Retinal Displays (VRDs) or Retinal Scanning Displays (RSDs) use low-power lasers or LED and microelectro-mechanical mirrors to scan an image directly on the human retina. These displays are very light and can be very bright. This overcomes limitations of current state-of-the-art OSTHMDs, which are limited in brightness. Other advantages are low power consumption and large depth of field (Lewis, 2004; Urey, Nestorovic, Ng, & Gross, 1999; Viirre, Pryor, & Nagata, 1998).

HMPDs are similar to OSTHMDs except that the image is projected forward onto a screen and retro-reflected back to the user's eyes. It fundamentally possesses a unique combination of features. These include (Rolland et al., 1998; Hua et al., 2000; Hua et al., 2001):

- Correct occlusion - projecting the image allows for correct occlusion of the virtual object by the real object, like the user's hand. As the user reaches the hand out in front of a virtual object to grab it, the hand instantaneously and correctly occludes that object.
- Image plane position independent of screen position – the image plane can be set to any depth near the retro-reflective screen – in front of it, behind it, or on it. Unlike conventional projection systems, the HMPD can be set so that inconsistency between accommodation and convergence is kept small by keeping the distance between the image plane and objects small. Incidentally, if the image plane is between the user and the screen, the image is actually an optically real image, as if there were an invisible screen in mid air that the image is projected onto.
- Correction of optical distortion – the HMPD can be designed to minimize optical distortion. This eliminates the need for distortion compensation using software or firmware. Correcting for distortion with accuracy and speed is often not optimal with software or firmware.
- Absence of keystoning – Projection displays generally suffer from keystoning, a consequence of off-axis projection with respect to a user's eye points. However, HMPDs, as a result of on-axis projection, do not suffer from keystoning.

- Real world view synchronized with proprioception – as with OSTHMD, the HMPD provides for an unhindered and instantaneous view of the world and thus provides perfect synchronization with proprioception.
- Large FOV – the HMPD can be designed for larger FOV compared to the conventional OSTHMD counterpart. Using a flat combiner, HMPDs can be designed for up to 90 degrees compared to 40 degrees in standard HMDs (Ha, Rolland, & Davis, 2006).
- Retro-reflection from curved surfaces - retro-reflective screen can take nearly any shape without affecting image quality. Retro-reflective elements can even be painted onto surfaces, providing for a wide variety of applications.
- Diminished Reality - the system allows for diminished reality, where real objects are visually removed or camouflaged. As an example a haptic display apparatus can be covered with retro-reflective fabric that also serves as a screen.
- Large exit pupil and eye relief – the HMPD projection optics can be designed for larger pupil size and eye relief requirements than is possible with conventional HMD eyepiece design.
- Supports multiple users - Multiple users can see different perspective of the same virtual object, since projected images cannot be seen except by the user wearing the HMPD.
- Strategic placement of the retro-reflective screen - the retro-reflective material can be strategically placed so that computer generated images appear only in selected locations. Effectively, the displays switch themselves off when users look at each other or look around at other objects that do not have retro-reflective material on

them. This feature is similar to that provided by blue screen or chroma-key technology (Darken, Sullivan, & Lennerton, 2003).

- Large depth of field – the HMPD can be designed with a small iris for larger depth of field (Inami et al., 2000).

There are also current limitations with HMPDs that must be considered when selecting appropriate applications. These include (Rolland et al., 1998; Hua et al., 2000; Hua et al., 2001):

- Variation in size of reflected image – since the retro-reflective material is imperfect, images retro-reflected tend to vary in size depending on the position and orientation of the user and material.
- Image blur – the retro-reflective material is made of small beads or often imperfect corner cubes and, therefore, reflects light back in a finite solid angle contributing to image blur.
- Illumination - Because the image passes through a beam splitter/combiner twice (which reduces the original intensity by at least 75%), illumination is poor in its most basic form compared even to conventional HMDs, which generally also have brightness challenges. This can potentially limit the application of HMPDs in its current form to near-field (personal space or within arm's reach) use only. Approaches to increase the illumination efficiency of HMPDs are under investigation.

Additionally, Appendix C provides a comparison of the advantages and disadvantages of Head Mounted Displays versus Head Mounted Projection Displays. The visual display technologies canvassed in this section are employed widely in virtual object interaction research discussed in the next sections.

Tasks for Direct Manipulation of Virtual Objects

Work related to identifying VE tasks that are relevant to object manipulation are reviewed and discussed below.

Virtual Environment Performance Assessment Battery (VEPAB).

Lampton et al. (1994) developed a set of tasks, VEPAB, to support research on VE training technology. It measures vision, locomotion, tracking, object manipulation, and reaction time. Of these, vision, object manipulation, tracking, and reaction time are relevant to this research topic, that is, direct object manipulation. Among the VEPAB vision tasks, distance estimation or depth perception is of special interest because of the criticality in accurate judgment of object position required for natural manipulation. For the object manipulation VEPAB tasks the slide and dial are relevant to virtual cockpit controls. However, Schiefele (2000) found that buttons are the most often encountered type of control in typical flight cockpits. Therefore, push is another fundamental task that could be considered for this work. For tracking, the task of control and movement of a device to a stationary target is also highly relevant since for a virtual cockpit, one would need to localize stationary panel controls by translating the hand (the device) to the controls (the targets). Finally, both of the VEPAB tasks for reaction time, called simple and choice, are important since completion time is a measure of performance.

Augmented Reality Performance Assessment Battery (ARPAB)

Based on the pioneering work of Rolland (1995) and Ellis et al. (1995) on quantifying depth perception in AR systems, Kirkley (2003) developed the ARPAB to investigate the ability to identify objects, judge distances, and estimate sizes of objects in Augmented Reality (AR). Of these, the task of judging distance is relevant to this dissertation's topic, direct object manipulation. For this task, Kirkley used real objects, 3-D models of those objects, basic 3-D shapes, and flat geometric shapes. Objects were displayed at 10 to 110 feet. Participants reported estimated distances. While this study is not as relevant to object manipulation (since it deals with action space not personal space), the technique of using different types of real and virtual objects are salient and is considered.

Precognition Performance Assessment Battery (Precog PAB)

Fidopiastis (Fidopiastis, Meyer, Fuhrman, & Rolland, 2004a; Fidopiastis, Meyer, Fuhrman, & Rolland, 2004b) developed a set of tasks, Precog PAB, for assessing visual display technology. Among the measures performed are static visual acuity, dynamic visual acuity, and depth perception. The Precog PAB work and its results could provide a comparison model for task performance in direct object manipulation.

Interaction Techniques

Bowman (1999) identified four task categories and a taxonomy of interaction techniques for measuring performance in VEs. The task categories are travel, selection, manipulation, and system control. Of these, selection and manipulation are relevant to this dissertation. Bowman also included the final task of object release for the manipulation category.

<u>Current Work In Object Manipulation</u>

The following sections describe various selected work related to manual manipulation of objects. Kitamura's categories of object manipulation are first introduced. Then, a number of research efforts and applications are introduced, grouped into these categories, and discussed.

Classification of Object Manipulation Schemes

Kitamura (Kitamura, Tomohiko, Toshihiro, & Fumio, 1999) classified object manipulation into four categories (Table 2). The distinctions between the four are real or virtual nature of the hand, the tool, and the object. Kitamura's definition of real and virtual relates to physical existence.

Table 2. Kitamura's Object Manipulation Classification

Physical Items	Category A	Category B	Category C	Category D
Hand	Real	Real	Real	Virtual
Tool	Real	Real	Virtual	Virtual
Object	Real	Virtual	Virtual	Virtual

For direct, manual interaction with objects, Kitamura's classification scheme could be slightly simplified to accommodate current applications and research work that is discussed in the next sections. For this simplified scheme, the tool is grouped into the object class. This is done because for certain applications that need to be included in this work, there are no tools, virtual or real. Furthermore for applications where the tool and the object exist, it is the tool that is directly manipulated by the hand so the tool becomes the "object" of concern (Table 3).

Referring to Table 3, the first column lists the visual items. The hand or the object can be real (physically true) or virtual (computer models). For the first category, the environment is one of Reality (RE) where both items are real. The user, directly or through video, sees the physical hand and the physical object being manipulated. The second category is one of Mixed Reality (MR) with the real hand and the virtual object. The hand is seen directly or through video, but the object being manipulated is a computer model. The last category is one of Immersive Virtual Environment (IVE) with both the hand and the object being virtual. Both are seen as computer models. The categories, or columns, in Table 3 are also changed from generic alphabetic designations to descriptive types of Virtual Environments.

Table 3. Kitamura's Classification Scheme, Simplified

Visual Item	Category: RE	Category: MR	Category: IVE
Hand	Real	Real	Virtual
Object	Real	Virtual	Virtual

Applications or studies related to object manipulation are categorized below using the simplified scheme and definitions described above for the three categories, RE, MR, and IVE. A real hand or object means having direct optical view or video of the item. A virtual hand or object is seen as a computer representation, for example, an image rendered from a 3-D graphics model.

Object Manipulation in Reality Environment

For the first category of object manipulation, the environment is essentially RE where the seen hand is real, just as the seen object is real. A number of applications take advantage of the intuitiveness, high accuracy, and speed this technique affords. Advantages and disadvantages of video systems are discussed first and then Optical See-Through systems next.

Manipulation in RE via Video and Video See-Through Techniques

Video can be displayed on a screen or on a HMD. A number of medical applications, e.g. computer aided surgery (CAS), use this technique to indirectly view the area worked on through video instruments, e.g. a laparoscope (Birkfellner et al., 2002; van Koesveld, Tetteroo, & de Graaf, 2003; Wagner, Ploder, Enislidis, & Truppe, 1996; Wagner, Rasse, Millesi, & Ewers, 1997; Wanschitz et al., 2002; Wendt et al., 2003). The user generally looks to the side or up away from the work area to a monitor as the task is performed. Some implementations use head mounted displays or mirrors or both to minimize the displacement between the scene and the work area. The surgeons get a video-realistic view at the expense of mismatch between the hand position and the position of the image seen.

VSTHMD provides a way to fuse the real and virtual world and reduces the mismatch between the two. It allows the user to look in the same, albeit not the exact, direction as the position of the hands. The HIT lab at the University of Washington (Kato, Billinghurst, Weghorst, & Furness, 1999) applied this technique to select, visualize, translate, and rotate virtual models. The users manipulated physical plates with markers that were recognized and tracked by CCD cameras when detected. Virtual objects were rendered at these markers and the user perceived them as being attached to the plates. In effect, the user was indirectly selecting virtual objects by flipping appropriate real, physical plates so that the marker was detectable by the camera. The user was then able to visualize different perspectives of the virtual objects and arranged them by rotating, translating, and placing the physical plates anywhere in the environment. One application was a storybook that popped out 3-D images as each page was turned. Another application demonstrated was the visualization of buildings or furniture for planning layouts.

The HIT Lab's VSTHMD technique allowed for fusing the two worlds and for indirect manual manipulation of objects. However, the computer generated images were only approximately matched to the markers on the plates. High spatial registration accuracy was not a priority in these mainly visualization (as opposed to interaction) applications. In fact, computer generated images seemed to have jumped and popped in and out as the plate was moved or rotated at times, especially when the markers were unintentionally and temporarily or partially covered. Furthermore, even though the hand and the object were well matched to each other for the sense of vision, the VSTHMD did not provide a perfect match between vision and proprioception, which affected eye-hand coordination. Optical See-Through systems overcome this problem and are discussed next.

Manipulation in RE via Optical See-Through Techniques

A number of vehicle simulations exist whereby physical mockup of the cockpit is provided in addition to a projection display or an optical see through HMD. Unlike the case with VSTHMD, these provide a complete match between vision and proprioception since the user sees the real hands manipulating real objects. An example is the PC-based microsimulator (White, Wharton, Kotick, & Anschuetz, 2003), which contained a cockpit mockup and a minidome screen.

One of the more grandiose systems of this type employs the CAVE. Lehner and DeFanti (1997) from the University of Illinois used stereoscopic glasses (Crystal Eyes) in a CAVE environment with a cockpit mockup. The cockpit was provided so that the users could interact with physical controls. Users had direct views of their hands interacting with real cockpit controls but saw the rest of environment in 3-D with stereoscopic glasses and the projection screens on the walls.

While the CAVE environment with stereoscopic glasses allows for 3D and depth perception for the VE, spatial registration in this and similar applications is not accurate enough for direct manipulation of cockpit controls, nor is it as critical, since the VE represented is generally farther away than personal space. The user manipulates physical controls instead of directly interacting with the VE.

Rolland et al. (Rolland et al., 2002a; Rolland et al., 1997), Wright, Rolland, & Kancherla (1995), and Yeo et al. (1999) also applied the optical see-through technique for medical visualization, which did require high precision in matching virtual and real objects. Using an optical see through head mounted projection display, the user saw physical objects that had position sensors attached to them. The computer generated image in turn was attached to the physical object being manipulated similar to the video see-through example with the plate and marker from the HIT lab. In one application this provided visualization of x-ray images or models relative to physical body parts for medical training. In effect, this allowed the user to visualize bone or joint movement inside the body.

This technique provided an excellent tool for medical visualization since the HMPD could be optimized for correct occlusion, convergence, accommodation, and stereo disparity. The effective image plane for the virtual object, for example, x-rays, could be accurately placed inside the physical object, for example, a mannequin’s leg.

However, similar to the case with VSTHMD and the HIT lab's work, the virtual object was manipulated indirectly through interaction with the physical object, the mannequin's leg. The virtual object was fixed to a specific real object for the purpose of visualizing different perspectives with accurate spatial relationship (Bailot, Rolland, Lin, & Wright, 2000; Argotti, Davis, Outters, & Rolland, 2002; Santhanam, Willoughby, Kaya, Shah, Meeks, & Rolland, 2008; Wright, et al., 1995).

Hua (Hua et al., 2002; Hua, Gao, & Rolland, 2003) applied HMPD technology to an AR board game application, which did not have to fix the virtual object to the real object. The application also required the user to perceive the position of virtual objects relative to physical objects precisely. A computer generated 3-D game board, similar to a chessboard, was projected onto a tabletop retro-reflective screen. The player wore the HMPD and moved real stones, similar to chess pieces, and was able to place the pieces precisely on the virtual board's reticules. The real game pieces appropriately occluded the virtual board. This research demonstrated the capabilities of augmenting the real environment with computer generated images, natural occlusion of virtual objects by real objects, and interaction in MR.

The above applications in RE showed that surgeons can perform highly accurate dextral tasks with off-axis visualization using video. New techniques provided additional ease of use with mirrors or HMDs. VSTHMDs are applied to other domains that demonstrated further reduction of the discrepancies between vision and proprioception. Optical see-through solutions using the HMPD completely eliminated this vision-proprioception discrepancy. The AR board game (Hua et al., 2002; Hua et al., 2003) further demonstrated how the HMPD could support accurate spatial tasks. The user saw the Virtual Environment (VE) with enough precision to place real stones onto virtual object reticules. These applications give evidence that precision may be achievable for spatial tasks interacting with real objects relative to virtual objects. Although, precision demonstrated above may indicate that such a task is feasible, by definition, the objects manipulated were real, not virtual. The next sections describe work related to manipulation of virtual instead of real objects.

Object Manipulation in Immersive Virtual Environment

The second scheme for object manipulation falls on the right-hand side of Milgram's Reality-Virtuality continuum, and provides the user with an Immersive Virtual Environment (IVE). Both the hand and the object to be manipulated are virtual, i.e., computer representations (see Table 3). A number of applications take advantage of the flexibility and reconfigurability of this purely virtual environment. Advantages and disadvantages are discussed below.

Beier (2000) produced car simulations using BOOMs or HMDs. The user was completely immersed except one hand was kept on a physical device, for example, the steering wheel or side grip on the BOOM. Applications of this type are typically anthropometrics, for example, reachability, to support analysis and design using virtual prototyping. A radio button, for example, is virtual and the driver has to demonstrate that the button can be reached comfortably. These applications demonstrate that virtual object spatial position can potentially be perceived with enough accuracy for evaluation of vehicle designs. Actually interacting with these virtual objects, however, would require highly accurate visual perception, especially depth perception, and was not covered in these car simulation applications.

Kirkley (2003) did address accuracy of visual perception, specifically depth perception. He used a HMD and a Retinal Scanning Display (RSD) in the VE and measured participants' depth perception. Kirkley found that users can judge distance most accurately with targets that are real objects, then 3-D models, then 3-D shapes, and finally, flat shapes. However, the work was relevant for 10 to 110 feet, which is basically action space, not personal space.

Surdick, Davis, King, and Hodges (1997) did conduct similar depth perception experiments in personal space. Surdick et al. found that the use of perspective (linear perspective, foreshortening, and texture) was the most effective for 1 meter to 2 meters viewing distance. In this specific application, effectiveness was related to the ability to perceive the change when the object depth was changed. This was compared with other apparently less effective depth cues of brightness, relative size and height, and stereopsis. The work was encouraging in that perspective cues could be perceived by all participants and were easily incorporated in computer simulation. However, while this work was within the realm of personal space, it was just beyond arm's reach.

Hu (Hu, Gooch, Creem-Regehr, & Thompson, 2002) conducted experiments to determine the effectiveness of stereopsis, shadows, and interreflections cues for conveying distance information within arms reach. Of these, stereopsis was found to be the most significant cue and by itself can provide both relative and absolute distance cues. Shadows and interreflections also provided significant cue but only for relative distances. The graphed data published by this work appeared to show that with stereopsis alone, accuracy was on the order of 25 mm and precision, 15mm. The experiments were conducted with nominal Interpupil Distance (IPD) settings so that stereoscopy cues provided were not perfect. This could account for as much as 7mm of the error seen.[1]

Rolland, Meyer, Arthur, and Rinalducci (2002b) and Rolland, Gibson, and Ariely (1995) conducted depth perception experiments using stereoscopy as the only cue but did so with the correct IPD setting on an optical bench. The computer rendered virtual objects of different shapes side by side. Accuracy (or average error) and precision (or standard deviation of error) of perceived depth were measured to be approximately 2mm and 8mm, respectively. Effects of eyepoint location in HMDs were also investigated (Rolland, Ha, & Fidopiastis, 2004). This and other studies in personal space discussed above dealt with perception only. The next set of work discussed below considers the interaction aspects.

von Wiegand et al. (1999) used Crystal Eyes Shutter glasses and a haptics device, the Phantom, in a setup called the Virtual Workbench. The set up was such that the user looked at a monitor with the shutter glass displaying stereoscopic images. The hand and the haptics device were not visible to the user, but were visually represented by the computer. The task was to locate the virtual probe on nodes on an electronics circuit board. Accuracy was determined to be within +/- 5 mm towards the center and +/- 10 mm towards the edge.

---

[1] IPD used to generate the computer graphics was fixed at 65 mm. The test subject's IPD ranged from 55mm to 70mm. Therefore, IPD error can be as much as (55-65)/65=10%. Maximum final distance between the objects was 70 mm in the vertical direction, which is the direction measured. Error due to incorrect IPD, therefore, may account for as much as 10%*70mm=7mm.

Arsenault and Ware’s (2000) fish tank VE also employed shutter glasses and a phantom haptics feedback device. This experiment involved the classic Fitt’s tapping task whereby participants tapped back and forth between two targets and the time intervals between taps were measured. This task is of interest to this dissertation because it required the eye-hand coordination skill that is similar to pushing buttons in a 3-D environment. Experimental results showed that having a true visual perspective from head tracking improved performance time by 9% and haptics feedback by an additional 12%. Both Arsenault’s fish tank and Von Weigand’s Virtual Workbench require the hand to be constantly fixed on the haptics device. This limited their applications. For example, localization of objects outside the device’s small range was not feasible.

Latham (1998) constructed an IVE that spanned the entire length of both arms’ reach. This VE provided for interaction with knobs, dials, and buttons that were physically real, but were visually virtual. The user wore an opaque HMD and saw a virtual 3-D model of the controls, but the hand felt the actual controls. This system was called Touched Objects Positioned In Time (TOPIT $_{tm}$). In this application, a servomechanism device placed actual physical controls in their correct positions, as the immersed user translated the hand while observing its virtual counterpart reaching for the virtual controls. The visual display provided a computer generated stereoscopic image of both the user’s virtual hand and the virtual instruments. The system was designed for virtual prototyping applications such as evaluating different cockpit panel designs.

The depth perception experiments and virtual environment test beds described above provide evidence supporting the feasibility of manual manipulation of virtual objects. Empirical data for precision and accuracy of depth perception were quantified and were in the sub-centimeter range. From the action or interaction perspective, it is interesting to note that both the virtual workbench and the TOPIT incorporated and emphasized haptics feedback. In fact, direct manual manipulation of objects without haptics feedback is rare in the literature. One of these experiments (Schiefele, 2000) produced results that indicate that haptics is necessary for IVE and is discussed in the next section.

Comparison of RE with IVE Schemes for Object Manipulation

Lok (2002) conducted a study that compared different techniques for the simple task of arranging blocks in AR. For one case, the participant manipulated real objects while looking at a video of the hand and the object (basically RE condition). For the other extreme the participant saw a virtual representation of the hand interacting with the virtual block (basically IVE condition). There were also hybrid cases in between where the real hand was disguised or otherwise visually represented differently (basically approximating MR condition). The results showed that RE is most effective, followed by hybrid (a step towards MR), then IVE. Lok concluded that training and simulation VEs would benefit from having real, instead of virtual, objects if they are to be manipulated.

Schiefele (2000) and colleagues at the Technical University Darmstadt developed a virtual cockpit simulation for the Airbus A340 commercial aircraft and conducted a similar experiment. The experiment had three conditions: RE, IVE, and IVE with simple haptics feedback. The tasks were pushing a button, turning a dial, and flipping a switch. For the RE condition, the user directly saw the hand and the controls. For the IVE condition, physical controls were not present. The user wore an opaque HMD and saw virtual controls as well as a computer model of the hand being tracked. Finally, the third condition was the same as the IVE one except a plastic panel was placed where the virtual controls were supposed to be. Experimental results showed that on average, for task completion, RE took about 1.5 seconds, IVE with haptics feedback, 3.5 seconds, and IVE alone (without hatpics feedback), 5.5 seconds. These results reinforce VE design guidelines for making as many objects real as possible (Lok, 2002; Lok, Naik, Whitton, & Brooks (2003)) and reaping the benefits of having touch feedback (Burdea et al., 1994) from the added physical surface (Lindeman, 1999; Burdea et al., 1994). Schiefele also found that it was almost impossible to adjust and localize to a virtual object in IVE without any haptics (touch) feedback and concluded that it was absolutely necessary.

The research above compared RE and IVE techniques. Each effort gives evidence and preference for the fidelity and performance afforded by RE techniques. However, the portability, reconfigurability, and cost savings provided by the IVE technique drove researchers to enhance IVE toward the performance of RE techniques. Schiefele found a simple haptics solution to improve the IVE technique, literally half way there, from 5.5 seconds down to 3.5 seconds with the ultimate goal of matching the 1.5 seconds benchmark in the RE condition. The next section discusses MR techniques, exploring that optimal middle ground and exploiting techniques that combine the fidelity of RE with the portability and flexibility of IVE.

Object Manipulation in Mixed Reality

This section takes a glimpse at a potentially optimal approach, MR with the Real hand manipulating Virtual objects (R/V) (see Table 3). These efforts begin to explore environments whereby direct view of one object, e.g., the hand, is provided next to a computer representation of another object, for example, virtual button. The literature abounds with virtual object manipulation but more so with indirect or unnatural methods. There are a few experiments in depth perception that provide evidence for the feasibility and effectiveness of MR environments. There are also prototype applications that have been published, albeit without experimental data. These MR works are discussed.

Rolland (1995) conducted experiments with real and virtual objects set side by side to measure depth perception by stereopsis. Experimental results indicated that the average measured shifts in perceived depth of virtual objects with respect to real objects were on the order of 50 mm. Virtual objects were seen farther away than real objects when both objects were presented at the same depth. A number of artifacts as noted by Rolland (1995) may have contributed to this discrepancy between the perceived depth of real and virtual objects. This includes optical distortion in the display, change in IPD of users and convergence not accounted for, unrealistic illuminance, and collimation of the images. Also, the methodology used – the method of constant stimuli, a standard in psychophysics – may not have yielded the most stable results because of the handling of size as a cue to depth (Rolland, et al., 2002b). The next experiments discussed below do account for some of these potential error factors.

Ellis and Menges (1997, 1998) conducted several extensive experiments whereby participants adjusted a real-object probe to match the distance of a nearby virtual or real object via a carefully designed Optical See-Through Head Mounted Display (OSTHMD). Monocular, biocular, and stereoscopic viewing conditions were used. For the biocular condition, the image presented was one that would be seen if the camera were between the real eyes. They were offset laterally so that convergence was correct. Experimental results showed that depth precision was 1.5 mm for RE and 3.0 mm for MR for the stereopsis condition. Accuracy was almost perfect for binocular (stereopsis) viewing. Accuracy was also near perfect for biocular viewing, which may be attributed to correct convergence setting. For monocular viewing, the distance perceived dropped back towards (not to) the physical wall behind the virtual object about 2.2m away, since no depth cues were provided.

Aside from the quantitative results above, a few findings are worth mentioning. First, Ellis and Menges (1997, 1998) suggested that monocular results may be explained by the “specific distance effect”, which is associated with tonic accommodation and vergence which relax to approximately one to two meters in the absence of distance cues. Second, a related but surprising finding is that all participants confirmed that the apparent distance of the monocularly viewed virtual object appeared to be driven around by the physical cursor. That is, where the real object was, and correspondingly where the user’s eyes were accommodating and converging to, affected where the participant thought the virtual object was when there are no other depth cues were present. This also confirms Ellis et al.’s (1995) earlier findings that judged distance of nearby objects is associated with changes in ocular convergence. Finally, the most striking phenomenon was that a monocularly presented object, which has no depth information but is placed next to the physical cursor was reported by all users to appear to have a definite depth (Ellis & Menges, 1997, 1998).

In addition to perceptual work, there have also been efforts with interaction in MR, albeit the manipulation is more indirect than direct. That is, hand gestures, as metaphors recognized by the computer, control the virtual object, not direct, natural action. One example is the responsive workbench and the work of Kruger (Kruger et al., 1995). This environment used Crystal Eyes Stereoscopic shutter glasses, Pohemous trackers, and Cybergloves. The bench was built with a projector, a large mirror, and a special glass plate as a tabletop. Two- and one-handed gestures were used, for example, to pan or zoom and did not require high accuracy of matching the real hand with the virtual objects.

Leibe (Leibe et al., 2000) went a step further than the responsive workbench discussed above by providing an untethered interaction without the use of any input device held in the hand. A camera-based system was used to recognize 3-D and 2-D hand gestures so the user could interact with the virtual object. This was more effective than glove interfaces, which caused subtle changes to the recognized hand gestures, affecting precision of fine manipulation. Although this technique was effective for its application, as with the responsive workbench, it relied on indirect hand gestures, not direct, natural, manual manipulation, for interacting with the virtual object such as those required for virtual buttons, dials, and switches.

The Research Gap – Direct, Manual Manipulation of Virtual Object

As shown in the previous sections, the literature abounds with the advantages of RE, i.e., direct view of the physical hand and the physical objects, for manual interaction (Birkfeltner et al., 2002; Lehner and DeFanti, 1997; White et al., 2003). The accuracy of the visual cues and the empirical evidence of task efficiencies (shortest time to complete correctly) are undisputable as compared with other techniques discussed. However, the physical environment, as discussed in Chapter One, would have to have a large footprint to house all the objects the user must interact with in personal space. One exception is the use of a physical tool to indirectly interact with the virtual world, but this technique is not as natural and is unacceptable for applications such as training where eye-hand coordination is crucial.

The literature also reveals evidence that a purely virtual environment can provide for effective interaction in certain applications (Beier, 2000; Latham, 1998; Lok, 2003). For manual interaction with virtual controls such as a vehicle cockpit, however, it is clear that haptic displays have to be added. Even with haptics, efficiency is still far from the ideal performance of manipulating real objects. Haptics displays, obviously, added more footprint as well as rigidity to the VE.

For the MR case, the literature is sparse. Visual perception experiments for this case reveal that current technology could provide the accuracy necessary for manipulating virtual cockpit controls (Ellis & Menges, 1998; Rolland et al., 2002). However, current implementations of these MR are somewhat limited to gesture recognition instead of direct manual interaction (Kruger et al., 1995; Leibe et al., 2000).

Figure 3 summarizes the literature discussed previously and where they fall in the simplified object manipulation scheme discussed earlier. A new column, “Research or Applications”, was also added. Under this column, manual manipulation refers to direct interaction with the object. Visual perception refers to work measuring depth perception only, i.e., no interaction. Finally indirect control refers, for example, to gesture recognition used to indirectly manipulate virtual objects.

Research or Applications	Category: RE	Category: MR	Category: IVE
1. Manual Manipulation	Birkfellner (2002) HIT Lab (1999) White (2003) Lehner (1997) Rolland (1997) Hua (2002) Lok (2003) Schiefele (2000)	Reseach Gap	Beier (2000) Von Wiegand (1999) Arsenault (2000) Latham (1998) Lok (2003) Schiefele (2000)
2. Visual Perception	Kirkley (2003) Ellis (1997)	Rolland (1995) Ellis (1997)	Kirkley (2003) Surdick (1997) Hu (2002) Rolland (2002)
3. Indirect Control	Not Reviewed	Kruger (1995) Leibe (2000)	Not Reviewed

Figure 3. Mapping of Current Research to Kitamura's Object Manipulation Schemes.

Literature related to indirect control was not searched and reviewed for all categories since it is not relevant for this dissertation. Referring to Figure 3 above, the only area that has generally not been explored in the literature is direct interaction with virtual objects using the directly seen, real hand. This research gap, from the discussions above, has the potential to provide for effective object manipulation with the portability and flexibility provided by MR.

CHAPTER THREE: HYPOTHESES AND METHODOLOGY

A discussion leading to the research methodology and hypotheses is provided in this chapter. The experimental apparatus centers on providing highly accurate cues to support effective and direct interaction with virtual objects. Because VEs provide two-dimensional images that can not perfectly provide for all three-dimensional, binocular cues, visual displays are carefully considered for the experimentations. Other sensory modalities are added and effectively implemented in the test bed design to improve task performance. Effective virtual object interaction techniques and research methods are judiciously selected to build upon previous works while narrowing the research gap in direct interaction with virtual objects.

This chapter begins with the manner in which the findings in the literature drove the formulation of this dissertation. Based on this, the research questions and hypotheses are then introduced. A test bed that supports the research questions is depicted. Finally, experimental designs that test these hypotheses are described.

Implications of Literature Reviewed for Direct Object Manipulation Research

Based on the literature reviewed in Chapter Two, this dissertation draws on a number of points to shape the proposed research questions, hypotheses, test bed development, and experimental design. This section starts with guidelines drawn from the literature reviewed on vision, haptic, audition, and proprioception. Visual display technology comparison follows. Next, the tasks relevant to virtual object manipulation, based on previous works, are also suggested. Then, this dissertation points to empirical evidence (from Chapter Two) and suggests that direct virtual object manipulation, although not yet attempted in the literature, may be feasible. In this section, the author also provides one calculated estimate using the Model Human Processor (MHP; Card, Moran, & Newell, 1983) to predict performance time for object manipulation in RE compared with those in IVE and MR.

Guidelines for Research Considering Sensory Modalities

Special attention must be paid to design and use of visual display systems. A number of conclusions can be drawn from the findings in the literature (Chapter Two) regarding vision relevant to the test bed design and experimentation for this dissertation.

- Optical alignment of display systems must be checked or calibrated.
- Optical focus should be set at or near the plane of computer generated images. Focus settings that are either too close or too far could cause accommodative adaptation with prolonged exposure; for example, it should not be set closer than reading distance of about 250mm or beyond infinity.
- To reduce risk of neural adaptation, focus setting must provide a sharp image.
- Image plane and virtual 3-D objects must be placed strategically so as to minimize unnatural vergence and accurately to avoid mismatched vergence.
- Binocular displays should be used only if a biocular solution is inadequate.
- Exposure time must be limited based on the type of display used.
- Visual stress should be measured and monitored.
- Acuity tests and vergence tests should be performed to assess visual aftereffects.

- For binocular displays, depth of field (or equivalently change in vergence) must be limited to a comfortable range depending on viewing conditions.

Additionally, Vestibular-Ocular Reflex (VOR) adaptation discussed in Chapter Two should also be considered carefully. For this dissertation, total elimination of the discrepancy between the vestibular and visual cues would require a perfect head tracking system with zero latency and a frameless image generation system (essentially instantaneous scene refresh). While the task of manipulating buttons, dials, and switches requires only minimal and slow head movement, the possibility exists that some VOR adaptation will occur and will contribute negatively to task performance.

The literature reviewed in Chapter Two indicates strong evidence for effectiveness of VEs that provide for multi-modal presentation, specifically, vision, audition, haptics, and kinesthetic proprioception to support direct manipulation of objects. These findings also suggest that there is a multidimensional trading space for substituting or complementing one sensory for another when certain cues are absent or are lacking.

Comparison of Display Technology for Direct Object Manipulation

Direct manipulation of an object requires precise knowledge of its location in all three dimensions. Of importance and more difficult to provide and calibrate accurately is depth. Based on Cutting and Vishton’s model (1995), the most important cues for depth perception in personal space are occlusion (closer objects visually blocking farther ones) and binocular disparity (stereopsis). The next important cues are relative size and motion perspective, which have been discussed previously as cues that are easier to implement in all visual display systems. After these, the next most important are accommodation and convergence.

So, aside from motion parallax and relative size, which are not issues for visual displays, the remaining four cues that require careful attention are occlusion, stereo, accommodation, and convergence. As Cutting and Vishton noted, occlusion is the most important and is the most effective depth cue for objects in any space. Ellis and Menges (1998) also found that task errors in VE can stem from missing or incorrect occlusion cues. Stereo is the second most important. Depth through stereo can be achieved without accommodation and vergence, as is the case with autostereograms, e.g., random-dot stereograms (Liu, Stark, & Hirose, 1991). Finally, correct accommodation and convergence are important not only for depth perception but also for reducing visual load (Shibata, 2002; Wann et al., 2002; Ellis & Menges, 1997). Visual display systems used for VE applications in personal space must carefully consider these four cues.

For direct manipulation of objects, this dissertation draws from the literature two other crucial factors besides depth perception – correction of optical distortion and synchronization of visual and proprioception modalities. Robinett and Rolland (1992) claimed that one of the most common visual errors in VE is ignoring the distortion caused by the optics. Rolland et al., (2000) further suggested that simulation developers using HMDs may not have access to the distortion function. Even if they do have the function, performance penalties may prevent them from correcting for the distortion. Regarding synchronization of vision and proprioception, eye-hand coordination is essential for reaching, grabbing, and manipulating objects. Accordingly, having instantaneous visual displays and proprioception cues eliminate dissociations of seen and real hand displacement that could induce sensory rearrangement (DiZio et al., 2002). Likewise having a direct view of the hand could reinforce and enhance these two modalities as Popescu et al. (2002) has found for vision and audition.

Finally, support for multiple users is also an important feature for a wide variety of applications, such as team training in VE. This dissertation also includes this capability as a desirable feature for comparing visual display technologies.

Table 4 lists the display technologies discussed in Chapter Two with key capabilities that support virtual object manipulation. The visual display capabilities most salient to near-field depth perception and natural, direct manipulation of virtual objects are listed in the first row of the table. These include Correct Occlusion, Stereoscopy, Correct Accommodation, Correct Convergence, Minimal Optical Distortion, Proprioception and Visual Synchronization, and Multi-User Support. Based on the discussions above, this dissertation suggests that HMPD technology can provide all these cues simultaneously and with the high perceptual fidelity necessary for virtual object manipulation in personal space.

Table 4. 3-D Display Technology and Characteristics for Virtual Object Interaction

		3D Display Characteristics						
Virtual Reality Continuum	3D Display Technology for virtual object Interaction	Correct Occlusion	Stereopscopy	Correct Accommodation	Correct ConVergence	Minimal Optical Distortion	Proprio-Visual Sync	Supports Multi-User
	Video Screen (2D)	X		X	X			X
Reality	Video See-Through (2D)	X		X	X			X
	Optical See-Through HMD		X	X	X		X	X
	Video See-Through HMD	X	X					X
	CAVEs	X	X				X	
Mixed Reality	Virtual Workbench	X	X				X	
	Autostereoscopic Display	X	X	X	X		X	
	Volumetric Display	X	X	X	X		X	X
	Virtual Retinal Display			X	X	X	X	X
	HMPD	X	X	X	X	X	X	X
Virtual Environment	Opaque HMD	X	X	X	X			X
	BOOM	X	X	X	X			

Tasks For Direct Object Manipulation

There is some overlap among the four sets of tasks reviewed in Chapter Two for VE evaluation. Each of the four sets, the VEPAB (Lampton et al., 1994), the ARPAB (Kirkley, 2003), the Precog PAB (Fidopiastis et al., 2004a; Fidopiastis et al., 2004b) and the Interaction Techniques (Bowman, 1999, 2002), could be used to assess certain aspects of the task of direct manipulation of virtual objects. A combination of these tasks is appropriate for this dissertation. Kirkley's ARPAB and Fidopiastis' Precog PAB essentially emphasize perceptual tasks, one of which is distance estimation. Lampton's VEPAB and Bowman's Interaction Techniques include interaction tasks, one of which is of particular interest for this dissertation – the manipulation task. This dissertation chooses to explore both types of tasks: Depth perception and object manipulation.

Why Explore the Research Gap?

Drawing on the literature review in Chapter Two, this section explores potential advantages of MR with the real hand manipulating virtual objects such as those found in a virtual cockpit.

First, Lok (2003) showed that interaction using real objects and video is more effective than interaction using avatars, i.e., virtual representations of the hand and the object. As a result, he suggested that designers develop VEs that provide for more physically real objects. Bowman (2002) generated guidelines for VE interaction techniques and also suggested using natural hand techniques and direct manipulation instead of using tools for manipulation of objects within arm's reach. Popescu et al.

(2002) also suggested that eye-hand coordination was important in direct manipulation tasks. DiZio et al. (2002) further suggested that virtual-visual displacement can cause unwanted adaptation of proprioception. This dissertation draws from these works and suggests that perhaps, the hand, the most important object in the VE, should also be real in order to improve effectiveness of spatial task performance and minimize adaptation over purely virtual environments. Having a direct view of the hand and therefore full fidelity for the visual perception could provide better accuracy, shorter task completion time, and less adaptation problems.

Next, from the literature cited previously, it is clear that correct occlusion, stereoscopy, accommodation, and convergence are highly desirable cues for object manipulation but are often omitted or else incorrectly provided due to limitations of conventional visual display technology. Cutting and Vishton's model (1995) also identified these cues to be among the most important in personal space. Based on the works of Ellis & Menges (1997, 1998), where it was shown that transparency in virtual and real objects creates the false depth illusion, this dissertation suggests that proper occlusion could eliminate this error completely. Also based on the works of Ellis and Menges (1997, 1998) and Rolland et al. (2002b), where they independently showed that stereoscopy alone provided for sub-centimeter accuracy and precision, this dissertation also suggests that this additional cue provides adequate precision for effective cockpit instrument interaction. Also based on experimental results of Ellis and Menges (1997) for biocular viewing, it is suggested that correct convergence and accommodation settings also contribute to accurate depth perception, and more importantly, reduce eyestrain.

From a quantitative standpoint, Ellis and Menges (1997) reported experimental results of 3.0 mm precision and perfect accuracy for depth perception of real and virtual objects side by side with binocular viewing alone. Since typical cockpit buttons, dials, and switches are at least three times larger than this tolerance of 3 mm, it appears that such an application is feasible. This does not yet take into consideration the findings of Ellis and Menges (1997) regarding adjacent physical objects. Having a physical object, like the finger, near the virtual object, like a virtual button, may provide even more realism and more perceptual accuracy in the judgment of virtual object depth. This also has not accounted for additional cues such as haptics, which Schiefele (2000) has suggested can reinforce visual cues as well as increase usability by providing a physical surface on which to rest the finger or the hand. Therefore, it appears that there is ample evidence suggesting high effectiveness in a well-designed MR where the hand is real and directly seen and the object is virtual.

Finally, this dissertation also suggests that direct view of the user's hand provides for a more efficient environment than a purely virtual environment based on the Model Human Processor (MHP). This model provides for prediction of performance time relatively accurately, especially for simple tasks (Eberts, 1999). The model takes into account perceptual processing time, Tp, of about 100 ms on average, cognitive processing time, Tc, of about 70 ms on average, and motor processing time, Tm, of about 70 ms on average. Based on this model and conservative assumptions, estimated times for task completion were calculated (Appendix D). Simple button pushing performance time was predicted to be on the order of 2.0 seconds for the RE condition, 2.4 seconds for the MR condition, and 3.0 seconds for the IVE condition. These estimates show noticeable improvement of MR over IVE and approaching the ideal condition of RE.

Research Questions

The works related to object manipulation described above provide ample evidence that an intuitive and effective MR could be designed to provide for direct manipulation of virtual objects. This dissertation's preliminary analysis also suggests that providing a direct view of the hand, for example, in MR, can be more effective and more efficient than displaying a corresponding model, for example, in IVE.

From a review of display technology, OSTHMD, HMPD, and RSD are the only technologies that provide direct optical view of the real world. Of these, neither OSTHMD nor RSD provide for correct occlusion, which has been shown (Ellis & Menges, 1997) to significantly affect visual perception (see Table 4). HMPD, on the other hand, has this and all other features previously identified as potentially beneficial for virtual object manipulation.

Review of the literature regarding vision also reveals many potential pitfalls with stereoscopic displays, likely due to improper design and use, rather than fundamental limits. Biocular (non-stereoscopic) displays are generally easier to implement, are more forgiving to inappropriate use, and place less stress on vision as a result, but lack depth cues. Stereoscopy, on the other hand, has been found to provide strong depth cues, but requires careful design, calibration, and use to avoid inaccurate and conflicting cues. Specifically, accommodation and convergence are cross-linked and mismatches of these cues with each other, with stereoscopy, or with other depth cues can contribute significantly to eyestrain.

Drawing from the summaries above, an overall research question is formed for this dissertation. Can one naturally, effectively, and directly interact with virtual objects in personal space? This is broken down into two pairs of research questions. Each pair includes one question dealing with depth perception and the other dealing with interaction. The first set of questions is carefully designed to address extreme ends of the Reality-Virtuality (R-V) continuum. This establishes solid reference points in the familiar RE and IVE, environments where ample research has already been conducted (see Figure 3). These questions seek primarily to confirm previous conclusions in the literature, slightly extending upon their findings, and linking them to this dissertation's construct. A second purpose is to provide solid reference points and establish robust experimental procedures for similar but more complex research.

The second set of research questions directly addresses the research gap (see Figure 3), again exploring depth perception and object interaction. It explores the full R-V continuum, spanning from RE, to MR, to IVE. In addition to the fundamental visual cues, other salient sensory modalities (audio and haptics) are included and controlled. All combinations of environment types with sensory modalities are investigated and their effects compared. This set of questions seeks to substantially extend the body of knowledge in virtual environment interactions by providing a full set of empirical data covering the R-V spectrum and sensory modalities.

Question One: Depth Perception in Immersive Virtual Environment (IVE)

What is the accuracy of perceived depth of virtual objects in personal space? This question is broken down further. First, what level of accuracy of depth perception is achievable with standard Optical See-Through Head Mounted Displays (OSTHMDs)? OSTHMDs have optical distortion. Furthermore, accommodation and interpupil distance (IPD) adjustments settings are inexact. Measuring depth perception using OSTHMD empirically provides a baseline of current capability. Second, is this level of accuracy significantly enhanced when using a display system and software design that carefully considers correction of optical distortion, matching of field of view and IPD between the user, the optics, and the graphics software, correct focal point (accommodation), correct convergence, correct disparate images, and proper spatial registration, alignment, and calibration? These are salient to a number of the cues that Cutting and Vishton (1995) suggested are most important for depth perception in personal space.

Question Two: Interaction in Reality Environment (RE)

Can one interact with virtual cockpit panel controls as timely and as precisely without haptics feedback, given complete and perfect visual cues? Based on previous works conducted by Lok (2002) and Schiefele (2000), haptics could be expected to enhance task performance. However, the visual cues in Lok's and Schiefele's experiments were not ideal. Measuring task performance in RE establishes the reference point and quantifies the contribution of the second sensory modality, haptic cue, in a perfect visual environment.

Question Three: Depth Perception in Mixed Reality (MR)

Given correct binocular visual cues and kinesthetic proprioception, what is the accuracy of perceived depth of virtual and real objects? This primary research question is broken down into detailed questions that provide further insight into the contribution of different VEs and the cue of kinesthetic proprioception to depth perception. Does the MR environment significantly enhance accuracy of depth perception over the pure IVE? Ellis and Menges (1997) found that having a real object next to a virtual one does affect depth perception. Does the addition of kinesthetic proprioception significantly enhance the precision of depth perception? Kinesthetic proprioception provides a spatial cue that is fully synchronized with the visual cue when using an optical see-through system, such as the HMPD. Finally, what is the accuracy of depth perception for RE and is it significantly different than that for the MR or IVE with or without synchronized visual and proprioception cues? The cues provided in all environments are essentially identical except for the fact that in MR and IVE, the 3-D object is computer generated and produced from 2-D displays.

Exploring these questions on one common test bed with one set of experiments that span all virtual environment types can perhaps link findings from previous, separate research that dealt with only one type of environment each (Rolland et al., 1995, 2002; Ellis et al., 1995). With the additional proprioception cue, this research can also extend previous research and perhaps partially explain the accuracy gap between virtual and real environments. Finally with a comprehensive look across all virtual environment types, using a highly accurate HMPD calibrated to each individual, perhaps such empirical findings can expand our understanding about which cues or techniques potentially contribute to enhancement of depth perception. Figure 4 expounds on the research gap diagram formulated earlier in Chapter Two (Figure 3) and highlights the key area in VE depth perception that this dissertation addresses.

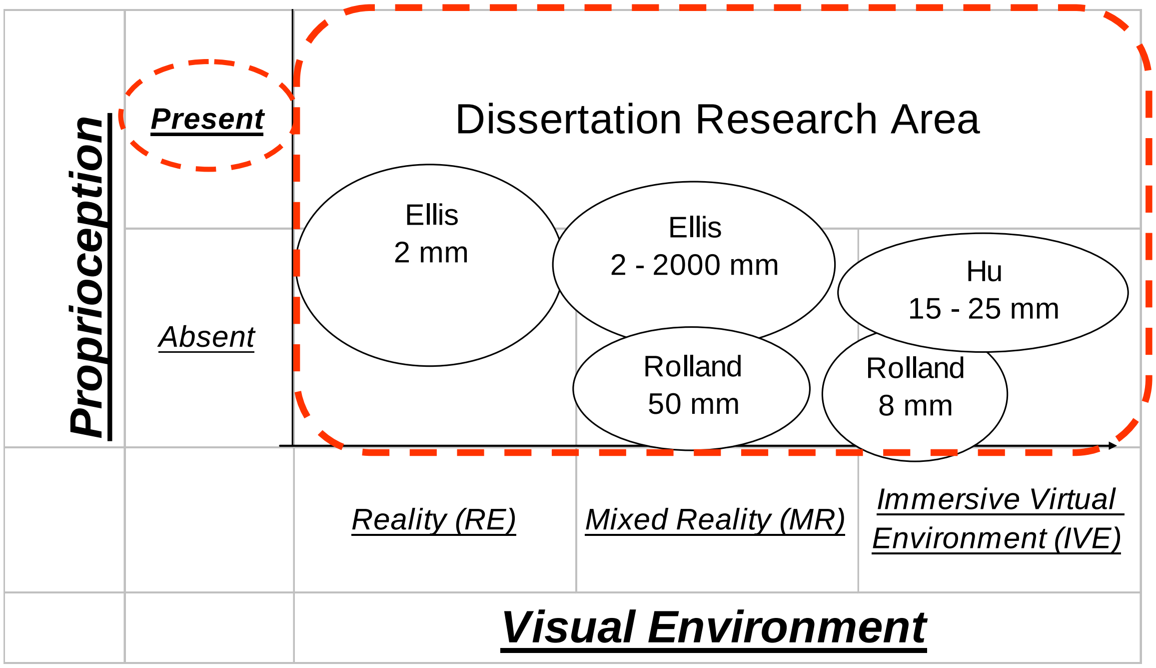


Figure 4. Dissertation Research Area, VE Depth Perception in Personal Space.

Question Four: Interaction in Mixed Reality (MR)

This research question is more practical from a flight training simulation standpoint. It deals with interaction with virtual and real objects. Can one manipulate typical virtual controls as effectively, i.e., as accurately, and as efficiently, i.e., as timely, as one does real counterparts?

This primary research question can also be broken into several detailed questions to assess the contribution of each modality on task performance. First, given all the proper, synchronized visual and kinesthetic proprioception cues in a MR, what are the performance time and error rate of the task of interacting with simple virtual cockpit controls? A MHP calculation (see appendix D) along with previous work (Schiefele, 2000) provide estimates for performance time ranging from 1.5 seconds to 5.5 seconds.

Second, do these performance measures improve with the addition of audio cues? That is, what is the contribution or effect of adding audio cues for this task? Audio provides additional temporal information so it could potentially contribute to performance time enhancement.

Similarly, given visual cues and simple touch haptic cue, what are the performance time and error rates for this task? How do they compare to the MR condition where only visual cues are provided? That is, what is the contribution of adding the additional haptic cue for this task? Haptic cues provide spatial information so they can potentially improve performance with respect to task error (Lok, 2002; Schiefele, 2000; Arsenault & Ware, 2000).

Fourth, given all feasible cues in MR – visual, audio, and haptics – what is the performance time and error rate for the task? How does it compare to the MR condition where only visual cues are provided? That is, what is the contribution of simultaneously adding audio and touch haptic cues for this task? Since all the cues are provided, task performance in this condition can potentially be significantly faster (shorter performance time) and more accurate than the visual-only condition.

Next, how does performance in this MR, where all the visual, audio, and haptic cues are present, compare to that in RE? Since all cues are present in both conditions, performance differences between RE and MR may be less pronounced.

Finally, the condition with all cues present except haptics is of special interest because it could potentially provide for an effective MR environment without the bulkiness and rigidity of the physical cockpit mockup or of the haptics apparatus. For flight simulation training applications, does task performance degrade significantly in the MR environment without haptic cues compared to the Reality (RE) environment?

Figure 5 expands on Figure 3 from Chapter Two and depicts some of the previous works in virtual object interaction highlighting the area of research for this dissertation.

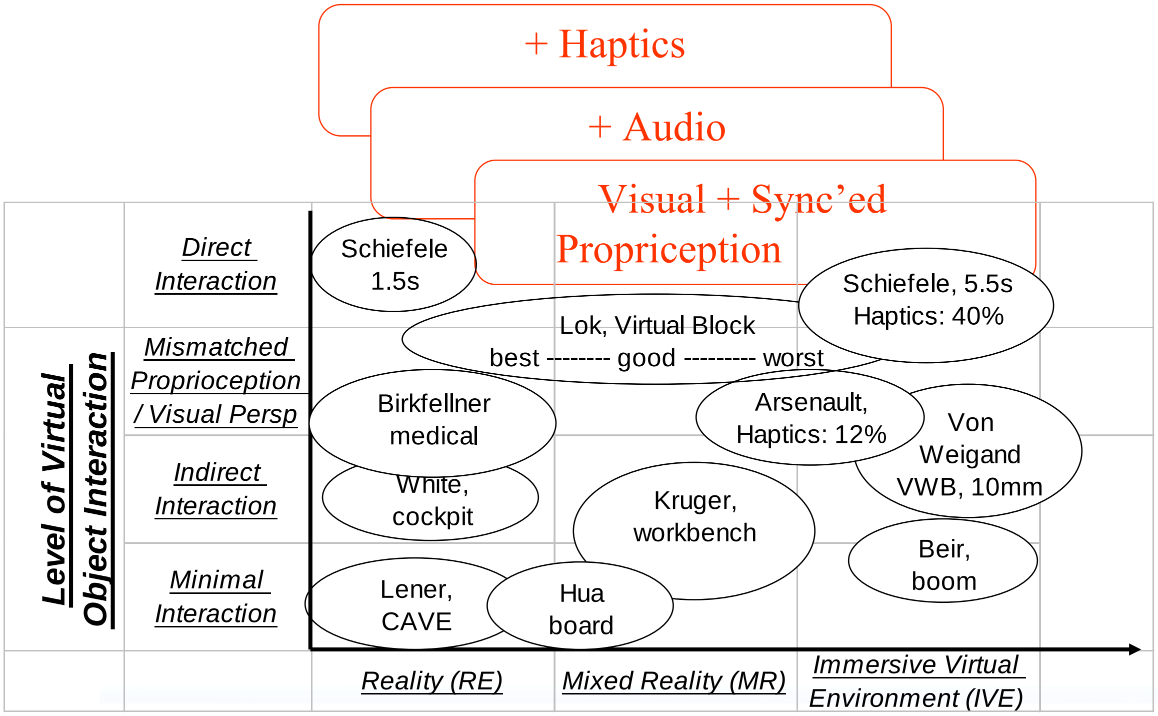


Figure 5. Dissertation Research Area, Object Interaction in Personal Space.

Experimental Design

Four experiments were designed to address the research questions previously discussed. The first two experiments explored the end points of the R-V continuum, depth perception in IVE and interaction in RE. The last two experiments also included depth perception and interaction tasks and explore the entire R-V continuum and all salient sensory modalities.

Apparati

The primary apparati used for collecting experimental data included VE equipment for the main data collection procedure and eye examination instruments for pre- and post- experiment procedures.

Visual displays for collecting experimental data included an nVision Hi-resolution Datavisor optical see-through HMD (OSTHMD) and a HMPD weighing less than 500g with Organic Light Emitting Diode (OLED) displays and provided by the Optical Diagnostic and Analysis (ODA) Lab at UCF. The optical see-through HMD and the HMPD had adjustments for head size and position and for IPD to match those of each test participant. Since the HMPD could be worn with prescription glasses, focus adjustment in lieu of the eyeglasses was unnecessary.

Image Generation hardware and software for the experimental test bed included a computer PC with Windows XP and a Matrox 256MB dual output graphics card. The computer simulation software used was DiSTI’s GL Studio, which provided a user-friendly Graphical User Interface (GUI) for creating 3-D cockpit models. The software application supported generation of disparate binocular views and provided for programmable adjustment of field of view, aspect ratio, IPD, and perspective to easily match those of the visual display with the experiment participants.

An Intersense 600 system was used to track the hand. This was a hybrid acoustics- and accelerometer-based system. The sonic discs (acoustics) were used to compensate for drift and provide absolute distance information. The inertial cube provided relative positional distance and angular rotation information.

An infrared detector with audio signal feedback was used to detect the presence of the hand touching the virtual object, thereby, breaking the infrared beam. Sony Digital Hi-Definition camcorder and HDTV were used for monitoring and, as a secondary method, to capture distance measurements during the experiment and to verify the results afterwards.

A Snellen Eye Chart, Dolman Depth Perception Box, and the Stereo Fly Test were used to measure each participant's static acuity, depth perception, ability to perceive 3-D via stereoscopy, respectively. These are used to screen out participants who had less than normal vision, and, as a safety measure, to look for signs of aftereffects if any.

Experiment One: Depth Perception in IVE, OSTHMD vs. HMPD

The first experiment measured depth perception in IVE and compared the results between two types of displays, the OSTHMD and the HMPD. The OSTHMD had visual cue settings at nominal values while the HMPD provided for more exact adjustments. Figure 6 depicts the two types of visual displays used in Experiment One.

Figure 6. Experiment One – nVision OSTHMD (left) vs. ODA Lab HMPD (right).

Experiment One: Hypothesis

Given correct binocular visual cues afforded by the HMPD, perceived depth of virtual objects is significantly more accurate than that afforded by the OSTHMD where visual cues are not ideal.

$$H_a\text{:}\ \ \text{Error}_{\text{HMPD}} < \text{Error}_{\text{OSTHMD}}$$

This hypothesis is mainly based on the preciseness of the cues provided by the HMPD compared to that of the OSTHMD. Based on previous VE depth perception experiments (Rolland, 1995), it is expected that the accuracy using OSTHMD — where accommodation and convergence settings can only be approximated and where optical distortion exists — would be on the order of centimeters (tens of millimeters). Also based on previous experiments (Rolland et al., 2002; Ellis & Menges, 1997), it is expected that visual display and software correctly aligned and calibrated to match the user would provide significant improvement in precision over OSTHMD. This design provides for comparison of the two cases, HMPD with optimal visual cues vs. OSTHMD with nominal cues, in one common experimental setting.

Experiment One: Design

The experiment had two treatments, one for each type of visual display. The design was based on work previously done by Rolland et al. (2002b) and Ellis and Menges (1997) but used the HMPD and off-the-shelf OSTHMD instead of laboratory optical display systems. The experiment measured depth perception in IVEs. The participants adjusted the position of one virtual object, an octahedron, until its depth is believed to match that of another nearby stationary object, a cylinder.

The following points helped to shape a robust experimental design and to plan for slight adjustments during the data collection phase as necessary due to unforeseen constraints, e.g., unanticipated apparatus limitation, participant variability, or erroneous assumptions discovered.

Within-Subject Design

A within-subject design was used to minimize the effect of variation due to individual participant differences. Depth perception using the method of adjustment (Rolland et al., 2002) of one virtual object to match that of the other was employed because of its proven reliability. Three object sizes were used to determine the effect, if any, of size on perceived depth. The lateral distances between the objects were proportionally wider for larger objects to maintain consistency. These lateral distances between the objects were also roughly comparable to the widths of the objects to reasonably balance out the scene.

Size Familiarity in Depth Experiment

Cutting and Vishton (1995) suggested that relative size is the fourth strongest cue in personal space, above accommodation and convergence. Since the participants may judge distance based on their familiarity with every day object size, this could introduce a confound in the experiment. Hence, the stimuli were chosen to be simple generic shapes, a cylinder and an octahedron. These do not relate to any physical object. The participants were told to base their alignment on the center of mass of each object. Accordingly, the volumes of the two objects were designed to be the same. The octahedron was also chosen (e.g., instead of a cube) because it had edges corresponding to the center of mass that can be referenced for alignment. The cylinder was also chosen to have a small diameter so that its center can be more accurately judged. Three object sizes were used to determine the effect, if any, of size on perceived depth.

Random Presentation in Depth Experiment

The octahedron, the object to be adjusted, was presented on either side of the cylinder, left and right, randomly and equally. Likewise, the octahedron was presented closer or farther away than the cylinder in a random and balanced manner. This also helped with subsequent analyses for determining whether dominant eye affected perception on one side or the other, or if front and back adjustment made any difference.

Optimal Settings and Mitigation of Side Effects.

In designing the experiment, a conscious decision was made to attempt to optimize all conditions. The HMD was set as closely as possible to correct accommodation, convergence, and IPD settings instead of nominal settings as is typically done. This strategy provided for comparisons of best cases instead of nominal cases for each type of display. Furthermore, this strategy minimized the chance of side effects due to typical, but improper, use of HMDs.

Population Sample

The ideal population for this study is military pilots. However, from a practical standpoint, a combination of active and inactive military personnel, civilian personnel, and graduate students were used. Additionally, since the population was not all pilots, the tasks for the experiments were designed, necessarily, to be simple and generic instead of more dynamic or more specific to specialized flight instruments.

Screening Tests

A number of pre- and post-experiment procedures were conducted. Participant Screening Tests were conducted prior to experiments. Participants were screened on acuity tests using the Snellen chart and with stereoscopy using the Dolman apparatus and the stereo fly test. Only participants with normal vision were used for data collection and analysis.

Calibration and Familiarization Sessions

Each participant for each session went through a calibration routine to align the HMD/HMPD with the eyes. Familiarization sessions were also conducted for each task. Each participant was allowed as much time as necessary to comprehend the task and explore the use of the equipment. For each experiment, one or two alignments taking about two minutes before the actual experiment started was generally sufficient.

Simulation Sickness Questionnaire (SSQ)

Side Effects Tests were conducted for each participant. Simulation Sickness Questionnaires (SSQs; Kennedy, Lane, Berbaum, & Lilienthal, 1993), acuity tests, depth perception tests, eye-hand coordination tests, and vergence tests were performed before and after the experiments to determine if aftereffects existed. Additionally, postural stability test was performed if necessary, since HMD-based devices could be expected to produce higher level of sickness compared with standard simulators (Kennedy, Dunlap, Jones, & Stanney, 1996; Drexler, Kennedy, & Malone, 2005). Each participant also completed a set of background questionnaires.

Lighting

The experiment was conducted in a confined booth completely covered with black linen so that no unwanted visual cues were present. The displays were relatively dim and therefore, the booth created more vivid stimuli by providing a dark environment.

Experiment One: Task and Independent Variables (IV)

The task for each participant was to adjust one virtual object, an octahedron, until the participant perceived that it was at the same depth as another nearby, fixed, virtual object, a cylinder. The task used the standard method of adjustments (Rolland et al., 2002b; Fidopiastis et al., 2004a, 2004b). The task was repeated multiple times. Independent Variables (IV) were visual display type at two levels, OSTHMD and

HMPD. The HMPD provided high visual fidelity. It was optimized for correct accommodation and convergence at 0.8 m, and had low optical distortion. The OSTHMD provided normal fidelity, typically found in high-end off-the-shelf displays. Accommodation and IPD were adjusted at nominal settings for each user. The OSTHMD also had some optical distortion that was unavoidable. The participant, with the help of the researcher, adjusted accommodation and IPD settings until the visual environment became relatively clear and centered.

Experiment One: Dependent Measures and Stimuli

This experiment included three dependent measures for each experimental unit, the participant. These were bias, accuracy, and precision as defined below for this study.

Error of depth perception was the primary quantity of interest. For each adjustment, the error was determined by measuring the final distance in depth (horizontally in the sagital plane) between the displaced octahedron and the fixed cylinder. It was arbitrarily defined as positive error distance if the octahedron was closer to the participant than the cylinder and negative error distance if farther away.

For each participant, these error distances were averaged over the repeated adjustments, to calculate the first dependent measure, bias. Bias can be positive or negative, since the error can be positive or negative. It is a measure of how far, on average, a participant overshoots, if positive, or undershoots, if negative, in perceiving depth.

Accuracy was the second dependent measure. For this study, it was defined as the average of the absolute values of the error distances. It was calculated for each participant by taking the absolute value of the error distance for each adjustment and averaging these values over all repeated adjustments. Accuracy can only be positive since it is the average of only positive values. It is a measure of average error of depth perception for each participant.

Precision was the final dependent measure. For this study, it was defined as the standard deviation of the signed (positive or negative) error distances. For each participant, the standard deviation of the error distances for the repeated adjustments were calculated. Although the error distances could be positive or negative, precision could only take on positive values since it was determined by calculation of standard deviation. Precision is a measure of the variability of depth perception for each participant.

The stimuli for the experiment were virtual objects, a cylinder fixed in position and an octahedron, which could be translated in one dimension, back and forth, by the participant. For each participant, the octahedron was on the left side of the cylinder in half the trials and on the right for the other half. It was alternated randomly from one side to the other. The depth of the fixed cylinder was placed at 0.8 m and the octahedron was randomly placed in front of and behind the cylinder and at different distances ranging between 0 and +/- 100 mm.

Experiment One: Participants and Procedures

Eight participants were recruited from pools of graduate students and researchers. As necessary, some were screened out due to failure of the eye exam. Each participant went through both treatments, OSTHMD and HMPD, in a within-subjects design.

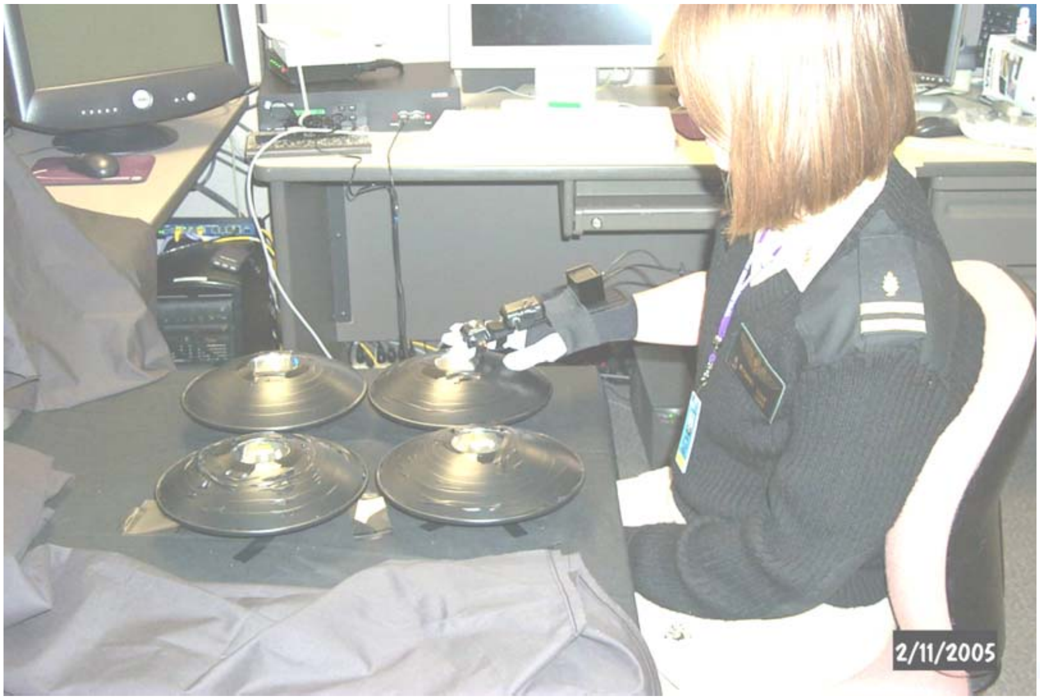
The initial depth of the cylinder was placed at 0.8 m. The participants repeated the adjustment task sixteen times with the octahedron on the left and sixteen more on the right. The adjustment task was repeated again with two other sizes of the same objects. Participants performed the task in an OSTHMD with approximated settings for accommodation and convergence. They also performed the same task wearing the HMPD where the experimenter set these accurately. The order of presentation of the octahedron on either side of the cylinder, the order of the visual display types used, and the order of the presentation of the sizes of the objects were randomized. The number of adjustments totaled 192 for each participant.

16 adj X 2 sides X 3 sizes X 2 Displays = 192 adjustments per participant

Each participant for each session also went through a calibration routine. Adjustments for IPD, focus, and convergence were necessary to match the HMPD and OSTHMD with the participants and the graphics software.

Experiment Two: Interaction in RE, Haptic Cue vs. None

Experiment Two measured accuracy and performance time of interaction in RE. It compared these measures for two conditions, one with a haptic cue present and the other with it absent. Figure 7 shows the participant's hand that was position tracked, the mirage dishes that presented the stimuli (optically-virtual images), and the removable, clear, plastic panels that provided for the simple haptic cue. The picture was taken with room lights on for clarity, but the experiment was conducted in the dark with only the black light turned on to illuminate the white glove and the florescent stimuli.



2/11/2005

Figure 7. Experiment Two – Object Interaction in RE, Haptics vs. None.

Experiment Two: Hypothesis

Error and performance time is significantly improved when haptic cues are provided, even in RE, where visual cues presented are perfect. This prediction is made primarily because haptics is a spatial cue, which can potentially improve spatial error as found by Schiefele (2000). It provides confirmation of action taken and can also potentially reduce task completion time as found by Arsenault and Ware (2000).

H_a: $\text{Error}_{\text{haptics}} < \text{Error}_{\text{no-haptics}}$

H_a: $\text{Performance Time}_{\text{haptics}} < \text{Performance Time}_{\text{no-haptics}}$

Experimental Two: Design

A within-subject design was developed. As discussed earlier, this minimizes the variation due to individual differences. In addition to some of the design consideration discussed earlier, below were additional concerns in development of Experiment Two.

Tracking Limitation

Tracking systems technology was still somewhat immature. This study, like many others employing HMDs and tracking systems, could have easily become highly dependent on the accuracy, alignment, and calibration of the tracking system. Position tracking was not the focus of this study. Therefore, tracking system was avoided as much as feasible. This experimental design limited the use of tracking systems. Only the hand was tracked. Tracking the head would have required complex alignment and calibration in six degrees of freedom for the head with the environment and would introduce additional, unwanted errors unnecessarily. For this experiment, optical mirages were used. Each optical mirage was generated by reflecting the real physical object, the button, off of two curved mirrors to present an optically real image. The set of curved mirrors, also called a mirage dish, displayed the stimuli with perfect visual cues no matter where the head was positioned so head tracking was not necessary. Hand tracking was carefully observed during data collection and raw positional and rotational data was analyzed in detail to account for any misalignment.

Experiment Two: Task and Independent Variable

The task was activation of a keyboard button, such as those found in flight cockpits or computer terminals. The first part of this experiment measured performance time and error rate. The second part was independent of the first and measured spatial error. The independent variable was haptic cue feedback. This variable had two levels, absence and presence of a haptic cue.

Experiment Two: Dependent Measures and Stimuli

Dependent measures for this experiment were performance time, error rate, bias, accuracy, and precision. Performance time measured the time it took to push the button. Error rate was the percentage of task errors, that is, missing the button. Bias, accuracy, and precision, as before, were defined as the average, average of absolute values, and standard deviation of the error distances respectively. The error distance was measured from the center of the finger tip to the center of the stimuli, the buttons.

Experiment Two: Participants and Procedure

Ten volunteers were recruited for the performance time and error rate portion. An additional ten were recruited for the accuracy tests. The participants were recruited from the pools of graduate students and researchers.

Each participant tapped a physical point on the table and then a virtual or a real button at the cue of a sound provided by the experimenter. Each participant repeated the task fifteen times with simple touch haptics feedback present and fifteen times without for the mirage button. Each repeated the task with the real button. Five of the participants received the haptic-present condition first, and the other five received the haptic-absent condition first. All received an additional condition using the real button last, just as a reference point for verifying the data collected. The total experiment time lasted about 15 minutes for each participant.

The button was presented as an optical mirage so that the set of visual cues presented were complete and ideal, i.e., occlusion, stereoscopy, accommodation, and convergence were perfect. This mirage also allowed for direct view of the hand and for correct synchronization of proprioception and visual cues. The mirage also provided for the exclusion of normal haptic cues. The simple touch cue was added for the haptics-present condition by placing a clear thin plastic panel where the mirage was.

All lights were turned off so no other visual cues were present. Black lights, white gloves, and florescent markers were used to highlight the hand and the stimuli. No other cue feedback was provided.

The second part of this experiment measured spatial error. Participant time was ignored so that spatial error could be more accurately measured. Ten volunteers participated, and none from the first group were reused. Four squares representing buttons were presented as optical mirages. The participants were advised that the squares would be presented visually but that haptic feedback may or may not be present. They were asked to align the index finger on top of each square. The position of the fingertip was measured using the IS600 tracker, which was specified with 1 mm accuracy at steady state and with overall test bed accuracy at 2 mm. Each participant performed this localization task four times for each of the four buttons presented and for each treatment.

(4 touches) x (4 squares) x (2 treatments) = 32 total touches per participant

Experiment Three: Depth Perception, Full Factorial

Experiment Three built upon the tools and results from Experiment One to provide a complete set of empirical data on depth perception spanning all three environment types and including the additional, salient sensory cue of proprioception. It was conducted on one common test bed, using one highly reliable participant pool, and leveraged proven methods in the literature.   The method of adjustment for depth perception was used and the stimuli were an octahedron and a cylinder of four degrees and one degree in widths, respectively (Rolland et al., 2002). Figure 8 depicts the virtual stimuli used for this experiment including calibration lines which facilitated alignment of the virtual and real environments.

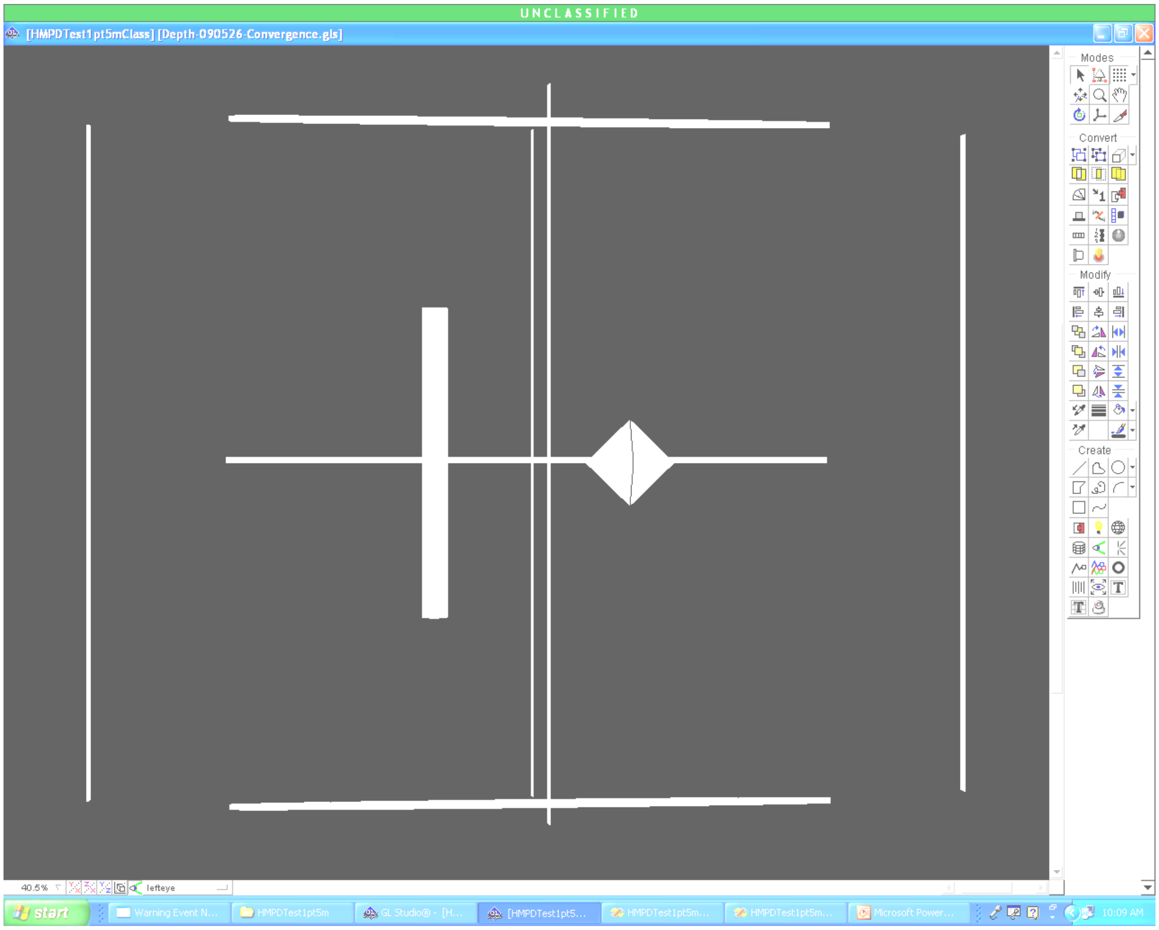


Figure 8. Experiment Three – Depth Perception, RE vs. MR vs. IVE.

Experiment Three: Hypothesis

Given the correct binocular visual and kinesthetic proprioception cues, the user can perceive depth with sub-centimeter accuracy between a real and a virtual object, between two real objects, and between two virtual objects. Additionally, MR may provide a significant increase in depth perception accuracy compared to that provided by IVE. This hypothesis was partly based on previous works of Ellis and Menges (1997, 1998) where real visual cues were found to affect perception of virtual cues. The addition of kinesthetic proprioception also provided an extra cue, which could potentially improve accuracy of depth perception. Furthermore, accuracy of depth perception in the RE was expected to be perfect, i.e., within measurement system accuracy of 2 mm. Finally, it was expected that accuracy in RE would be comparable to that of the MR condition where visual and kinesthetic proprioception cues were synchronized and visual cues for both environments are comparable in fidelity.

Hypothesis 3a: With accurate visual cues for virtual objects, mean depth perception errors are small enough for training applications, i.e., less than 10 mm.

$$H_a\text{: Error} < 10\text{mm (for MR and IVE)}$$

Hypothesis 3b: Proprioception significantly improves depth perception, i.e., the error is less given the added cue.

$$H_a\text{: Error}_{\text{proprioception}} < \text{Error}_{\text{no-proprioception}} \text{ (for MR and IVE)}$$

Experiment Three: Design

In addition to some of the design considerations discussed in the previous experiments, the points below provide for an improved, more robust experiment.

HMPD Brightness, Adjustments, and Calibration

While the HMPD provided for accurate adjustments and alignment for convergence, focus, and IPD, it had brightness and contrast issues in the previous experiments. The HMPD, like other displays, gradually lost its brightness and contrast over time with used. For Experiment Three, the displays were replaced with a new series of Organic Light Emitting Diodes (OLEDs) provided by eMagin Inc. that dramatically improved brightness.

With the unexpected brightness improvement, it was discovered that other, more reliable methods for adjustment and calibration can be used. Since the HMPD design was a hybrid of the HMD and projection display, theoretically, adjustment and calibration methods for either types could be used. Because of the low brightness level in Experiment One, only the user could see the image, and not the researcher standing by. Experiment One relied heavily on typical HMD methods. These methods largely require the user to wear the HMD to make the alignment and adjustments, for example verifying focus adjustment subjectively. The projection display methods, on the other hand, allow the user and the researcher to simultaneously see the same image and provide for real-time adjustment by the user and verification by the researcher. Experiment Three included frequent verification of alignment before, during, and after experiment sessions that were not previously feasible.

Full Factorial Design

A full factorial experimental design was employed (Table 5). In research terms, this design increased the volume (or information value) while reducing the noise (or contribution of unwanted variation). The design also provided for thorough analysis on the effect of each factor, or independent variable.

Table 5. Depth Perception, Full Factorial Design

		Visual Environment		
		Reality (RE)	*Mixed Reality (MR)*	*Immersive Virtual Environment (IVE)*
Proprio-ception	*Present*	Group 4	Group 5	Group 6
	Absent	Group 1	Group 2	Group 3

Across-Subject Design.

The use of an across-subjects experiment eliminated the confound with respect to fatigue effects, which was observed previously in Experiment One. It also eliminated carryover or practice effects from one treatment to the other, which may also have existed in Experiment One. An additional benefit was the VE exposure time. Since each participant underwent only one treatment, VE exposure and any side effects were minimized. One downside of the across-subject design was the requirement of many more participants, especially with a large number of treatments for the same statistical power as a within-subject design. Non-equivalency between groups was also a concern. For these reasons, the participant pool had to be carefully selected. An experimental venue that could provide for a large pool of highly reliable participants was crucial.

Participant Screening

To further minimize non-equivalency between groups, participants needed to have already been fully examined for visual capabilities or else, they had to be carefully tested and if necessary, screened out. Each participant had to pass the visual acuity test using the Snellen Eye Chart, to demonstrate stereoscopy abilities using the Stereo Fly Test, and to be able to accommodate for near and far focusing. Additionally, participants with IPDs too wide or too narrow for HMPD adjustments were also screened out.

Motion Parallax

Motion parallax, according to Cutting and Vishton (1995), is the third strongest cue for depth perception in personal space. The fidelity of this cue, however, could depend heavily on the accuracy and responsiveness of the head tracking system. For this reason, head tracking was not employed. The participant's head was fixed on a chin rest.

Experiment Three: Task and Independent Variables

As in Experiment One, the task was to adjust one object until the participant perceived that it was at the same depth as another nearby fixed object using the method of adjustment (Rolland et al., 2002b). Results from Experiment One provided the data needed to estimate the number of repetitions. Error measurements were made in Experiment One and the standard deviation, σ, was estimated to be approximately 6 mm. Half width was chosen at 2 mm, the approximate accuracy of the test bed. The minimum number of adjustments necessary for each participant was calculated below. This number of repeated adjustments provided for 95% confidence in estimating the population mean, i.e., the accuracy of depth perception, to within +/- 2mm.

$$N = \left[\frac{Z_{\alpha/2}\, \sigma}{H} \right]^2 = \left[\frac{1.96 * 6}{2} \right]^2 = 36 \text{ adjustments}$$

Experiment One also exhibited some fatigue or boredom effect. To balance the accuracy of the measurements with the potential fatigue factor, the number of adjustments was chosen to be about half the calculation above to twenty repetitions. This reduction in the repetition increased the half width from 2.0 mm to 2.6 mm or, equivalently, decreases the confidence in the mean estimation from 95% to 86%.

Independent variables, also called experimental factors, for this experiment were virtual environment type and kinesthetic proprioception cue. Virtual environment type had three levels. V/V, for virtual object next to virtual object, represented IVE. R/V, for real object next to virtual object, represented MR. R/R, for real object next to real object, represented RE. The proprioception factor had two levels, presence or absence of the cue. For the presence condition, the participant's hand moved inch for inch with the octahedron. For the absence condition, the participants pressed arrow buttons on the keyboard to move the octahedron and did not get any relevant, direct proprioception feedback.

Experiment Three: Dependent Measures and Stimuli

As in Experiment One, error distances from participant adjustments were collected and bias, accuracy, and precision were calculated and used as dependent measures. These measures were computed for each participant using the average, average of absolute value, and standard deviation of the error distances, respectively. Also similar to Experiment One, the stimuli were an octahedron and a cylinder. Both were virtual objects for the IVE condition and both were real objects for the RE condition. The cylinder was real and the octahedron was virtual for the MR condition. Again, the octahedron was alternated randomly left, right, in front of, and behind the cylinder, which was fixed at 0.65 m away, approximately the extreme reach of the participants.

Experiment Three: Participants

The results of Experiment One provided the data needed to estimate the sample size for this experiment. The sample size estimate technique depended on the statistical analysis used, the statistical power and significance criteria chosen, and an estimation of the effect size.

For the purposes of sample size estimation, ANOVA, F-Test, for comparing multiple means was assumed to be the method that would be used for statistical analysis. This in turn, required other assumptions. For each dependent measure, the variances for the treatments were assumed to be equal. Each measure was assumed to be normally distributed. Finally, each treatment was assumed to be the same size.

In addition to the above assumptions, statistical power and alpha value had to be chosen. Power, P, represents the probability of correctly rejecting the null hypothesis, H_o, and was chosen to be 0.8, as generally done in this field of study (Cohn, 1988). Alpha, α, represents the probability of falsely rejecting the null hypothesis, H_o, and was chosen to be 0.05 also as generally done in the field.

Effect Size, d, was defined as σ_m / σ, for the F-Test. The variable, σ_m, is the standard deviation for all experimental population means. The variable, σ, is the standard deviation within one of the populations and was assumed to be the same for all for the purpose of sample size estimation. Effect size had to be estimated. Based on results from Experiment One, σ can be reasonably estimated to be 6 mm. σ_m was more difficult to estimate. Conservative estimates for the mean accuracy in the four (k=4) experimental treatments were 14, 8, 2, and 2 mm. This yielded $\sigma_m = 4.8$ and $d = 4.8/6.0 = 0.8$.

Given the above criteria and assumptions, the sample size needed was (Cohen, 1988, Table 8.4.4, p. 384 and Table 8.3.14 p. 315 for d=.8, $\alpha = 0.05$, Power = 0.8, and u = k-1 = 3):

N = 6 participants per population or treatment

Experiment Three design included 6 treatments. To balance out and slightly match the number of participants required by the next experiment, Experiment Four, N was chosen to be one higher, N = 7. Total sample size for this experiment was 7 times 6 treatments, which equated to 42 participants.

Experiment Three: Procedures

The participant adjusted the position of one object, an octahedron until its depth was believed to match another close by, stationary object, a cylinder. The experiment quantified error of depth perception for all three types of Virtual Environments – RE, MR, and IVE. Half of the participants performed the task with kinesthetic proprioceptive cue and the other half performed the task without the feedback. With the kinesthetic proprioceptive cue, the participant moved the octahedron by translating a computer mouse which was calibrated, inch for inch, with the movement of the octahedron. Without the proprioceptive cue, the participant indirectly moved the object using computer keyboard buttons. The up and down buttons moved the octahedron slowly while the left and right buttons moved it quickly. For this experiment, the HMPD and head were fixed. The chin was rested on a fixed reference point.

Experiment Four: Object Interaction, Full Factorial

Experiment Four was also a full-factorial design that built on the methods and results obtained from Experiment Two. Figure 9 illustrates the experimental test bed set up in a night vision goggle training laboratory at the Naval Air Station in Jacksonville, Florida. It depicts the HMPD, retro-reflective fabric, the chin rest, the black light, and the real stimuli used.

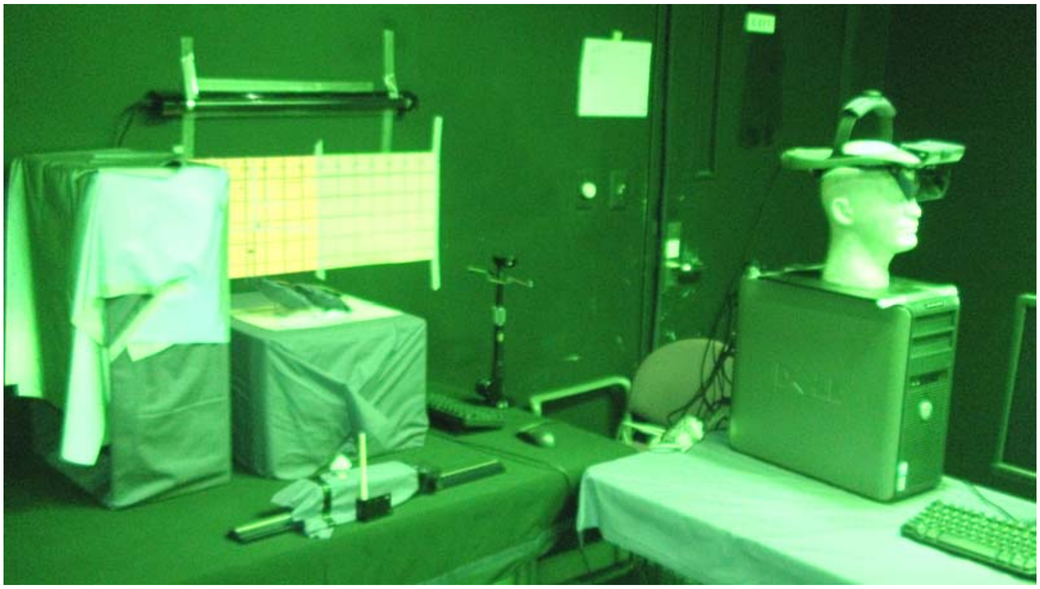


Figure 9. Experiment Four – Object Interaction, RE vs. MR.

Experiment Four: Hypothesis

The following hypotheses were drawn for the research questions discussed earlier.

Hypothesis 4a: Given accurate visual cues, including stereoscopy, accommodation, convergence, and occlusion and appropriate and synchronized proprioception cue, interaction with a simple object is highly accurate, in the sub-centimeter (millimeters) range.

$$H_a : \text{Error} < 10 \text{ mm}$$

Hypothesis 4b: Performance time and error for RE is significantly less than that for MR. RE provides perfect visual cues compared with MR with near-perfect cues. A

MHP calculation shows a slight decrease in performance time in the MR condition (see Appendix D).

H_a : $\text{Error}_{RE} < \text{Error}_{MR}$; H_a : $\text{Performance Time}_{RE} < \text{Performance Time}_{MR}$

Hypothesis 4c: The addition of an audio cue significantly improves task performance time. This prediction is mainly asserted because audio is a temporal cue, which could contribute to shorter completion time (Shilling & Shinn-Cunningham, 2002).

$$H_a : \text{Performance Time}_{audio} < \text{Performance Time}_{no\text{-}audio}$$

Hypothesis 4d: The addition of a haptic cue significantly improves accuracy of task performance. This prediction is made because haptics is a spatial cue, which could contribute to spatial accuracy (Schiefele, 2000; Lok, 2000).

$$H_a : \text{Error}_{haptics} < \text{Error}_{no\text{-}haptic}$$

Experiment Four: Design

An across-subjects design was developed for this experiment for similar reasons outlined previously. The design was a full factorial experiment with three factors of two levels each for a total of eight treatments (Table 6).

Table 6. Experiment Four Design, Full Factorial

Audio Absent					*Audio Present*			
		Visual Environment					*Visual Environment*	
		Reality (RE)	*Mixed Reality (MR)*				*Reality (RE)*	*Mixed Reality (MR)*
Haptics	*Present*	Group C	Group D		*Haptics*	*Present*	Group G	Group H
	Absent	Group A	Group B			*Absent*	Group E	Group F
5 participants x 8 groups = 40								

Experiment Four: Tasks and Dependent Measures

The task was simple, direct interaction with real and virtual objects, which were keyboard buttons. Independent variables were virtual environment type, haptic cue, and audio cue. Virtual environment type included MR and RE. The MR condition was represented with the real hand manipulating the virtual object, or the R/V condition. The RE environment type was represented with the real hand manipulating the real button, or the R/R condition. The haptic cue had two levels, presence and absence. Likewise, the audio cue had two levels, presence and absence.

For estimation of the number of repeated actions for this manipulation experiment, the same approach as used in Experiment Three was employed. A number of assumptions were necessary. First choosing the half-width, H, to be 100 milliseconds sufficed since a difference of one-tenth of a second is practically insignificant for the task of manipulating cockpit controls. It was also reasonable to assume that the variation in performance time between the same, simple, repeated task of manipulating buttons for one participant was about a quarter of a second based on preliminary measurements conducted on the test bed. The minimum number of repeated actions necessary for each participant was as calculated below. This number provided for 95% confidence in estimating the mean, i.e., the performance time of an individual, to within plus or minus a tenth of a second.

$$N = \left[\frac{Z_{\alpha/2}\ \sigma}{H} \right]^2 = \left[\frac{1.96 * 250}{100} \right]^2 = 25 \text{ actions}$$

Experiment Four: Dependent Measures, Cues and Stimuli

Dependent measures for this experiment were performance time, bias, accuracy, and precision. These terms were defined previously for Experiment Two and reused for this experiment.

Cues for this experiment included visual feedback synchronized with proprioception, haptic feedback, and audio feedback. Visual feedback included, among other visual cues, true and synchronized occlusion (interposition of one object in front of another). Incorrect occlusion cues could contribute to depth perception error (Ellis & Menges, 1997). Timely feedback of action taken could reinforce and compliment other visual cues and therefore improve task performance. Diotic audio feedback (simple mono ding sound) was provided. Audio, a temporal cue, enhances awareness of the virtual environment (Shilling & Shinn-Cunningham, 2002) and could improve task performance, specifically, reaction time. Finally, passive touch feedback was also provided. This feedback could significantly enhance performance of virtual object interaction (Schiefele, 2000).

Real and virtual buttons similar to those found in flight cockpits were used as stimuli. Real buttons were presented using the mirage dish similar to that described in Experiment Two. Virtual buttons were presented by the computer software application, GLStudio and the HMPD. Haptic feedback was provided by placing a clear, essentially invisible, plastic panel where the mirage and the virtual buttons were. Audio feedback was provided by an infrared sensor and a speaker set to provide a ding sound when the infrared beam was broken.

Experiment Four: Participants

Results from Experiment Two provided critical data needed to estimate population size for this experiment. Similar methods and assumptions previously discussed were used to estimate the number of participants required.

ANOVA, F-Test, was assumed for the purposes of sample size estimation. This Power Analysis also assumed equal variance, normality, and equal sample sizes.

Effect Size, d, is defined as σ_m / σ for the F-Test. The standard deviation for performance time within a population, σ, could reasonably be estimated to be half a second. Standard deviation across treatments, σ_m, could be determined from mean performance times, which can conservatively be estimated to be at 2.2, 1.8, 1.4, and 1.4 seconds. This yielded $\sigma_m = 0.33$ and $d = 0.33 / 0.50 = 0.7$.

Given the above criteria and assumptions, the sample size needed, for k = 8 treatments, was (Cohen, 1988, Table 8.4.4, p. 384 and Table 8.3.18 p. 323 for $d = 0.7$, $\alpha = 0.05$, Power = 0.8, and $u = k-1 = 7$):

N = 4 participants per population or treatment.

The number of participants was chosen to be one more than minimally required, which equated to five participants per treatment. With eight treatments, the total number of participants for Experiment Four was 40. This sample size was similar to the total number of participants in Experiment Three, which was a practical convenience. Participants from Experiment Three were reused as much as practical and randomly assigned to the treatments for this experiment.

Experiment Four: Procedures

For this experiment, the participant's head was also fixed. The hand and finger were not fixed and its position was measured relative to the fixed head and the fixed stimuli.

Similar to the traditional Fitt's tapping task conducted by Arsenault and Ware (2000), each participant tapped two buttons, alternating from one to the other, 32 times. Half of the participants tapped buttons that were real. The other half tapped buttons that were virtual. Half received haptic cue feedback and half did not. Likewise, half received audio feedback and half did not. In addition to the visual cues of accommodation, convergence, and proprioception provided in Experiment Three, the participant also received an additional visual cue, occlusion (i.e., the participant's hand blocked the view of the object).

Overall Experimental Procedure and Design

The sample sizes for Experiment Three and Experiment Four were designed to be similar. These experiments shared as many of the participants as possible. Each treatment group from Experiment Three was randomly and evenly distributed among treatment groups in Experiment Four. Each participant for Experiments Three and Four underwent, roughly, the procedures as detailed in Table 7 lasting about 84 minutes. Table 8 summarizes the four experiments including the respective hypotheses, experimental tasks, control factors, and dependent measures. Lastly, Table 9 describes some of the uncontrollable independent variables (IV) previously discussed, techniques to measure these, and approaches in Experiments Three and Four to mitigate their effects.

Table 7. Steps and Approximate Duration for the Experiments

Step #	Procedure	Approximate Duration (minutes)
	Pre-Experiment Procedures	
1	Informed Consent, Explanation, Questions & Answers	10
2	IPD Measurement	1
3	Break (return later at assigned date/time)	2 (or days)
2	Background Questionnaire	2
3	Snellen Eye Chart Test	2
4	Dolman Test	5
5	Pre-Test SSQ	3
6	Break, if needed	5
	Depth Perception (Experiment One or Three)	
7	HMPD/Participant Alignment and Calibration	5
8	Practice Session	3
9	Depth Perception Experiment	10
11	Break, if needed	10
	Virtual Object Interaction (Experiment Two or Four)	
12	HMPD/Participant Alignment Verification	2
13	Practice Session	3
14	Object Interaction Experiment	10
15	Post-Test SSQ	3
16	Break	15
	Post-Experiment Procedures	
16	Snellen Eye Chart	2
17	Vergence Test	2
18	Eye-Hand Coordination Test	2
	Total Estimated Time	84

Table 8. Summary of Experimental Designs

Experiment	**Task**	**Factors**	**Measures**	**Hypothesis (H_a)**
1) Depth Perception, IVE	Adjust Objects for Equal Depth	Visual Cue – OSTHMD vs. HMPD	Bias, Accuracy, Precision	$\text{Error}_{\text{HMPD}} < \text{Error}_{\text{OSTHMD}}$
2) Object Interaction, RE	Tap mirage buttons	Haptic Cue – Presence vs. Absence	Bias, Accuracy, Precision, Performance Time	$\text{Error}_{\text{haptic}} < \text{Error}_{\text{no-haptic}}$ $\text{Time}_{\text{haptic}} < \text{Time}_{\text{no-haptic}}$
3) Depth Perception, IVE, RE, MR	Adjust Objects for Equal Depth	VE Type by Proprio-ception	Bias, Accuracy, Precision	$\text{Error} < 10\text{ mm}$ $\text{Error}_{\text{proprio}} < \text{Error}_{\text{no-propprio}}$
4) Object Interaction, RE, MR	Fitt's Tapping Task	VE Type by Haptic cue by Audio Cue	Bias, Accuracy, Precision, Performance Time	$\text{Error} < 10\text{ mm}$ $\text{Error}_{\text{RE}} < \text{Error}_{\text{MR}}$ $\text{Error}_{\text{haptic}} < \text{Error}_{\text{no-haptic}}$ $\text{Time}_{\text{audio}} < \text{Time}_{\text{no-audio}}$

Table 9. Uncontrolled Independent Variables (IV) and Mitigation Approach

Uncontrolled IV	**Measurement Technique**	**Experimental Approach**
Acuity	Eye Chart	Eliminate
Fusion Ability	Stereo Fly Test	Eliminate
Standard Depth Perception	Standard Dolman Box Test	Factor in Analysis
Nearfield Depth Perception	Dolman Box Test at 0.8 m	No-treatment Baseline, Distributed b/w treatments
Sim Sickness	Pre/Post SSQ	Data Analysis
IPD	IPD Meter	Eliminate
Dexterity	Observation	No time limitation
Practice/Boredom	Data observation	Practice, limited repetition

Statistical Analysis Plan

Confidence Interval, ANOVA, K-W, and Multiple Comparisons

95 % confidence intervals were used to gain insight on the average error or performance time for each experiment. Multiple comparison techniques were appropriate for determining which treatments differed in mean (Mendenhall & Sincich, 1994). However, it was appropriate only after rejecting the hypothesis of equality between treatments. This test for equality can be done for more than two (k>2) treatments by using the ANOVA F-Test technique or the Kruskal-Wallis (K-W) Test. The K-W test

would be used only if non-normality or unequal variances existed. The outline below provides the steps for analysis for comparison of the mean of the measures (bias, accuracy, precision, and performance time) between the treatments within each experiment.

1. Test for Normality, Shapiro-Wilk test.
2. Test for equal population variances, Bartlett's test for homogeneity of variances.
3. If all treatment distributions are normal and all five variances are equal,
    a) Use ANOVA F-Test (One-Way, Two-way, and General Linear Model)
    b) If the ANOVA results reject the null hypothesis, i.e., at least one treatment mean was different among all, then use Tukey's multiple comparisons procedure to determine which means are different.
4. If any of the treatment distributions are not normal or if all treatment variances are not equal,
    a) Use non-parametric Kruskal-Wallis Test for Shift in Population Locations for Independent Random Samples.
    b) If the K-W results rejects the null hypothesis, i.e., indicates that at least one treatment mean was different, use a non-parametric multiple comparison procedure to determine which treatments differ.

CHAPTER FOUR: RESULTS

This dissertation was structured such that latter experiments, which were more complex and thorough, can gain from the methods, results, and lessons learned from the prior experiments, which were more conservative and narrowly focused. Population estimation for the latter experiments leveraged the empirical data collected from the prior experiments. A fatigue or boredom effect was visible in the first two experiments and drove the design of the last two. Fragility of the experimental apparatus was discovered in the first experiment that resulted in some data loss. The ODA Lab HMPD prototype provided for Experiment One had been developed as a proof of concept system and was not originally designed to be used in human factor studies. Substantial time and cost was invested in fixing, integrating, testing, aligning, and calibrating the HMPD for significantly more reliable, latter experimentations. A software routine was developed in the process to help maintain calibration on the HMPD. The most important factor gained from the first experiments was the level of interest of participants. When the participants were serious and fully engaged in the experiment, the data collected was significantly more robust. This drove a change of venue. The first experiments were conducted in a research laboratory. For the last experiments, the test bed was transported to a Navy base where the participant population was primarily military. The results from this set of four experiments built from one to the next.

For each participant and each treatment, the first three measures listed below were calculated for statistical analysis. For Experiments Two and Four, the last two measures listed below were also added for analysis.

1. Bias: average of the errors for the *r* repetitions
2. Accuracy: average of the absolute values of the errors for the *r* repetitions
3. Precision: standard deviation of the errors for the *r* repetitions
4. Performance time: mean of the performance times for the *r* repetitions
5. Error Rate: number of times error was made divided by *r* repetitions

Experiment One: Perception, OSTHMD vs. HMPD

Experiment One measured depth perception in IVE, that is, with two virtual objects side by side as stimuli. Two different visual display types were used, the OSTHMD with nominal visual cue settings and the HMPD with highly accurate settings. Comparisons were made to determine the effect on depth perception when the visual display was optimally designed and carefully adjusted. Data was collected and analyzed from four participants in a within-subject study. The participants were graduate students and researchers.

The experimental design and written procedures were followed closely. Some additional implementation details and observations during data collection are worth noting up front.

Since the HMPD can be worn with prescription glasses, focus adjustment was unnecessary and was accurately preset. The HMPD IPD was also visually and accurately aligned with the participants' eyes. For the OSTHMD, these settings were set nominally and the participants subjectively made adjustments themselves for the clearest and most comfortable viewing. Each experimental session was conducted in a dark, quiet room so that no other unwanted cues were present.

Both the left and right HMPD OLED displays suffered from poor brightness. Even with maximum settings, the HMPD was still significantly dimmer than the OSTHMD. Furthermore, one HMPD display was dimmer than the other. Additionally, the IPD adjustment for the HMPD was limited to 68 mm and precluded participants with wider IPD from participating. The HMPD also flickered on and off, either due to inadequacy of the USB (Universal Serial Bus) power or display interface cable. The OSTHMD also had limitations in head size adjustment and focus adjustment. The focus adjustment was subjective and it was not clear if participants were able to fully adjust the settings. Software adjustment was available to offset different individual IPD, but setting the IPD on the OSTHMD was another subjective alignment that was difficult to perform. Although Experiment One was relatively successful, it was wrought with alignment and troubleshooting issues.

Each participant was allowed as much time as needed to practice the task until comfortable. However, not all were able to completely and clearly see the computer generated images, either due to the inability to fuse the images or else inability to focus on the images. Towards the end, some participants also showed signs of boredom or fatigue, making the data collected less reliable.

As a result of the screening procedure, technical issues with the fragile displays, and the unreliability of some data collected, not all participants were included in the final analysis for Experiment One. The final sample size for statistical analysis was n = 4 for this within-subject, depth perception experiment.

Data Analysis

For this depth perception experiment, unreliable data from equipment misuse necessitated exclusion of a number of participants in the final data analysis. Where there were questionable data, the entire set of data for that participant was discarded. The final analysis revealed an average error of about 3mm (SD = 5mm) for bias for both the OSTHMD and the HMPD. Precision averaged about 8mm (SD = 3 mm) for the HMPD and about 11mm (SD = 3 mm) for the OSTHMD. The results of this depth perception experiment showed that correct stereoscopy, accommodation, and convergence settings provide approximately sub-centimeter bias and precision, which were consistent with the results of Ellis and Menges (1997) and Rolland et al. (2002b).

It is noted that the findings in this depth perception study did not produce statistically significant differences. The study was unable to find significant differences in error in depth perception for the HMPD vs. that for the OSTHMD. Table 10 displays the bias and precision measures, in millimeters, for each of the three sizes of stimuli, for each visual display, and for each participant. Averages of the measures are in the final column.

Table 10. Bias and Precision of Depth Perception, HMPD vs. OSTHMD

		Subj A			Subj B			Subj C			Subj D			
Display	Measures	Size1	Size2	Size3	Size1	Size2	Size3	Size1	Size2	Size3	Size1	Size2	Size3	Average
HMPD	Bias	5.9	7.6	2.1	-10.8	1.8	-4.0	10.9	1.4	5.5	5.4	4.6	2.0	2.7
HMPD	Precision	5.4	5.1	5.5	13.1	9.9	12.6	6.6	8.3	6.8	7.3	6.2	7.1	7.8
OSTHMD	Bias	6.6	1.8	-4.0	-0.7	0.5	-2.3	4.4	17.9	2.5	2.2	2.4	2.9	2.9
OSTHMD	Precision	6.3	9.9	12.6	14.8	7.2	8.2	14.4	17.9	14.7	7.5	9.1	9.1	11.0

Lastly, there was no learning effect observed from the time series plot of the results for either experiment. This was expected since the participants did not get any feedback on how well they performed after each repetition. In fact, the time series plot showed a slight decrease in performance, which may indicate an effect due to boredom or fatigue.

Observations

This experiment served as a first test of the apparatus, the test setup, and participants' response. The results indicated little or no signs of eyestrain using binocular displays for six ten-minute sessions for each participant. A number of equipment use problems were discovered that accounted for a large variation in the measured results. Factoring out equipment misuse and eliminating participants with less than perfect vision, initial results of this depth perception experiment does approximate that conducted by Ellis and Menges (1997) and Rolland et al. (2002).

This experiment searched for evidence supporting improved depth perception using the HMPD over that using the standard OSTHMD. Statistically significant differences in depth perception between the HMPD and OSTHMD were not observed in this experiment. Further experiments, perhaps with a larger number of participants and a better controlled virtual environment, as planned later for Experiment Three, may provide statistically significant findings.

These VE depth perception results, however, did include some data points where sub-centimeter accuracy was achieved with careful alignment and proper use for both the HMPD and the OSTHMD. Although not statistically significant, the results also indicated a slight improvement in precision with the HMPD (μ = 8 mm, SD = 3 mm) over the OSTHMD (μ = 11 mm, SD = 3 mm).

Experiment Two: Interaction, Haptics vs. None

This experiment included two phases. The first phase measured performance time and error rate. During the first phase, each of five participants tapped an optical mirage of a keyboard button. The participants repeated the task with simple touch feedback provided by a clear (almost invisible) plastic sheet. Another set of five participants performed the same tasks, but with the haptic cue provided first. For each treatment, the participants contributed a number of repetitions, r =15. Figure 10 illustrates the simple tapping task for the haptic present condition provided by the clear plastic panel on top of the mirage dish.

Figure 10. Experiment Two, Performance Time Measurement.

The second phase was independent of the first and measured spatial accuracy. Ten other participants, none from the first phase, aligned their fingers on four optical mirages of square shapes. This alignment was quantified to provide a measure of the accuracy of interaction with the object. Again, for a balanced procedure, five participants performed the task first with the touch cue absent and repeated the task with the touch cue present. The other five started with the touch cue present first and repeated the task with the touch cue absent. For each treatment, the participants contributed a number of repetitions, r = 16.

An Intersense IS600 system was used to track the hand. It was a hybrid acoustics and accelerometer based device with highest accuracy in the steady state mode.

Therefore, for optimal spatial accuracy, the participants were asked to align and hold their fingers still for a few seconds. This was also the reason for having two phases of this experiment. The first phase included a dynamic task (tapping) for measuring performance time and the second included a static task (holding finger steady) for measuring spatial error. Two different sets (ten participants each) were used to eliminate any carry over effect going from one phase to the other.

Real keyboard buttons were used to generate the corresponding optical mirages. These buttons were the type of controls typically found in a flight cockpit. A Snellen Eye Chart and a Dolman Depth Perception Box measured static acuity and depth perception, respectively, and were used to screen out participants who had less than normal vision.

A black light, florescent paint, and white gloves were used to highlight the stimuli. The area was kept completely dark and quiet so that no other visual or audio cues were present.

Data Analysis

For the first part of the interaction experiment, no errors were observed, i.e., each of the participants touched the mirage button on every trial. The results also revealed no significant differences in performance time. Performance time averaged about one second, whether the real button or the mirage was used, and whether or not haptics feedback was presented.

The second part of the interaction experiment did reveal statistically significant differences ($\alpha = 0.05$) in all of the measures—accuracy, bias, and precision—but only in one spatial dimension, *Z*, which was the axis perpendicular to the square presented and was upright vertical in orientation. Errors in the *X* axis, which projected outward in front of the participant, and that in the *Y* axis, which projected away to the left of the participant, showed no significant differences between the haptics-present and the haptics-absent conditions. Errors in the total distance, *D*, also showed no significant difference. Precision for the total distance averaged about 0.1 inches, or 2.5 mm (see Table 11) whether haptic cue was present or not.

Table 11. Mean Spatial Error for Interaction Experiment (in inches)

	Bias (*b*)				Accuracy (*a*)				Precision (*p*)			
	X	Y	Z	D	X	Y	Z	D	X	Y	Z	D
No Haptics	0.21	0.14	0.17	0.44	0.26	0.20	0.19	0.43	0.12	0.13	0.09	0.09
Haptics	0.17	0.16	0.06	0.36	0.23	0.20	0.08	0.36	0.14	0.12	0.04	0.08
Difference	0.04	-0.01	0.12	0.08	0.03	0.01	0.11	0.07	-0.02	0.01	0.06	0.00

The Wilcoxon signed ranks test (Mendenhall & Sincich, 1994) for matched pairs using the one-tailed test for shift in distribution was employed for statistical analysis. The bold numbers in Table 12 contribute to *T*-, which is the rank sum of the negative differences and is the test statistic. The last row provides for T_o, the criteria for the test statistic. Of the twelve columns, only the three *Z* columns had *T*- less than T_o, for *alpha* = 0.05. Therefore, only the differences in the Z direction are statistically significant.

Table 12. Wilcoxon Signed Ranked Test for Matched Pairs

		Ranking											
		Bias (*b*)				Accuracy (*a*)				Precision (*p*)			
		X	Y	Z	D	X	Y	Z	D	X	Y	Z	D
	Participants	A	B	C	D	E	F	G	H	I	J	I	L
	A	**8.0**	3.5	6.5	2.0	**4.0**	3.0	6.5	2.0	0.0	2.0	5.0	**1.5**
	B	**7.0**	9.5	8.0	8.0	**6.0**	9.0	8.0	8.0	**8.0**	5.0	5.0	**5.0**
	C	5.0	5.0	6.5	7.0	8.5	7.0	6.5	7.0	5.0	0.0	7.5	1.5
	D	9.0	**6.5**	10.0	9.0	10.0	**5.5**	10.0	9.0	**2.0**	**4.0**	5.0	0.0
	E	**6.0**	1.5	2.5	**4.0**	**8.5**	4.0	1.5	**3.5**	**3.5**	0.0	0.0	0.0
	F	10.0	**9.5**	9.0	10.0	6.0	**8.0**	9.0	10.0	1.0	6.0	9.0	7.0
	G	**1.5**	**8.0**	1.0	**6.0**	**1.0**	**5.5**	1.5	**6.0**	**7.0**	2.0	3.0	**4.0**
	H	**4.0**	**1.5**	**2.5**	**5.0**	**3.0**	**1.5**	4.0	**5.0**	0.0	**8.0**	7.5	0.0
	I	**3.0**	**6.5**	4.5	**1.0**	**6.0**	0.0	5.0	**1.0**	**6.0**	2.0	**1.0**	**3.0**
	J	1.5	3.5	4.5	3.0	2.0	1.5	3.0	**3.5**	3.5	7.0	2.0	6.0
Average	Wilcoxen T+	25.5	31.0	52.5	39.0	26.5	24.5	55.0	36.0	9.5	24.0	44.0	14.5
	Wilcoxen T-	29.5	24.0	2.5	16.0	28.5	20.5	0.0	19.0	26.5	12.0	1.0	13.5
	To @ alpha = 0.05	11.0	11.0	11.0	11.0	8.0	11.0	8.0	11.0	6.0	6.0	8.0	4.0

Observations

This experiment searched for evidence supporting improved task performance attributed to the addition of haptic feedback. The results are somewhat inconclusive for these specific objectives, but do provide some clear insights into the overall research question of whether VEs can support virtual object interaction and if visual cues alone are enough.

First, overall, statistical significance was not observed in performance time, error rate, bias, accuracy, or precision of object interaction between the haptic-present and haptic-absent conditions. However, the error rate was zero, the spatial error was small, in the three-millimeter range, and performance time was near optimal, about one second, even in the haptic-absent condition. Because these errors and completion time are relatively small and the visual environment, RE, was ideal, there was little room for further improvement in task performance with the addition of haptic cue. The added touch cue did not significantly affect task performance overall for RE. Further experiments, as planned for Experiment Four, with a larger number of participants and including a less optimal visual environment, for example, MR, may provide statistically significant findings.

Although overall results were not significant, Experiment Two did show statistically significant improvement in precision with a haptic cue added for one inconsequential spatial axis. This was the Z or vertical axis. The result simply validated the test equipment accuracy and the method and does not necessarily suggest a practical improvement in task performance with haptic cue added. A difference in the depth axis would have been more significant.

Experiment Three: Perception, Full Factorial

By design, Experiment Three gained from previous experiments. Sample size estimation took advantage of the data gathered from Experiment One. Graduate student participants used in Experiments One and Two were not ideal and thus military personnel were recruited as participants instead. Experiment Three was also adjusted to avoid boredom effects, another lesson learned from Experiments One and Two. Finally, the HMPD display was fixed and more care was given to calibration and measurement procedures as a result of large variation observed previously.

Experiment Three measured depth perception and compared the results across VE types and proprioception conditions in a full factorial design. VE types included R/R, R/V, and V/V which represented RE, MR, and IVE, respectively. The proprioception conditions included presence and absence of the cue. Figure 11 shows the measurement in the RE condition with proprioception cue present.

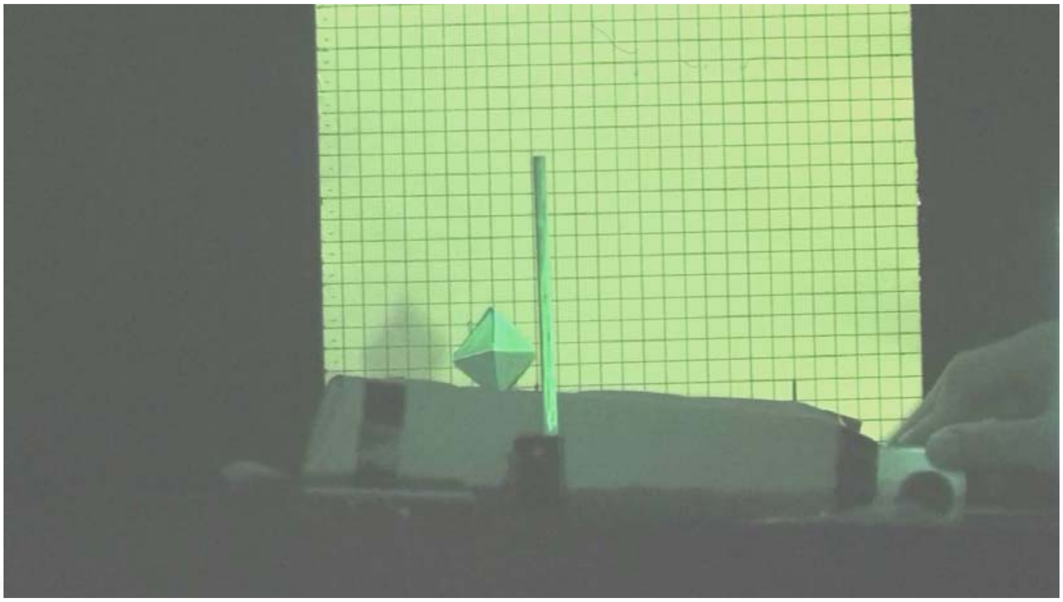


Figure 11. Experiment Three, Depth Perception, RE with Proprioception.

Data Collection

Repetition, Military Participants, and IPD Grouping

There was no learning effect observed in Experiments One and Two since the participants did not get any feedback on how well they performed after each adjustment. However, fatigue or boredom effect may have been present. Therefore, Experiment Three only used 20 repetitions, instead of 96 as done in Experiment One. As discussed in Chapter Three, 20 repetitions still allowed for adequate measurement accuracy, 2.6 mm, at 95% confidence interval.

For Experiment Three, military personnel were recruited. A military installation with a pool of well disciplined students and instructors is an ideal experimental environment. Since these participants were from the military, they had already been screened and had the eyesight required for the study. Almost all were familiar with the experiment pretests including the Snellen eye chart, the Dolman box, and the Stereo Fly test. Many expressed that they considered these procedures to be as important as similar eye exams. They followed the procedures intently and asked relevant questions to make sure they had every opportunity to perform their best.

One by one, the participants' IPDs were measured on the first day. Participants were grouped by IPD size. Their experiment date and time were scheduled according to IPD. This minimized the number of times the HMPD and the graphics software had to be readjusted from one participant to the next.

Optical Alignment

The HMPD used in the study was repaired shortly before Experiment Three data collection started. The microdisplays were replaced, which improved brightness dramatically. Since the HMPD projected an optical real image, it could be seen by both the participant wearing the HMPD and the researcher standing directly behind, but only when the HMPD was bright enough. The room for the experiment was completely dark, black on all walls, floor, and ceiling, and normally used for Night Vision Goggle (NVG) training. This setting allowed the researchers to see the display images, dimly, with the participant wearing the HMPD. Experiment Three capitalized on this capability to perform quick verifications of alignment before, during, and after experimental sessions. Software routines were coded to facilitate rapid, fine compensation for participant variations and slight HMPD maladjustments. Keyboard functions were programmed to display calibration lines, cross-hairs, and boxes with fine translational and angular adjustment of graphics images for each eye as needed. The participant was still relied upon to provide feedback on whether the image looked clear and aligned, but the researcher could confirm separately with the dim images seen.

Measures (Responses, y), Sample Size (N)

Experiment Three had six treatments. Each treatment had seven participants randomly assigned. Each participant, for each experiment, contributed 20 data points from which the following was calculated.

1) Depth Perception Bias: average of error for the 20 adjustments

2) Depth Perception Accuracy: average of absolute value of the error for the 20 adjustments

3) Depth Perception Precision: standard deviation of the error for the 20 adjustments

These bias, accuracy, and precision values were the dependent measures used for statistical analysis. Each of the six treatments included seven participants for a total sample size, N = 42.

Data Analysis

Microsoft Excel was used to calculate the measures from the raw experimental data. These tables of measures were imported into Minitab application software, which produced statistical quantities, tables, and plots for analysis and discussions.

Depth Perception Bias

Depth perception bias, for each participant, was a measure of how far, on average, the participant overshot or undershot when attempting to align the moveable octahedron to line up with the fixed reference depth, the cylinder. Bias was calculated by averaging the error distances, which were negative or positive, from the 20 repeated adjustments. Positive bias value would indicate that the octahedron or virtual object is, on average, adjusted closer to the participant. Since the dependent measure was continuous data, not ordinal data, ANOVA was used for statistical analysis with frequent checks for evidence of non-normality or unequal variances. Figure 12 depicts the histogram, normal probability plot and effects plots for the dependent measure of bias. The histogram and normal probability plot did not indicate signs of non-normality. The main effects plot indicated some differences in mean perception bias between the IVE condition and the RE and MR conditions.

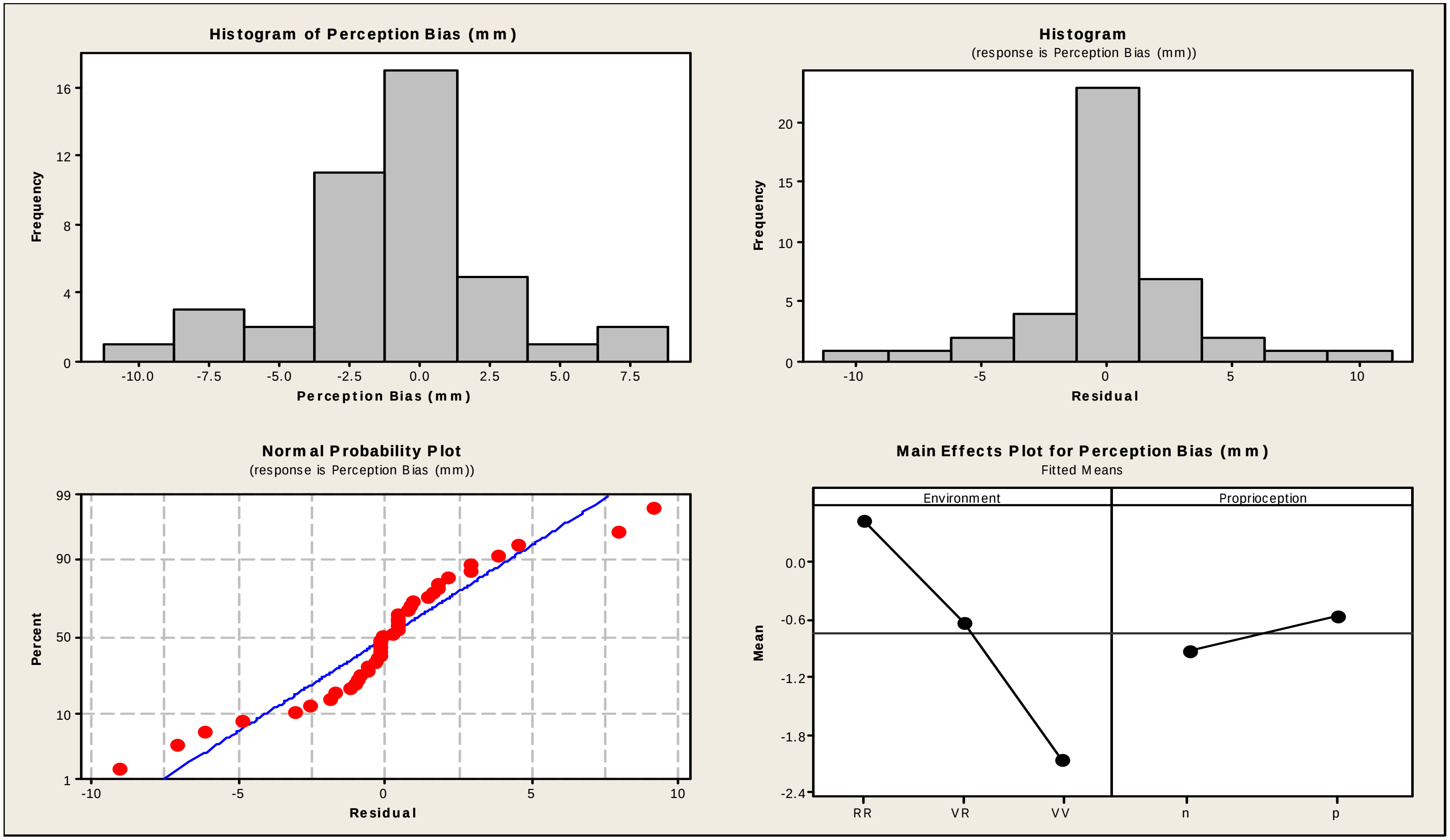


Figure 12. Depth Perception Bias Plots.

The ANOVA F-Test including all main effects and first order interaction was computed and shown in Table 13 and Table 14. The test did not show statistical significance for either of the two terms, environment type or proprioception, nor their interaction (environment type by proprioception).

Table 13. ANOVA for Depth Perception Bias (Full Factorial Fit)

```
Source                      DF  Seq SS  Adj SS  Adj MS     F      P
Environment                  2   43.55   43.55   21.77  1.80  0.180
Proprioception               1    1.49    1.49    1.49  0.12  0.728
Environment*Proprioception   2    0.82    0.82    0.41  0.03  0.967
Error                       36  435.53  435.53   12.10
Total                       41  481.38
```

Table 14. ANOVA and Confidence Interval for Depth Perception Bias

```
Source          DF       SS       MS     F      P
Environment      2   43.549  21.7745  1.90  0.164
Proprioception   1    1.486   1.4860  0.13  0.721
Error           38  436.349  11.4829
Total           41  481.384


                       Individual 95% CIs For Mean Based on
                       Pooled StdDev
Environment      Mean  ----+---------+---------+---------+-----
RR            0.42857                 (-----------*----------)
VR           -0.63571           (----------*----------)
VV           -2.05714  (----------*-----------)
                       ----+---------+---------+---------+-----
                        -3.2      -1.6      -0.0       1.6


                           Individual 95% CIs For Mean Based on
                           Pooled StdDev
Proprioception       Mean  ----+---------+---------+---------+-----
n               -0.942857  (--------------*--------------)
p               -0.566667     (--------------*--------------)
                           ----+---------+---------+---------+-----
                            -2.0      -1.0       0.0       1.0
```

Although not statistically significant, the data showed slightly larger bias in depth perception for the MR condition, -0.6 mm (SD = 2.2 mm), and even larger in the IVE condition, about -2.1 mm (SD = 2.2 mm), compared with the RE condition, 0.4 mm (SD = 1.0 mm). The negative signs indicated that on average, the virtual octahedron was adjusted farther away from the participant than the real or virtual fixed cylinder.

Depth Perception Accuracy

Similar to the bias measure, ANOVA and confidence intervals were used for statistical analysis of depth perception accuracy. Figure 13 depicts the histograms and plots for the measure of accuracy.

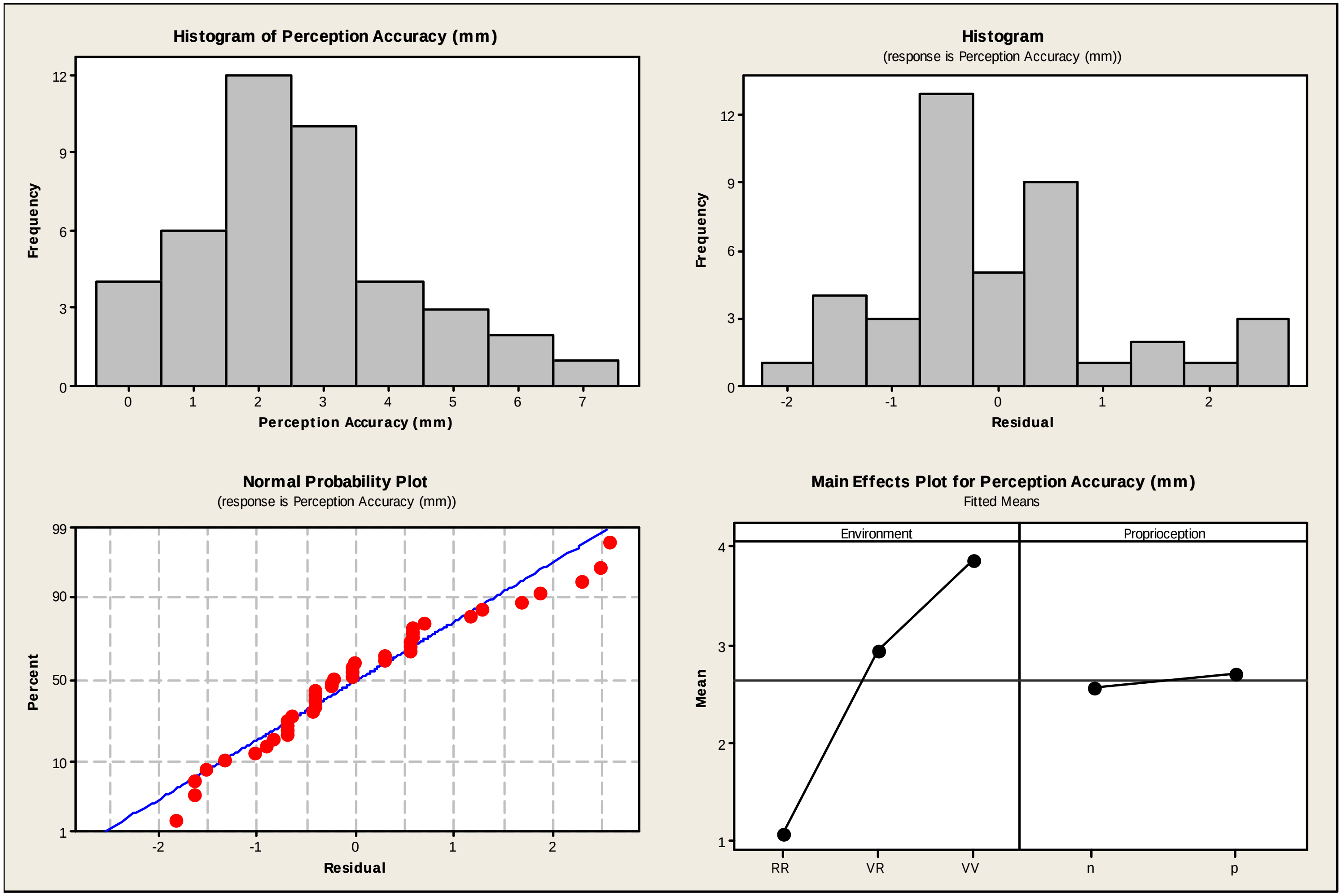


Figure 13. Depth Perception Accuracy Plots.

The analysis (Table 15) showed significant difference in mean accuracy between the environment types.

Table 15. ANOVA, Depth Perception Accuracy

```
Analysis of Variance for Accuracy (mm), using Adjusted SS for Tests

Source                      DF   Seq SS   Adj SS   Adj MS      F      P
Environment                  2   57.516   57.516   28.758  21.19  0.000
Proprioception               1    0.229    0.229    0.229   0.17  0.684
Environment*Proprioception   2    6.743    6.743    3.372   2.48  0.098
Error                       36   48.849   48.849    1.357
Total                       41  113.336
```

Table 16. ANOVA and Confidence Interval, Depth Perception Accuracy

```
Source          DF       SS       MS      F      P
Environment      2   57.516  28.7579  21.19  0.000
Proprioception   1    0.229   0.2288   0.17  0.684
Interaction      2    6.743   3.3717   2.48  0.098
Error           36   48.849   1.3569
Total           41  113.336

                      Individual 95% CIs For Mean Based on
                      Pooled StDev
Environment     Mean  ------+---------+---------+---------+---
RR           1.07143  (----*----)
VR           2.95000                (-----*----)
VV           3.88571                        (----*-----)
                      ------+---------+---------+---------+---
                          1.2       2.4       3.6       4.8
```

ANOVA and confidence interval (Table 16) for the main effects were computed to determine if the three types of environment exhibited a significant difference in accuracy. The analysis indicated a statistical difference in at least one of the environment types. The 95% confidence interval plot indicated that the RE condition is statistically different from the MR and IVE conditions. The RE condition, where both of the stimuli, the octahedron and the cylinder, were real, showed high accuracy, 1.7 mm (SD = 0.9

mm). Mean accuracy for the MR and IVE conditions were 3.0 mm (SD = 1.2 mm) and 3.9 mm (SD = 1.7 mm), respectively.

Although at the α = 0.05 level, the analysis (Table 16) indicated no significant difference for the interaction term, the P-value of 0.09 is relatively close to the α value. The interaction term would be significant if a slightly higher alpha value had been chosen. Figure 14 provides the interaction plots, which showed that RE clearly afforded better accuracy than MR and IVE in the proprioception-absent condition and that RE and MR afforded better accuracy than IVE in the proprioception-present condition. These two observations were statistically significant, as shown by the ANOVA analysis and confidence plot (Table 17).

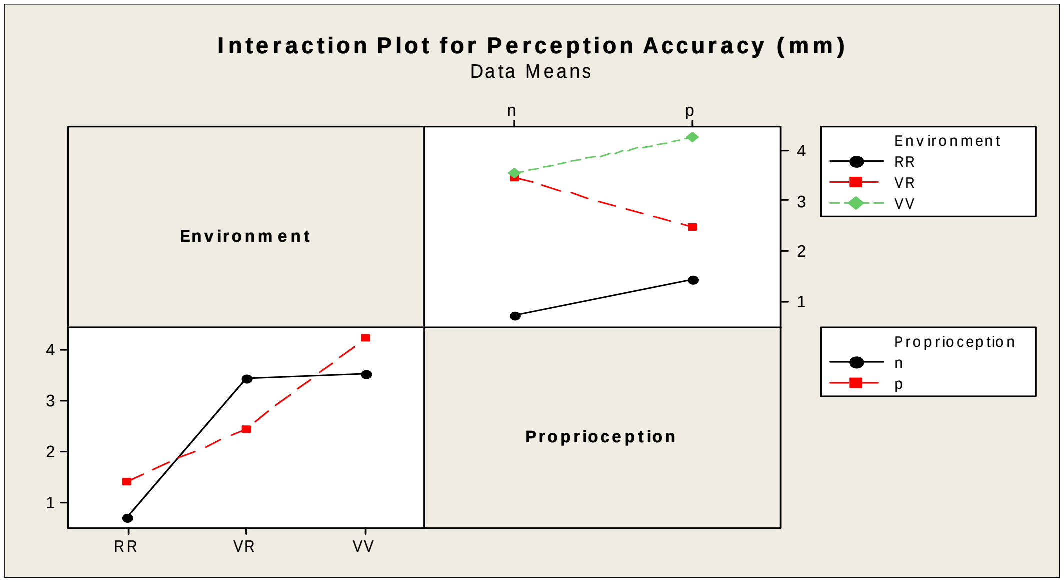


Figure 14. Depth Perception Accuracy Interaction Plots.

Table 17. ANOVA for Depth Perception Accuracy

```
ANOVA: Perception Accuracy (mm) vs. Environment (Proprioception Absent)
Source       DF     SS     MS      F      P
Environment   2  35.87  17.93  11.26  0.001
Error        18  28.68   1.59
Total        20  64.55
                        Individual 95% CIs For Mean Based on
                        Pooled StDev
Level  N   Mean  StDev  --+---------+---------+---------+-------
RR     7  0.714  1.113  (------*-----)
VR     7  3.443  1.370                    (------*------)
VV     7  3.529  1.291                     (------*-----)
                        --+---------+---------+---------+-------
                        0.0       1.5       3.0       4.5
```

```
ANOVA: Perception Accuracy (mm) vs. Environment (Proprioception Present)
Source       DF     SS     MS      F      P
Environment   2  28.39  14.19  12.67  0.000
Error        18  20.17   1.12
Total        20  48.56
                        Individual 95% CIs For Mean Based on
                        Pooled StDev
Level  N   Mean  StDev  -----+---------+---------+---------+----
RR     7  1.429  0.535  (------*------)
VR     7  2.457  0.526          (------*------)
VV     7  4.243  1.673                         (------*------)
                        -----+---------+---------+---------+----
                           1.2       2.4       3.6       4.8
```

Depth Perception Precision

ANOVA and confidence intervals with corresponding plots were again used for analyzing depth perception precision. Figure 15 shows the plots and histograms for the data collected.

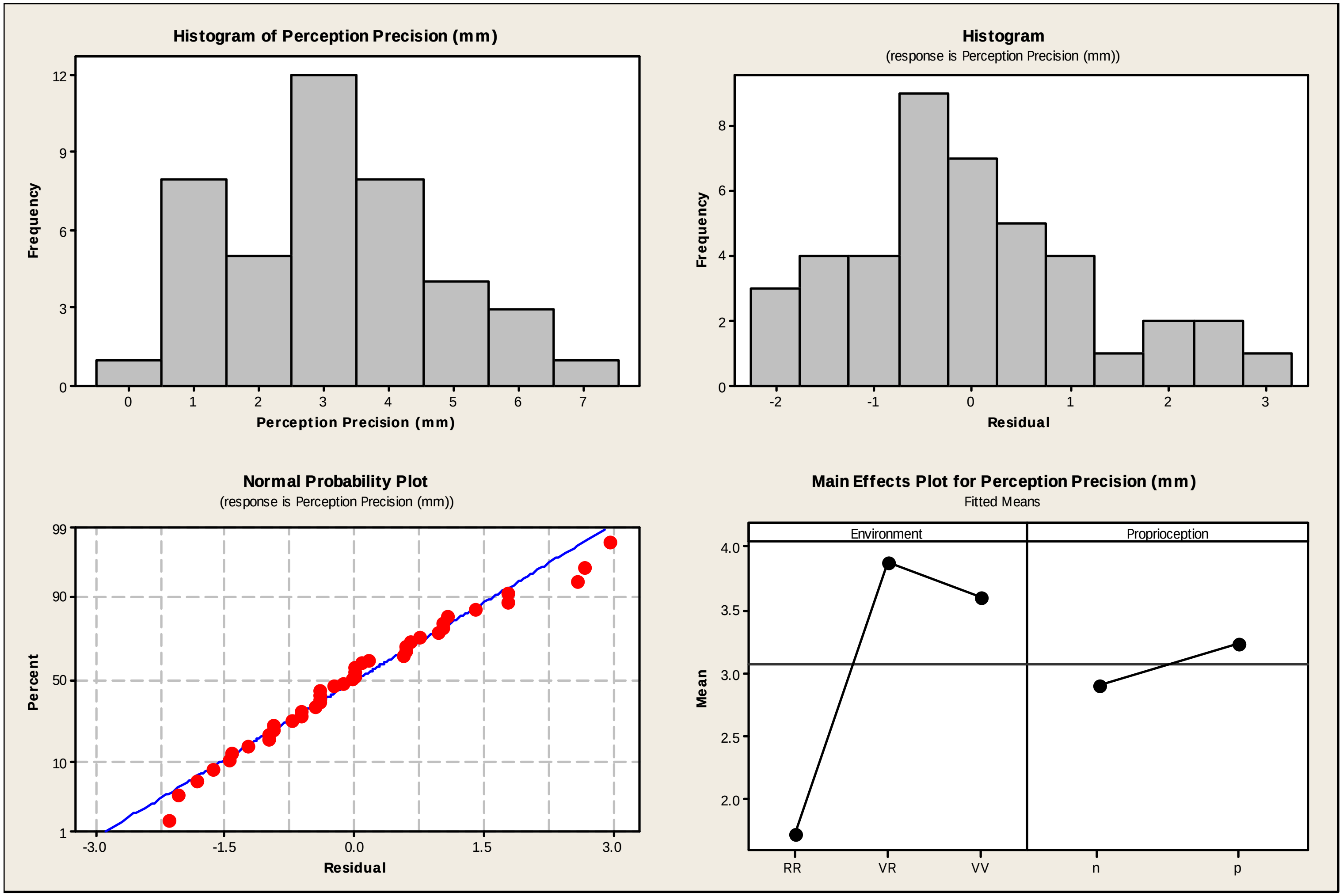


Figure 15. Depth Perception Precision Plots.

Environment type, again, had at least one mean that was significantly different and the interaction between the two factors was not significant (Table 18).

Table 18. ANOVA for Depth Perception Precision

```
Analysis of Variance for Precision (mm), using Adjusted SS for Tests

Source                      DF   Seq SS  Adj SS  Adj MS      F      P
Environment                  2   39.086  39.086  19.543  11.14  0.000
Proprioception               1    1.069   1.069   1.069   0.61  0.440
Environment*Proprioception   2    6.095   6.095   3.047   1.74  0.190
Error                       36   63.140  63.140   1.754
Total                       41  109.390
```

ANOVA and confidence intervals for the main effects indicated a significant difference in precision only in the RE condition. Precision for the MR and IVE conditions were not significantly different as shown in Table 19. Participants clearly performed with better precision for the RE condition, with a mean precision of 1.7 mm (SD = 1.1 mm). Table 19 shows the means for the environments and with confidence intervals. For the MR and IVE conditions, participants averaged 3.9 mm (SD = 1.3 mm) and 3.6 mm (SD = 1.7 mm), respectively, for precision.

Table 19. ANOVA and Confidence Interval for Depth Perception Precision

```
Source          DF       SS       MS      F      P
Environment      2   39.086  19.5431  10.73  0.000
Proprioception   1    1.069   1.0688   0.59  0.448
Error           38   69.235   1.8220
Total           41  109.390

                      Individual 95% CIs For Mean Based on
                      Pooled StDev
Environment      Mean   +---------+---------+---------+---------
RR            1.71429   (------*------)
VR            3.88571                         (------*------)
VV            3.60714                      (------*------)
                        +---------+---------+---------+---------
                      1.0       2.0       3.0       4.0


                         Individual 95% CIs For Mean Based on
                         Pooled StDev
Proprioception      Mean --+---------+---------+---------+-------
n                2.90952 (--------------*--------------)
p                3.22857         (--------------*--------------)
                         --+---------+---------+---------+-------
                         2.40      2.80      3.20      3.60
```

Observations

Hypothesis 3a, H_a: Error < 10mm, was supported by the empirical data and statistical analysis above. Whether bias, accuracy, or precision was used as the measure of spatial error, the results clearly showed that mean error for all types of environments, whether proprioception cue were provided or not, was less than 4 mm with a standard deviation of approximately 2 mm (1.6 mm for accuracy and precision and 2.2 mm for bias).

The empirical data showed that the difference in depth perception in RE compared with MR and IVE was statistically significant. Depth perception in RE was more accurate and more precise than that in MR and IVE. Although statistically significant, this difference was practically small, less than 3.0 mm (SD = 1.2 mm).

Hypothesis 3b, H_a: $Error_{proprioception} < Error_{no\text{-}proprioception}$, was not supported by the data. The results failed to show significant difference attributable to the addition of proprioception cue. While this hypothesis was not supported, it is noted that the spatial error measured turned out to be relatively small, less than 4 mm (SD = 2.2 mm) in all environments, RE, MR, or IVE. With such small error to begin with, further improvement attributable to the additional sensory modality of proprioception was difficult and was not found. Although not statistically significant, the results showed a slight difference in bias, which was a measure of overshoot when adjusting the octahedron. On average, the virtual octahedron, used in the MR and IVE conditions, was adjusted about 2 mm (SD = 2.2 mm) farther away from the participant compared with the real octahedron used in the MR condition. One explanation for the difference in bias could be the position of the retro-reflective screen, which was another real object and an unintended real image in the experiment. The image screen was placed slightly, about 20 mm, farther away from the participant than the fixed stimuli, the cylinder. Ellis and Menges (1997, 1998) showed that depth of a virtual object could be perceived further than actual position, if a real object further in depth was also in the scene.

The spatial errors observed were relatively small, comparable to experimental apparatus tolerances. Of the three measures for spatial error, precision is the most reliable measure. Precision is a measure of standard deviation and is affected only by the precision of the measurement equipment, measured at 2 mm, not the equipment's or setup's absolute accuracy, which was likely worse.

Experiment Four: Interaction, Full Factorial

Experiment Four measured spatial error and completion time associated with object interaction. Control variables included environment type, haptic condition, and audio condition in a full factorial design. Environment types included R/R and R/V representing RE and MR. Haptic conditions included the presence and absence of a simple touch cue. Audio conditions included presence and absence of a simple sound cue. Figure 16 illustrates the tapping task in the RE condition, with haptics and audio cues presented. Figure 17 shows the same task in the MR condition with only audio cue present.

Figure 16. Experiment Four, RE with Haptics and Audio Present.

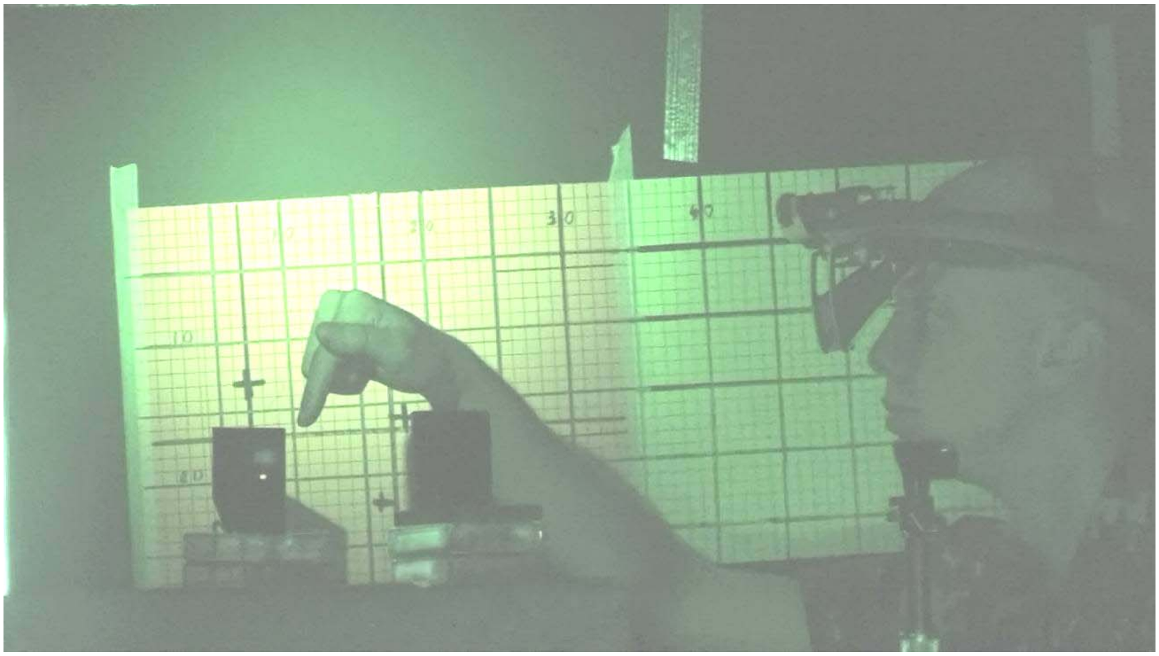


Figure 17. Experiment Four, MR with Audio Present.

Data Collection

Some anomalies observed in Experiments One and Two were avoided in this final experiment. These anomalies were due largely to a number of device limitation and use errors. The display device was repaired for dramatic increase in brightness which aided in alignment procedure that was drastically enhanced for this experiment. The experiment venue and participant pool were carefully selected. The experiment was carried out at the Navy base in Jacksonville using a highly dependable participant pool mostly of its military students and instructors. The procedure was revised to minimize the necessity for tinkering with the display or graphics software, while including frequent alignment check and quick adjustments as necessary. Real-time computer display of raw experimental data and images were observed carefully by the researchers for unanticipated anomalies. Participants were given ample opportunities to provide feedback and repeat data points where unintended mistakes were made. The sessions were scheduled to exclude the half hour before their lunch period and to end no later than 3:30 pm to avoid the need to rush or be careless. Since participants were more engaged in the beginning of the session, pre-experiment procedures and waits, conducted in separate rooms by research assistants, were kept short. These procedures contributed significantly to reliability of the data.

Software code was enhanced to aid in integration of the graphics image with the HMPD and compensate for differences in participants. To compensate for slight misalignment in the HMPD, the software rotated the right eye view nominally around the Z (depth) axis by 0.9 degrees and the X (horizontal) axis by -0.6 degrees. The Y (vertical) rotation depended on the IPD settings. Cross-hairs were displayed before each experiment and all three axes were verified. Fine adjustments were made from these nominal setting only when misalignments were observed.

Measures (Responses, y), Sample Size (n)

Experiment Four had eight treatments. Each treatment had five participants randomly assigned from the group used in Experiment Three. Each participant contributed 32 data points from which the following measures were calculated for statistical analysis.

1) Depth Perception Bias: average of the spatial error for the 32 button taps. The distance was measured only in the Z (depth) dimension.

2) Depth Perception Accuracy: average of absolute value of the error for the 32 button taps.
3) Depth Perception Precision: standard deviation of the error for the 32 taps.
4) Performance Time: Time to complete the 32 button taps.

Data Analysis

Similar to the results found in Experiment Three, spatial error recorded for all treatments were surprisingly low, comparable with the test apparatus accuracy of 2 mm. Because task performance was highly accurate in all treatments, near optimal, significant differences between treatments were difficult to find from the statistical analysis. On the other hand, performance time data varied widely between participants. Data analysis provided insight into which factor contributed to task completion time.

Object Interaction Bias

Bias provided a measure that indicated whether each participant, on average, overshot (negative value) or undershot (positive value) when interacting with the virtual or real button. It was calculated for each participant by averaging the error distances, negative if farther and positive if closer, between the tip of the finger and the center of the button. As done in Experiment Three, Minitab, ANOVA, confidence intervals, and related plots were employed for statistical analysis. Figure 18 depicts the plots for the dependent measure of bias for object interaction. The histograms and normality plots do not show signs of non-normality. The main effect plots show minimal effect from environment, haptic, and audio on the measure of bias.

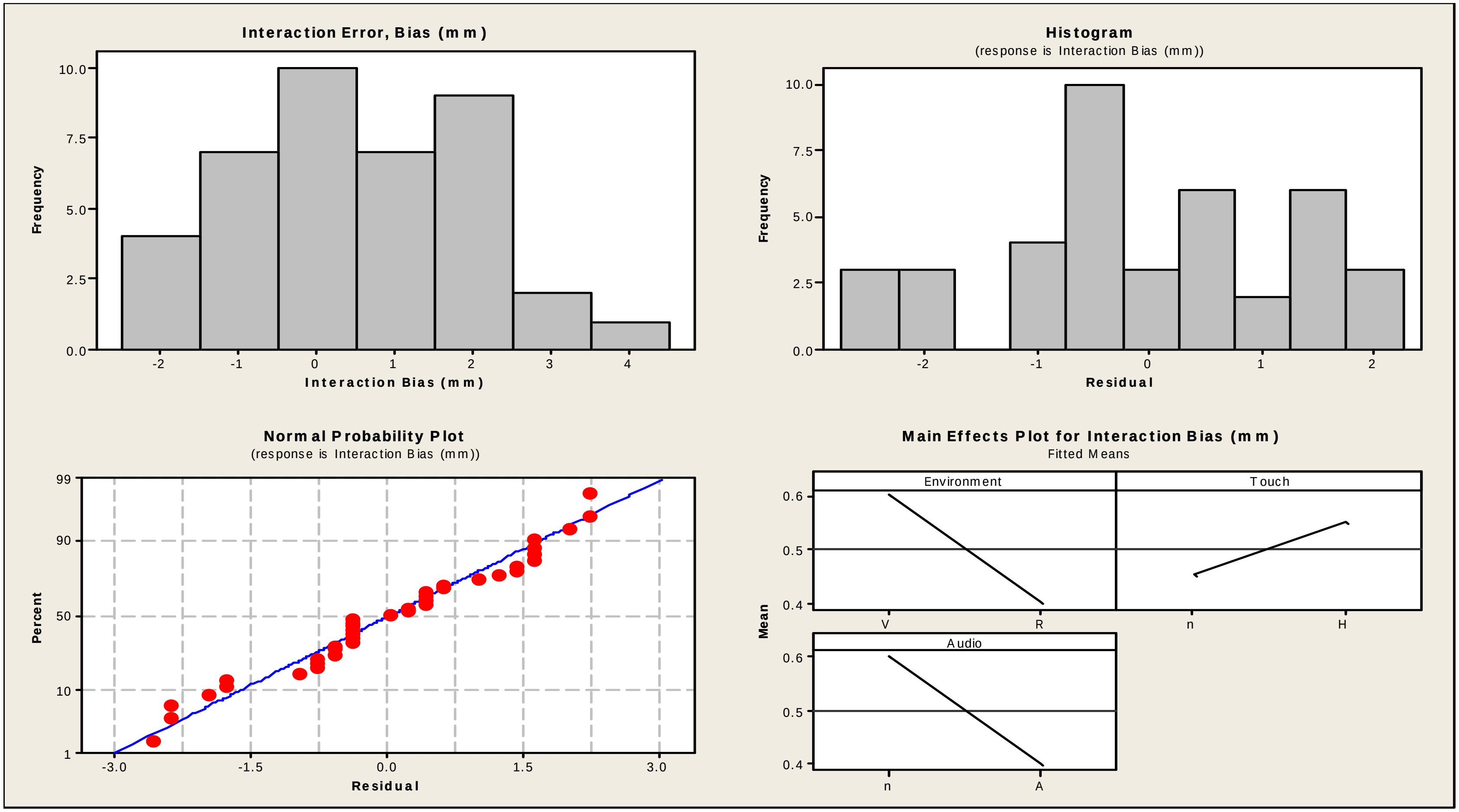


Figure 18. Object Interaction Bias Plots.

ANOVA and factorial fit for all main effects and first order interaction terms were computed. The results (Table 20 and Table 21) indicated no statistically significant main effect. It revealed one statistically significant and one marginally significant interaction term, Environment by Audio and Environment by Touch, respectively.

Table 20. Interaction Bias, Factorial Fit

```
Term                      Effect     Coef  SE Coef      T      P
Constant                          0.5000   0.2264   2.21  0.034
Environment              -0.2000  -0.1000   0.2264  -0.44  0.662
Touch                     0.1000   0.0500   0.2264   0.22  0.827
Audio                    -0.2000  -0.1000   0.2264  -0.44  0.662
Environment*Touch         0.9000   0.4500   0.2264   1.99  0.055  ***
Environment*Audio        -1.2000  -0.6000   0.2264  -2.65  0.012  ***
Touch*Audio              -0.3000  -0.1500   0.2264  -0.66  0.512
Environment*Touch*Audio  -0.1000  -0.0500   0.2264  -0.22  0.827
```

Table 21. Interaction Bias, ANOVA With First Order Interactions

```
Analysis of Variance for Interaction Bias (mm) (coded units)

Source              DF   Seq SS   Adj SS  Adj MS     F      P
Main Effects         3   0.9000   0.9000  0.3000  0.15  0.931
2-Way Interactions   3  23.4000  23.4000  7.8000  3.80  0.019
3-Way Interactions   1   0.1000   0.1000  0.1000  0.05  0.827
Residual Error      32  65.6000  65.6000  2.0500
  Pure Error        32  65.6000  65.6000  2.0500
Total               39  90.0000
```

Interaction plots were explored to assess the significance of the two interaction terms. The plots (Figure 19) showed that haptics may improve performance with the virtual button, but may actually diminish performance in the real environment. Likewise, the plots showed potential improvement with the audio cue added to the virtual-buttons condition and a decline with the real-buttons condition. However, these increases and decreases were small, from a spatial error bias of one to zero millimeter in improvement and from zero to one millimeter in degradation. This difference was actually below the experiment apparatus accuracy of about 2 mm. The improvement or decline observed in the Bias measure was thus not of practical significance and may be due to equipment variation or an unintended experimental procedure such as rounding errors.

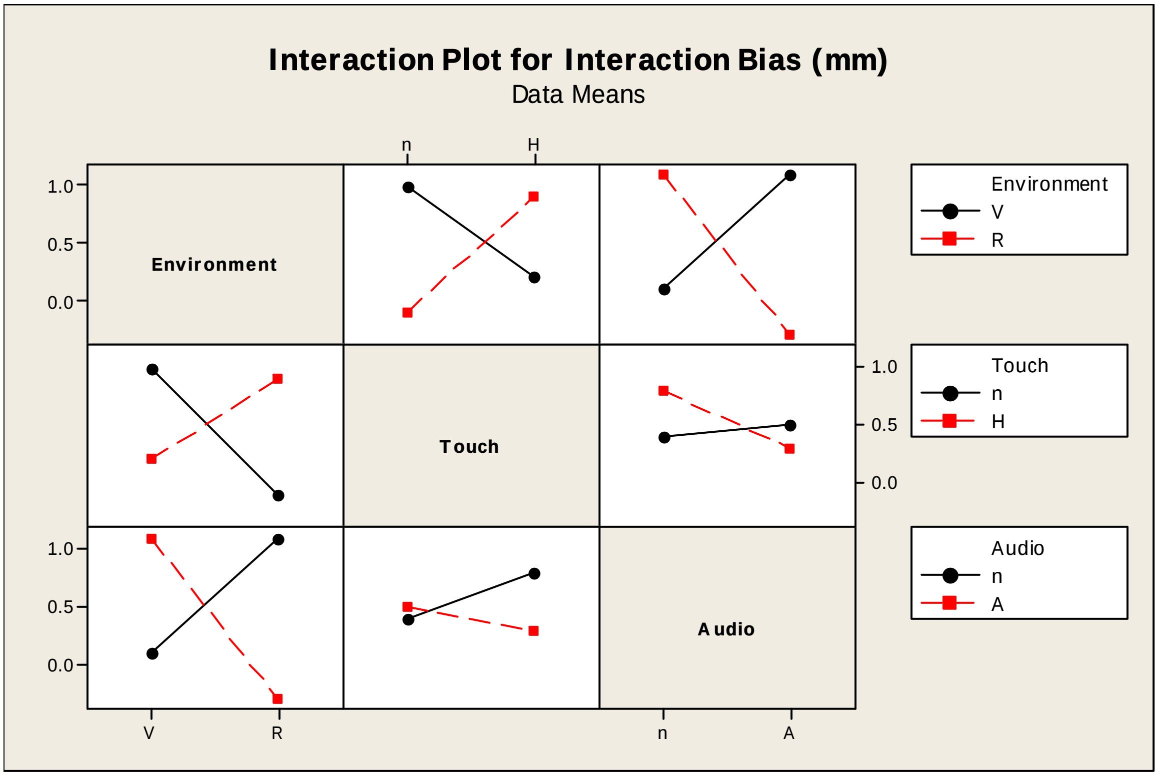


Figure 19. Object Interaction Bias, Interaction Plots.

ANOVA analyses (Table 22) were conducted separately for the haptic-absent, haptic-present, audio-absent, and audio-present sets of data, each containing 20 of the 40 total observation points (40 participants) to further evaluate the interaction terms. The analysis showed a statistical difference in performance accuracy between real and virtual buttons in the audio present condition. Again, the error was only about one millimeter, within measurement instrument accuracy, and thus was not of practical significance. The remaining three tests as described above showed no statistical significance.

Table 22. ANOVA, Object Interaction Bias, Audio & Haptics Present/Absent

```
One-way ANOVA: Interaction Bias (mm) versus Environment (Audio Absent)
Source       DF     SS    MS     F      P
Environment   1   5.00  5.00  1.97  0.178
Error        18  45.80  2.54
Total        19  50.80

One-way ANOVA: Interaction Bias (mm) versus Environment (Audio Present)
Source       DF     SS    MS     F      P
Environment   1   9.80  9.80  6.08  0.024
Error        18  29.00  1.61
Total        19  38.80

One-way ANOVA: Interaction Bias (mm) versus Environment (Haptic Absent)
Source       DF     SS    MS     F      P
Environment   1   6.05  6.05  2.80  0.112
Error        18  38.90  2.16
Total        19  44.95
One-way ANOVA: Interaction Bias (mm) versus Environment (Haptic Present)
Source       DF     SS    MS     F      P
Environment   1   2.45  2.45  1.04  0.322
Error        18  42.50  2.36
Total        19  44.95
```

Overall, average bias for each condition – virtual, real, audio present, audio absent, haptics present, and haptics absent – was less than 1 mm with standard deviation also less than 1 mm.

Object Interaction Accuracy

Similar to the previous measure, ANOVA, confidence intervals, and associated plots were used for statistical analysis on the dependent measure of object interaction accuracy. Again, accuracy was calculated from the average of the absolute values of the error distances computed for each participant. Figure 20 depicts the histogram and plots for the dependent measure of accuracy of object interaction.

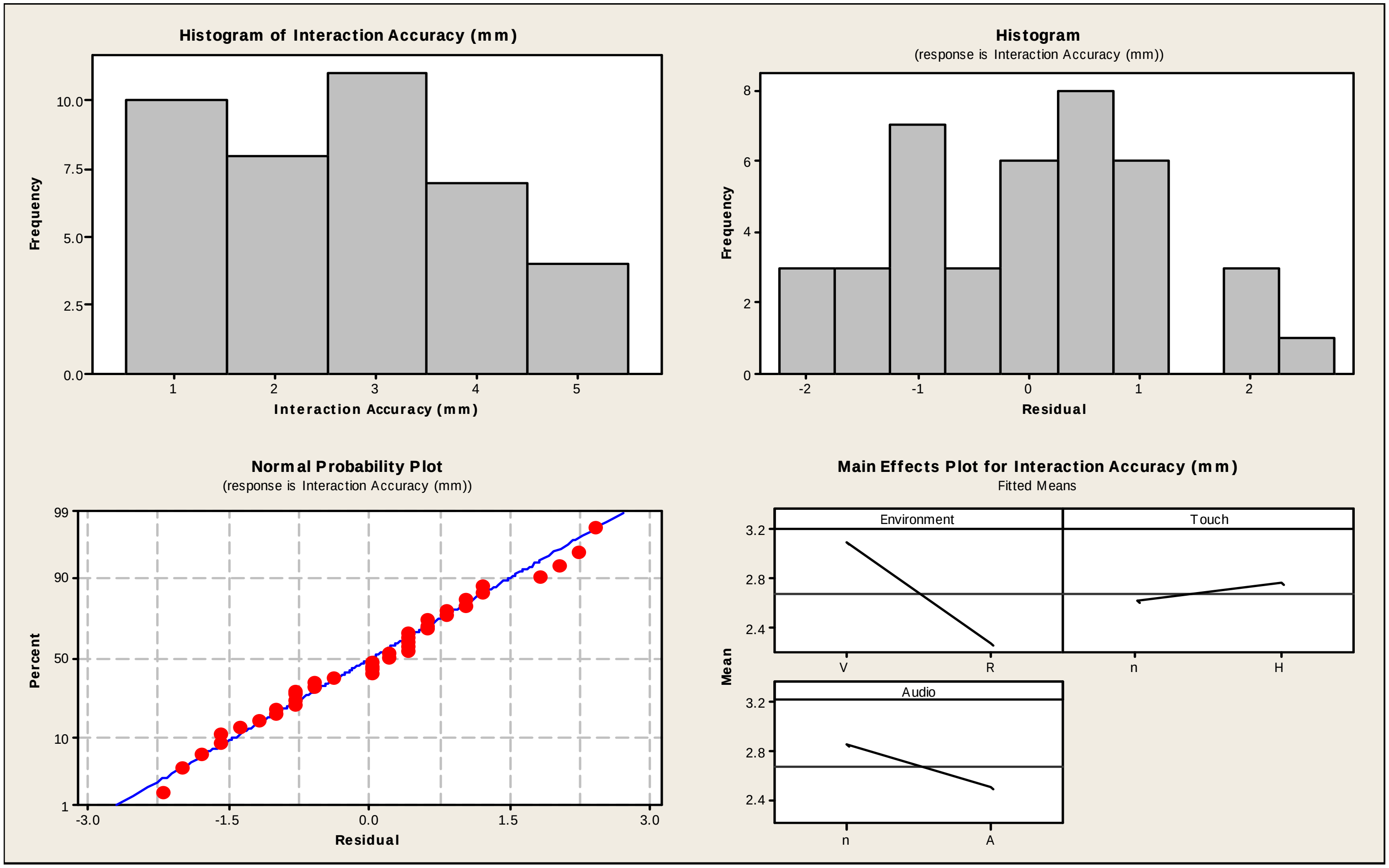


Figure 20. Object Interaction Accuracy Plots.

ANOVA and 95% confidence intervals (Table 23) showed no significant effect from any main factor or interaction terms. Mean accuracy was 3.1 mm (SD = 1.2 mm) for interaction with virtual buttons and 2.2 mm (SD = 1.3 mm) for interaction with real buttons.

Table 23. Object Interaction Accuracy, ANOVA and Confidence Intervals

```
Analysis of Variance for Interaction Accuracy (mm) (coded units)

Source              DF   Seq SS   Adj SS  Adj MS     F      P
Main Effects         3   8.6750   8.6750  2.8917  1.78  0.171
2-Way Interactions   3   5.8750   5.8750  1.9583  1.21  0.324
3-Way Interactions   1   0.2250   0.2250  0.2250  0.14  0.712
Residual Error      32  52.0000  52.0000  1.6250
  Pure Error        32  52.0000  52.0000  1.6250
Total               39  66.7750


                   Individual 95% CIs For Mean Based on
                   Pooled StDev
Environment  Mean  ------+---------+---------+---------+---
V            3.10                   (----------*----------)
R            2.25  (----------*----------)
                   ------+---------+---------+---------+---
                       2.00      2.50      3.00      3.50


             Individual 95% CIs For Mean Based on
             Pooled StDev
Touch  Mean  --+---------+---------+---------+-------
n      2.60  (---------------*----------------)
H      2.75      (----------------*---------------)
             --+---------+---------+---------+-------
             2.10      2.45      2.80      3.15


             Individual 95% CIs For Mean Based on
             Pooled StDev
Audio  Mean  --+---------+---------+---------+-------
n      2.85           (-------------*-------------)
A      2.50  (--------------*-------------)
             --+---------+---------+---------+-------
             2.00      2.40      2.80      3.20
```

Object Interaction Precision

ANOVA and confidence intervals were also used for statistical analysis for the measure of precision. Figure 21 provides the histograms and plots for object interaction precision.

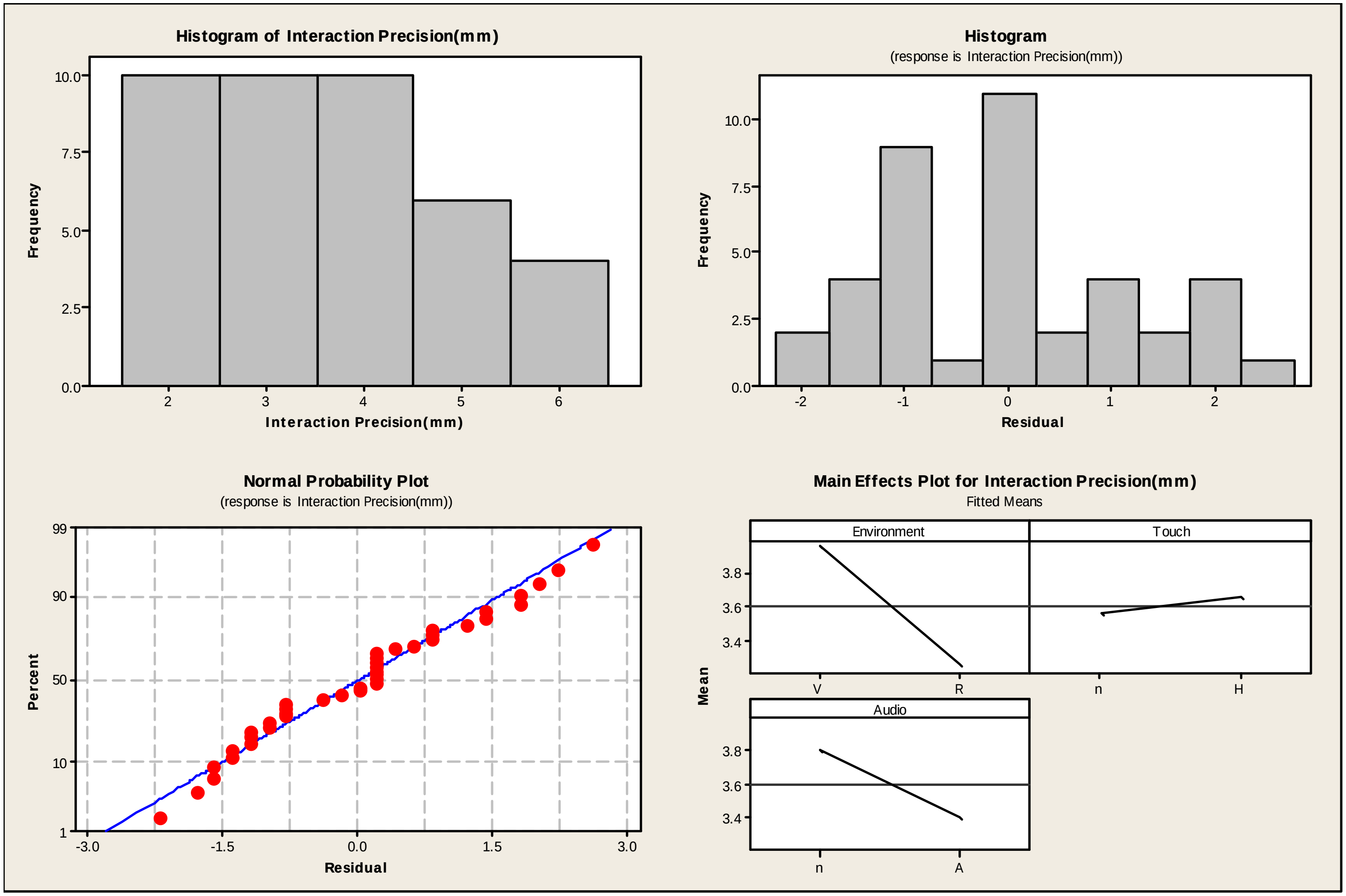


Figure 21. Object Interaction Precision Plots.

Similar to the previous measures, no main factor or interaction term were statistically significant (Table 24). Mean precision was 3.95 mm (SD = 1.2 mm) for interaction with virtual buttons and 3.3 mm (SD = 1.3 mm) for interaction with real buttons.

Table 24. Object Interaction Precision, ANOVA and Confidence Intervals

```
Analysis of Variance for Interaction Precision(mm) (coded units)

Source              DF   Seq SS   Adj SS  Adj MS     F      P
Main Effects         3   6.6000   6.6000  2.2000  1.25  0.309
2-Way Interactions   3   1.0000   1.0000  0.3333  0.19  0.903
3-Way Interactions   1   1.6000   1.6000  1.6000  0.91  0.348
Residual Error      32  56.4000  56.4000  1.7625
  Pure Error        32  56.4000  56.4000  1.7625
Total               39  65.6000


Analysis of Variance for Interaction Precision(mm), using Adjusted SS for Tests

Source       DF  Seq SS  Adj SS  Adj MS     F      P
Environment   1   4.900   4.900   4.900  2.99  0.092
Touch         1   0.100   0.100   0.100  0.06  0.806
Audio         1   1.600   1.600   1.600  0.98  0.330
Error        36  59.000  59.000   1.639
Total        39  65.600


                   Individual 95% CIs For Mean Based on
                   Pooled StDev
Environment  Mean  -------+---------+---------+---------+--
V            3.95                (-----------*-----------)
R            3.25  (-----------*-----------)
                   -------+---------+---------+---------+--
                        3.00      3.50      4.00      4.50


             Individual 95% CIs For Mean Based on
             Pooled StDev
Touch  Mean  -----+---------+---------+---------+----
n      3.55  (---------------*----------------)
H      3.65     (---------------*----------------)
             -----+---------+---------+---------+----
                3.15      3.50      3.85      4.20


             Individual 95% CIs For Mean Based on
             Pooled StDev
Audio  Mean  ---------+---------+---------+---------+
n       3.8            (-------------*-------------)
A       3.4  (-------------*-------------)
             ---------+---------+---------+---------+
                    3.20      3.60      4.00      4.40
```

Object Interaction Performance Time

Reaction or performance time was the duration taken to tap the buttons 32 times. Figure 22 depicts the histogram and plots for object interaction performance time.

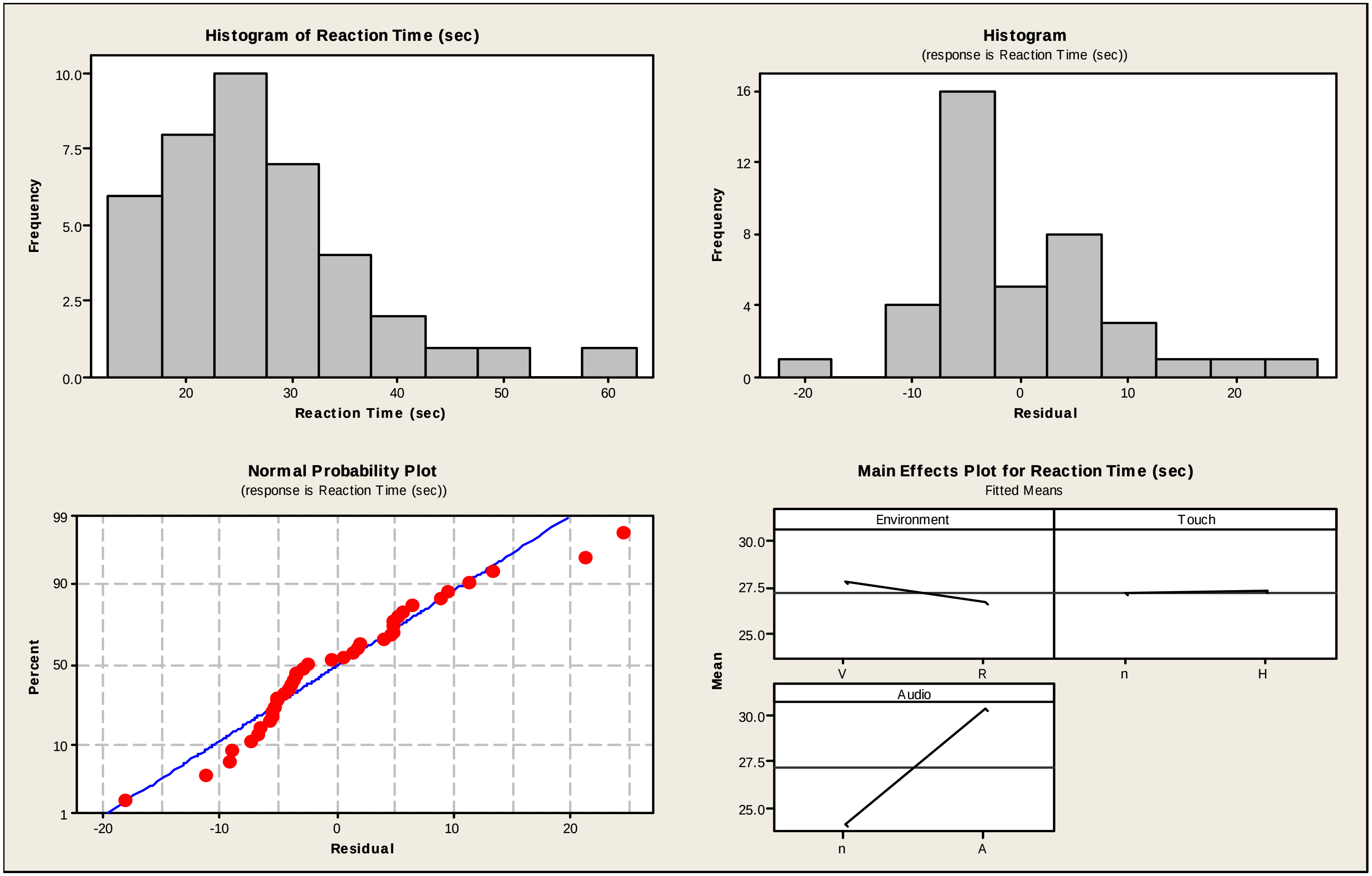


Figure 22. Object Interaction, Performance Time Plots.

ANOVA analysis (Table 25) revealed only one significant term, audio, which was a main factor. Surprisingly, the effect of adding audio was opposite of prediction. The addition of audio slowed performance time. The added cue increased completion time from a mean of 24 seconds (SD = 8 seconds) to 30.3 seconds (SD = 10.6 seconds).

Table 25. Performance Time, ANOVA

```
Estimated Effects and Coefficients for Reaction Time (sec) (coded units)

Term                      Effect    Coef  SE Coef      T      P
Constant                          27.198    1.473  18.46  0.000
Environment               -1.065  -0.532    1.473  -0.36  0.720
Touch                      0.065   0.032    1.473   0.02  0.983
Audio                      6.215   3.107    1.473   2.11  0.043
Environment*Touch         -4.775  -2.388    1.473  -1.62  0.115
Environment*Audio          3.915   1.957    1.473   1.33  0.193
Touch*Audio               -4.975  -2.487    1.473  -1.69  0.101
Environment*Touch*Audio    2.225   1.113    1.473   0.76  0.456
```

The analysis (Table 25) also showed two interaction terms, Environment by Touch and Audio by Touch, which were not statistically significant, but were close enough to warrant further exploration with interaction plots (Figure 23) and ANOVA analysis for each level of the Touch Factors independently (Table 26) . The plots showed a practically large increase in performance time, by about 12 seconds, when audio cue was added in the haptic-absent condition. The ANOVA analysis also revealed a statistical difference (p-value = 0.026) in performance time between audio-present and audio-absent in the haptic-absent condition. The three other tests, one for the haptic-absent condition and two for the haptic-present condition, as shown in Table 26, revealed no statistical significance.

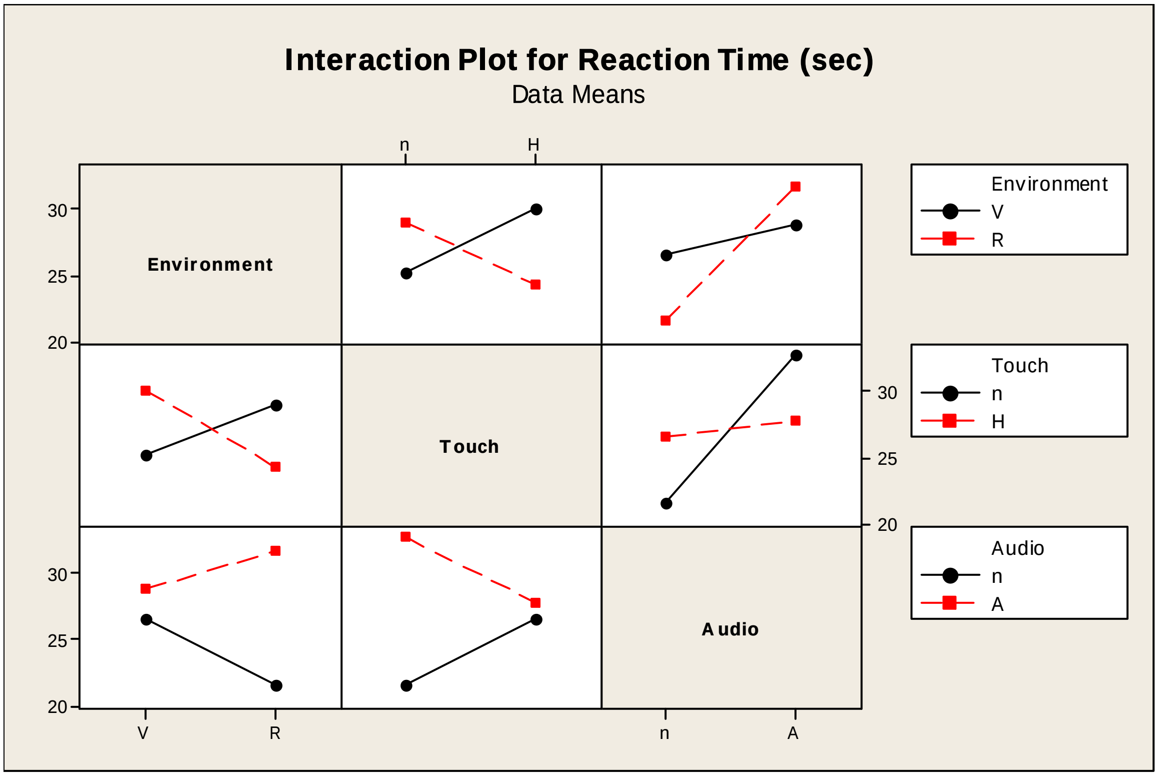


Figure 23. Performance Time, Interaction Plots.

Table 26. Performance Time, ANOVA, Haptics Present & Haptics Absent

```
One-way ANOVA: Reaction Time (sec) versus Env (Haptic Asbsent)
Source       DF    SS   MS     F      P
Environment   1    69   69  0.50  0.487
Error        18  2459  137
Total        19  2528
```

```
One-way ANOVA: Reaction Time (sec) versus Env (Haptic Present)
Source       DF      SS     MS     F      P
Environment   1   170.5  170.5  2.66  0.121
Error        18  1155.5   64.2
Total        19  1326.0
```

```
One-way ANOVA: Reaction Time (sec) versus Audio (Haptic Absent)
Source  DF    SS   MS     F      P
Audio    1   626  626  5.92  0.026
Error   18  1902  106
Total   19  2528
```

```
One-way ANOVA: Reaction Time (sec) versus Audio (Haptic Present)
Source  DF      SS    MS     F      P
Audio    1     7.7   7.7  0.10  0.750
Error   18  1318.3  73.2
Total   19  1326.0
```

Observations

Hypothesis 4a, H_a : Error < 10 mm, was supported by the data collected. Mean error for interaction with real or virtual buttons was, surprisingly, less than 4 mm with standard deviation of 1.6 mm or less whether bias, accuracy, or precision was used as the measure of spatial error. This held true for all experimental treatments, with real or virtual buttons and with or without haptic and audio cues.

Hypothesis 4b, H_a : $\text{Error}_{RE} < \text{Error}_{MR}$; H_a : $\text{Time}_{RE} < \text{Time}_{MR}$, was not completely supported by the empirical data. Spatial error, in terms of bias, accuracy, and precision were slightly less for interaction with real buttons compared to interactions with virtual buttons, but were not statistically significant. Although statistical significance was not found, it is noted that spatial error measured for all eight treatments were surprisingly found to be near optimal, less than 4 mm, not leaving much room for enhancement by the addition of other sensory modalities. Using an α level of 0.1 instead of 0.05 would have yielded different results for the measure of precision, but not bias nor accuracy, and again, the difference would still be practically small. For performance time, there was no evidence of any difference at all between interaction with real and virtual buttons. Again, while no significant difference was found, performance time was surprisingly near optimal, about one second per tap, for all treatments. The optimal result was likely due to the high fidelity of the visual cues provided for all treatments, MR or RE.

Hypothesis 4c, H_a : $\text{Performance Time}_{audio} < \text{Performance Time}_{no\text{-}audio}$, was also not supported by the data collected. On the contrary, the opposite was found to be true.

Audio is a temporal cue and its addition was predicted to enhance performance time. However, the results showed significant degradation in performance time with audio cue added. This was likely due to the experimental apparatus. The audio cues were not provided immediately when the buttons were touched. The audio feedback system, an Enforcer ® Alert System Model E-931CS22RC, had a noticeable delay of about 200 ms. Some participants hesitated and paced themselves waiting for the audio cue when tapping the buttons, especially when the haptic cue was absent.

Hypothesis 4d, H_a : $\text{Error}_{\text{haptic}} < \text{Error}_{\text{no-haptic}}$, was also not entirely supported. For interaction with virtual buttons, the addition of the touch haptic cue, on average, appeared to reduce the error, but by less than a millimeter. On the other hand, for interaction with real buttons, the cue appeared to increase the error, on average, by about the same amount. This relatively small difference is not of practical significance and could be due to an artifact of rounding errors during data collection and calculation or the experiment setup and apparatus.

A number of unexpected, subjective observations were made that are worth noting. Several participants expressed that they could not help but focus on the retro-reflective screen instead of the virtual object. Some took several tries over several minutes to wean themselves from focusing on the real screen to concentrating on the virtual buttons. Two expressed similarity of their experience with autostereograms, which generally take several minutes to get acclimated enough to focus at the proper depth and to see the stereo image.

Even though the room was kept completely dark so that the participant could not notice any real objects, the retro-reflective screen was lit, necessarily so, by the HMPD.

The hard, polished retro-reflective screen with high reflectivity had small engravings and minor scratches that were clearly visible. This was replaced during the experiment with a cloth material which had lower reflectivity and was less visible, but still not completely invisible. As discussed previously, having an unintended real object in the scene could have an effect on the results. For this experiment, the real screen may have made focusing on the virtual objects more difficult initially and some participants took longer to get adjusted to concentrating on the virtual buttons.

CHAPTER FIVE: CONCLUSION AND FURTHER RESEARCH

This dissertation set out to narrow a research gap in VE interaction seeking to quantify the spatial accuracy and performance time for direct interaction with virtual objects. Key to this research was development of a robust test bed that could provide highly accurate cues in personal space for object manipulation. Visual display technology was canvassed and compared. HMPD technology was suggested to provide for the most important cues visual cues in personal space as well as the most accuracy. As an optical see-through display, it is also one of the few technologies that allow for instantaneous synchronization with other salient sensory modalities in personal space, including proprioception and haptic.

A research scheme was architected to fully explore the feasibility of direct object manipulation using the highly accurate, light weight, proof of concept, prototype HMPD display and carefully designed and calibrated test bed. It included all sensory modalities and all VE types for comprehensive comparisons. Some performance differences were observed, but the results clearly show that mean error was surprisingly low, less than 4 mm (SD = 2.2 or less) and mean performance time is within 1 second (SD = 0.2 second) for the simple task of pushing a button using HMPD technology, regardless of the VE type (RE, MR, or IVE) or whether other sensory cues (haptics or audition) were provided.

Summary of Results

Four experiments were conducted to explore direct object interaction in personal space, to quantify accuracy and performance time achievable in VEs, and to gain insight on which factors contribute to these measures, positively or negatively. Each experiment built upon the previous, incrementally extending our knowledge of perception and interaction in RE, MR, and IVE, and with various sensory modalities included. Experiment One quantified error of depth perception in IVE with two different types of displays. Experiment Two quantified spatial error and performance time for object interaction in RE. Experiments Three and Four leveraged results from the previous two experiments as reference points in IVE and RE to explore the entire R-V spectrum and all salient sensory modalities.

Spatial Error for Depth Perception and Object Interaction

Using a display device that offered key visual cues in personal space, this study provided empirical evidence asserting that interaction in VE is highly effective, with a mean spatial error of 4 mm or less, given correct, accurate binocular visual cues – accommodation, convergence, stereoscopy, and occlusion – instantaneously synchronized with proprioception cues. The results were upheld regardless of visual environment type and whether additional sensory modalities were presented or not.

The results are summarized in Tables 27 and 28. These show low mean spatial error (bias, accuracy, and precision < 4 mm) in all environment types (RE, MR, and IVE) and for both tasks (depth perception and object interaction). For depth perception, the differences between RE and MR and between RE and IVE were statistically significant. The differences between MR and IVE were not.

Table 27. Summary of Depth Perception Experimental Data

Depth Perception Mean Error

	Bias (mm)	Accuracy (mm)		Precision (mm)	
RE	0.42	1.07		1.71	
MR	-0.63	2.95		3.89	
IVE	-2.06	3.89		3.61	
MR - RE	-1.05	1.88	$p < 0.001$	2.18	$p < 0.001$
IVE - RE	-2.48	2.82	$p < 0.001$	1.9	$p < 0.001$
MR - IVE	1.43	-0.94		0.28	

Table 28. Summary of Object Interaction Mean Error

Object Manipulation Mean Error

Buttons	Bias (mm)	Accuracy (mm)	Precision (mm)
Real	0.40	2.25	3.25
Virtual	0.60	3.10	3.95
Difference	0.20	0.85	0.70

Performance Time for Object Interaction

Performance time varied widely. Participants were asked to tap the buttons at a normal and comfortable pace. This pace varied widely even in RE. Regardless, the data showed a mean performance time of about one second per button tap (or button push). These results were similar for virtual or real buttons, with correct visual cues presented and synchronized with proprioception, and with or without haptic or audio feedback. These findings are consistent with Schiefele's (2000) findings, where mean performance time for manipulating real objects was measured at 1.5 seconds. The results, however, were, on average, shorter than Schiefele's findings for virtual object manipulation of 3.5 seconds. The main difference in the two studies was the visual display, its fidelity, and its synchronization with proprioception.

The data also showed a slight difference with audio cue provided, but in the opposite direction as predicted. This degradation in performance was likely due to the participants' reaction to the delay in the audio feedback system, which was not previously considered and was realized only after the empirical data was collected and analyzed. The difference was statistically significant for the haptic-absent condition, as shown in Table 29 for completion time of the 32 taps.

Table 29. Summary of Task Performance Time Data

Audio	***Performance Time (sec)***
Absent	33.6
Present	21.50
Difference	12.10 (p = 0.026)

(for Haptic-Absent Condition)

Reliability of the Experimental Data

Experiments Three and Four leveraged findings and methods of Experiments One and Two to produce highly reliable data. The key to attaining the high reliability was a dramatic increase in brightness in the display which facilitated careful optics alignment, vigilant adjustment of the display and graphics software to match the user, and judicious selection of participant pool.

The HMPD displayed a real, optical image that was seen clearly by the user and with some care, by the researcher where the room was completely darkened and the display brightness was set to maximum. Calibration cross-hairs were programmed and projected out by the left and right displays. The cross-hairs' exact positions in space, in all dimensions, X, Y, Z, and rotations, were unmistakable, as observed by the researcher and reported by the participants. This allowed for easy adjustment of the display before use, for verification during experimentation, and for rapid adjustment in between sessions as necessary.

Graphics software required detailed calculations. Field of view, IPD, and convergence settings were calculated, programmed and tested for accuracy. Integration with the HMPD was tested and measurements for the right and left displays were continuously verified. Software routines were coded for quick compensations for minor optical maladjustments, for participant's differences, and for occasional misalignments.

Perhaps the most important element contributing to the reliability of the experimental data was the quality of the participants. For Experiments Three and Four, participants were recruited from the Navy's C school at Naval Air Station (NAS), Jacksonville, Fl. With few exceptions, the participants were all highly engaged and took the experiments seriously. The students in the participant pool saw this as an assignment. The staff from the school saw the experiment as a challenge to outperform each other or boast of their performance. These military personnel had been through thorough eye exams beforehand. None had to be screened out due to poor eye sight. The participants expressed that they saw the depth perception and interaction experiments as another set of exams similar to previous exams that they wanted to do well on.

The accurate visual cues afforded by the HMPD, frequent verification and re-alignment through software routine, and quality participants contributed to the highly reliable empirical results. Consequently, Experiments Three and Four benefited from the best expected condition with visual cues and experimental controls carefully optimized. The low spatial error and fast completion time observed for each and every treatment in both Experiments Three and Four were very surprising. Experiments Three and Four unexpectedly reached a ceiling effect, which may very well be explained by the highly optimized experimental conditions achieved.

Additional Findings

The Simulation Sickness Questionnaire (SSQ) showed little side effects. Mean SSQ scores were less than five, indicating negligible symptoms (Kennedy et al., 1993). None of the 28 eight participants from Experiments One and Two indicated side effects. Two of the 42 participants in Experiments Three and Four indicated slight eye strain from trying to focus on the virtual buttons which were placed at 0.4 m directly in front of the participant. The two indicated that the symptoms immediately disappeared when they stopped focusing on the stimuli. Much care was given to alignment of the visual display, matching the optics parameter with the software and graphics, and adjustment for the user, which may have contributed to the lack of side effects.

The presence of real objects appeared to have noticeable effects, both positive and negative. Some participants expressed difficulty focusing on the buttons because they found themselves focusing on the screen about 0.3 m farther away. They expressed similarity with autostereograms, which generally take a minute or two to get used to before the stereo image is consistently seen. Some also expressed that being able to see the real hand helped them see the virtual buttons more easily. Before lifting the hand into the scene, they reported seeing double images of the buttons, which likely indicated that they were focusing on the screen until the hand came into the scene near the buttons. Others asked if the light could be turned on dimly as it helped them see the virtual buttons better. From a brightness standpoint, dialing up the room light made the virtual buttons less visible. On the other hand, it made nearby experimental apparatus and the hand more visible and the screen less noticeable. This also points to the positive effect of having nearby real objects next to virtual objects. It appears that participants tended to focus on real objects, which may have helped with awareness of and concentration on virtual objects nearby, but may have had negative effects with perception of virtual objects that were elsewhere (Ellis & Menges, 1997, 1998).

All participants in Experiments Three and Four expressed confidence in their ability to clearly perceive the virtual objects and know of their exact positions. Many compared the experiment to familiar eye exams and commented afterwards that they had “passed the test”, even though individual results were not provided to the participants.

Occlusion, or lack thereof, had a strong effect. After the experiment, some participants were asked to move a real reference rod right next to a virtual calibration line to verify alignment. It was noted that when the rod was placed right on the line, the participant tended to keep pushing the rod further away. Some expressed that the "line followed the stick". One potential explanation was occlusion. The participant may have expected the virtual line to eventually occlude the reference rod (which was not possible) and therefore kept on sliding it out. Another explanation would point back to the vividness or richness of the real object affecting the perception of the virtual object as discussed previously.

Two other subtle observations are worth noting. Pushing the reflective screen further away made the virtual objects blurry (due to non-perfect retro-reflection) but that did not affect accuracy. The few air crewmen in the participant group who had been trained in NVG seemed to be even more confident in the tests. There were not enough air crewmen participants in the study for a separate statistical analysis of this group, but the data did seem to show better performance.

Recommendations

It is evident that virtual object interaction with correct visual cues presented, including synchronization with proprioception, is highly accurate (4 mm) and highly effective. Performance in MR is comparable to RE, with less than three millimeter difference, whether mean bias, accuracy, or precision is used as the measure of spatial error. Likewise, task completion times for simple button pushes are similar for RE and MR, approximately one second. This performance level suits many VE applications adequately. Moreover, this study was performed with simple geometry models. Higher accuracy may be achieved if the VE takes advantage of richer models or if other visual cues, such as size and motion perspective are included. This performance level exceeds the accuracy requirements for many applications including cockpit simulations, where controls are set at about 12 mm apart, the minimum recommended layout for ergonomic designs (Boff et al., 1986). The HMPD and MR technology has matured to a level where more complex and practical experimentation can be conducted for application specific environments.

This study showed that in a high-fidelity visual environment, the addition of an inaccurate cue, specifically a delayed sound cue, degrades performance compared with absence of the cue. For VE designs involving button pushes, having no audio feedback may provide a more effective VE than having an unsynchronized one.

Further Research

This study was conducted with the goal of gaining a better understanding of which factors can improve MR towards the performance of RE. It has provided empirical evidence asserting that basic visual cue and synchronization are key to virtual object interaction. Performance in all environment types, RE, MR, and IVE was highly accurate and efficient, given salient, optimal visual cues. The empirical results observed seemed to have reached an unexpected ceiling effect due to optimal settings and selections of experimental controls. Further research is recommended using similar technology and test methods with varying, perhaps suboptimal visual cues in order to quantify their effects.

The task in this study was simple tapping or pushing of buttons. More complex manipulation tasks are recommended. Different types of virtual buttons, switches, and dials and different complex tools, such as virtual medical instruments would provide more comprehensive results that would be beneficial for a wide variety of VE communities. Perhaps performance for more complex tasks would not be so effective or efficient, even with optimal visual cues and experimental controls, and the ceiling effects may not be observed with the added complexity.

An unintended observation made in the study was the effect due to nearby real objects or their richness in visual cue. Scene richness seemed to enhance or degrade performance noticeably. Further research should include richness in the scene as a factor. Rich virtual and real objects should be looked at and compared with simple geometry. Perhaps the richness of the cues provided by the real or virtual object can partially replace the need for other more complex visual cues such as stereoscopy (Ellis & Menges, 1997, 1998).

Perhaps, the most practical next step is to push the development of the HMPD technology based on this successful prototype, which was not originally designed for human factor studies. Nevertheless, among the prototypes developed in the ODA Lab, this HMPD has been the most extensively used in human factor studies because of its compactness and extremely light weight that came as a consequence of using OLED microdisplays and its compactly integrated electronics. The HMPD technology holds much promise, combining the features of projection displays and HMD. It literally provides a real optical image in mid-air, similar to holograms or optical mirages. These images are optically real, whether generated by a computer or not, and from a physics perspective, interaction with its light rays is also physically real. This lends itself well to calibration and alignment routines compared with HMDs and may provide undiscovered benefits to more complex interactions in VE.

APPENDIX A. SIMULATION SICKNESS QUESTIONAIRE

Simulation Sickness Questionnaire (Kennedy et al., 1993)

C1. Simulation Sickness Questionnaire, Pre-exposure

Participant Number: ____ Time now:___________

Pre-exposure Symptom Checklist

Instructions: Please circle the severity of any symptoms that apply to you right now.

1.	General Discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Headache	None	Slight	Moderate	Severe
4.	Eye Strain	None	Slight	Moderate	Severe
5.	Difficulty Focusing	None	Slight	Moderate	Severe
6.	Increased Salivation	None	Slight	Moderate	Severe
7.	Sweating	None	Slight	Moderate	Severe
8.	Nausea	None	Slight	Moderate	Severe
9.	Difficulty Concentrating	None	Slight	Moderate	Severe
10.	Fullness of Head*	None	Slight	Moderate	Severe

*Fullness of head means internal pressure in head, similar to sinus pressure, such as one gets when hanging upside down

11.	Blurred Vision	None	Slight	Moderate	Severe
12.	Dizzy (Eyes Open)	None	Slight	Moderate	Severe
13.	Dizzy (Eyes Closed)	None	Slight	Moderate	Severe
14.	Vertigo**	None	Slight	Moderate	Severe

**Vertigo is a disordered state in which the person or his/her surroundings seem to whirl dizzily; loss of orientation that makes it difficult to perceive which way is up

15. Stomach Awareness***	None	Slight	Moderate	Severe

***Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea

16. Burping	None	Slight	Moderate	Severe

Are there any other symptoms that you are experiencing right now? If so, please describe the symptom(s) and rate their severity on the other side.

APPENDIX B. USE EXAMPLES OF THE TERMS “REAL” AND “VIRTUAL”

For this dissertation, the topic is manipulation of virtual controls (“virtual objects”) directly and naturally by the user's hand (a “real object”). One sees the hand (wearing the glove) faithfully as a “real image”. One also sees “computer-generated images” of the virtual buttons (“virtual objects”). Experiments with real hand (real object) manipulating real controls (real objects) will also be conducted for comparison. Here the user sees “real images” of both the hand and the objects.

Likewise, experiments with virtual hand (“virtual object”) manipulating virtual controls (“virtual object”) could also be conducted for comparison. The user sees models (“computer-generated images”) of both the hand and the objects.

Finally, there is one potentially confusing concept that is important to discuss and that uses the terms introduced at the beginning of Chapter Two - "optically real image" and "optically virtual image". It is in the comparison of Head Mounted Projection Displays (HMPDs) with other VE displays, especially HMDs. With HMDs, the computer-generated images, photographs, or video images seen by the observer are necessarily "optically virtual images". On the other hand, with HMPDs, similar images can be provided as "optically real images". In effect, "optically real images" of the display screen and the virtual objects in it (but not the physical screen nor the physical objects themselves) are presented in mid-air in front of the user. This is an important feature that has not seen much discussion in VE literature. Having an "optically real image", and therefore actual physical light rays focusing on and passing through it, allows one to naturally interfere with, or essentially touch the image, but without touching any physical screen nor objects.

No other VE display technology provides for an "optically real image" of the display screen in mid-air (i.e. without the physical screen in the same space). CAVEs provide "optically real images" but on a physical screen, not in midair. HMDs provide perceived images in midair but these are "optically virtual images", not "optically real images". This feature provides for experimentations involving interaction with virtual objects with high visual cue fidelity, i.e., with "optically real images", but with the total absence of haptics cue.

APPENDIX C. OSTHMD AND HMPD COMPARISON

The fundamental differences between a conventional optical see-through display and a head-mounted projection display coupled with retro-reflective material are listed in Table 30 (Rolland et al., 1998; Hua et al., 2000; Hua et al., 2001; Rolland et al., 2005).

Table 30. Advantages of HMPD vs. HMD

See-Through HMD	HMPD	
Eyepiece	Projective Lens	Fundamental Design Differences
Beam splitter reflects light into eye	Beam splitter reflects light away from eye, retro-reflective material reflects it back towards the eye	
No screen	Retro-reflective material	
No Occlusion	Correct Occlusion of virtual object by real objects	Fundamental Advantages of HMPD
Eyepiece design is heavier	Projective Lens are lighter	
Optical aberration and distortion from design are challenging to correct	Projection lens can be designed to more easily minimize aberrations with fewer elements and no distortion	
Provides only "optically virtual images" of the micro-displays (LCD/CRT screen)	Can provide "optically real images" of the micro-displays in front of the retro-reflective material or "optically virtual images" behind the material	
IPD can be set to match user but only approximately and is difficult to align (left/right adjustment of both beam splitters) with the user's eyes.	Reflection of the micro-display image out and away from the user facilitates alignment of the beam splitters to match the position of the user's eyes.	
Limited to 40 deg FOV using a flat combiner	Allows for wider FOV design, up to 90 deg, using flat combiner	
Real and computer-generated images are simply superimposed, which is visually incorrect. Computer generated image is visible at all look angle. Real and virtual environments are not mixed correctly.	Allows for diminished reality, i.e., visually remove or camouflage real objects. Strategic placement of reflective material effectively allows for proper mixing of virtual and real environment similar to capability provided by blue-screen technology	
Exit pupil and eye relief are limited by the eye piece allowable size	Can be designed for larger exit pupil and eye-relief for an equivalent optics size compared to eyepiece	

See-Through HMD	HMDP
Image plane can be set anywhere in the user space	Image plane should be set near the retro-reflective material to minimize blurring effects from diffraction; however depending on the application other settings are also allowed at the expense of some blurring
Resolution of micro-display is preserved	Imperfect reflection from retro-reflective material may produce image blur reducing effective resolution from that of the micro-display.
Bright images. Can be designed for near 0% to near 100% attenuation of light intensity for computer-generated image reflection from the beam splitter	Dim images. At least 75% attenuation of light intensity for computer-generated image due to reflection from and transmission through the beam splitter.
Fixed depth and size of computer-generated image	Variation in depth and size depending on position and orientation of the user if the retro-reflective material is imperfect.
Fixed exit pupil	Exit pupil shift as a function of distance from the screen if the retro-reflective material is imperfect.

Fundamental
Disadvantages
of
HMPD

APPENDIX D. ESTIMATE USING THE MODEL HUMAN PROCESSOR (HMP)

Table 31. Estimated time to push a button for IVE, RE, and MR

	Tasks Using Model Human Processor (MHP)			IVE (ms)	RE (ms)	MR (ms)
1	Perceive, transfer to working memory			Tp = 100	100	100
2	Decide to Locate Virtual Hand			Tc = 70		
3	Look at Virtual Hand			Tm = 70		
4	Perceive Virtual Hand		Localize	Tp = 100		
5	Decide to move hand			Tc = 70	Tc = 70	Tc = 70
6	Move hand			Tm = 70	Tm = 70	Tm = 70
7	Artificial Time Delay			Td = 60		
8	Perceive Match of hand and button			Tp = 100	Tp = 50	Tp = 100
9	Artificial Time Delay	loop=4	Select	Td = 60		
10	Decide if Match is good enough			Tc = 70	Tc = 70	Tc = 70
11	Select			Tm = 70	Tm = 70	Tm = 70
12	Perceive feedback of object selected	loop=2	Manipulate	Tp = 100	Tp = 50	Tp = 100
13	Decide to Manipulate			Tc = 70	Tc = 70	Tc = 70
14	Manipulate			Tm = 70	Tm = 70	Tm = 70
15	Perceive completion of Manipulation	loop=2	Release	Tp = 100	Tp = 50	Tp = 100
16	Decide to Release			Tc = 70	Tc = 70	Tc = 70
17	Release			Tc = 70	Tc = 70	Tc = 70
			Total	Tt = 3090	Tt = 1970	Tt = 2370

Assumptions:

1) The first loop is estimated to take about 1 second for RE so the author chose loop =4;
   Similarly, 2nd & 3rd loop =2
2) Perceptual processing time for matching the position of the hand and button, for selection, and for manipulation is less than 100 msec for the RE case because tactile feedback is present.
   This haptics cue provides for redundant information from a separate modality reducing response time (Boff, 1986, 11.420). Therefore, the author chose a reduced value of 50 msec, which is also The lower bound for perceptual processor time used by Card (1983).
3) For RE and MR virtual hand times don't apply
4) For RE, artificial delays for trackers and screen or frame update rates do not apply
5) If cues for IVE and MR are not perfect, there are more loops and the times are longer.
   The total time calculated does not account for this imperfection.

APPENDIX E. UCF INSTITUTIONAL REVIEW BOARD APPROVAL

University of Central Florida

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901, 407-882-2012 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

EXPEDITED CONTINUING REVIEW APPROVAL NOTICE

From : **UCF Institutional Review Board**
FWA00000351, Exp. 10/8/11, IRB00001138

To : **Long Nguyen** and Co-PIs: **Jannick P. Rolland, Linda C. Malone**

Date : **May 06, 2009**

IRB Number: **SBE-06-03534**

Study Title: **Analysis and Development of a Full Immersion Projective Head-mounted Display for Dismounted Infantry Training**

Dear Researcher,

This letter serves to notify you that the continuing review application for the above study was reviewed and approved by the IRB Chair on **5/6/2009** through the expedited review process according to 45 CFR 46 (and/or 21 CFR 50/56 if FDA-regulated).

Continuation of this study has been approved for a one-year period. The expiration date is 5/5/2010. This study was determined to be no more than minimal risk and the categories for which this study qualified for expedited review are:

4. Collection of data through noninvasive procedures (not involving general anesthesia or sedation) routinely employed in clinical practice, excluding procedures involving x-rays or microwaves. Where medical devices are employed they must be cleared/approved for marketing.

7. Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Subjects or their representatives must receive a copy of the consent form(s).

All data must be retained in a locked file cabinet for a minimum of three years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained on a password-protected computer if electronic information is used. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

To continue this research beyond the expiration date, a Continuing Review Form must be submitted 2 – 4 weeks prior to the expiration date. Use the Unanticipated Problem Report Form or the Serious Adverse Event Form (within 5 working days of event or knowledge of event) to report problems or events to the IRB. Do not make changes to the study (i.e., protocol methodology, consent form, personnel, site, etc.) before obtaining IRB approval. Changes can be submitted for IRB review using the Addendum/Modification Request Form. An Addendum/Modification Request Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

On behalf of Tracy Dietz, Ph.D., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 05/06/2009 01:14:21 PM EDT

Joanne Muratori

LIST OF REFERENCES

Allard, T. (1997). US Navy and Marine Corps requirements and challenges: Virtual environment and component technologies. *North Atlantic Treaty Organization Research and Technology Organization (NATO RTO), RTO-MP-54*, 2-1-2-5.

Argotti, Y., Davis, L., Outters, V., & Rolland, J.P. (2002). Dynamic superimposition of synthetic objects on rigid and simple-deformable objects. *Computers and Graphics, 26(6)*, 919-930.

Arsenault, R. & Ware, C. (2000). Eye-hand co-ordination with force feedback. *CHI Letters, 2,* 408-414.

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

Baillot, Y., Rolland, J.P., Lin, K., & Wright, D.L. (2000). Automatic modeling of knee-joint motion for the virtual reality dynamic anatomy (VRDA) tool. *Presence: Teleoperators and Virtual Environments (MIT Press), 9*(3), 223-235.

Beier, K. (2000). Virtual Reality Laboratory. University of Michigan [On-line]. Available: http://www-VRL.umich.edu

Biocca, F., Kim, J., & Choi, Y. (2001). Visual touch in virtual environments: An exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions. *Presence - Teleoperators and Virtual Environments, 10*, 247-265.

Birkfellner, W., Figl, M., Huber, K., Watzinger, F., Wanschitz, F., Hummel, J. et al. (2002). A head-mounted operating binocular for augmented reality visualization in medicine - Design and initial evaluation. *IEEE Transactions on Medical Imaging, 21*, 991-997.

Blade, R. & Padgett, L. (2002). Virtual environments standards and technology. In K.Stanney (Ed.), *Handbook of virtual environments - Design implementation and applications* (pp. 15-27). Mahwah, NJ: Lawrence Erlbaum.

Boff, K. R. & Lincoln, J. E. (1986). *Engineering data compendium: Human perception and performance*. Wright Patterson Air Force Base, OH: Harry G. Armstrong Aerospace Medical Research Laboratory.

Bowman, D. A. (1999). Interaction techniques for common tasks in immersive virtual environments: Design, evaluation, and application. Ph. D. Georgia Institute Of Technology.

Bowman, D. A. (2002). Principles for the design of performance-oriented interaction techniques. In K.Stanney (Ed.), *Handbook of virtual environments - Design implementation and applications* (pp. 277-300). Mahwah, NJ: Lawrence Erlbaum.

Burdea, G. & Coiffet, P. (1994). *Virtual reality technology*. New York, NY: John Wiley & Sons.

Cakmakci, O. & Rolland, J.P. (2007). Design and fabrication of a dual-element off-axis near-eye optical magnifier. *Optics Letters, (32)*11, 1363-1365.

Cakmakci, O., Vo, S., Thompson, K.P., & Rolland, J.P. (2008). Application of radial basis functions to shape description in a dual-element off-axis eyewear display: field of view Limit. *Journal of the Society of Information Display 16* (11), 1089-1098.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human computer interaction*. Hillsdale, NJ: Erlbaum Associates.

Clancy, T. (1999). Carrier: *A guided tour of an aircraft carrier*. New York, NY: Berkley Books.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences. (2nd ed.)* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cutting, J. E. & Vishton, P. M. (1995). Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion* (pp. 69-117). San Diego, CA: Academic Press.

DiZio, P. & Lackner, J. (2002). Proprioceptive adaptation and aftereffects. In K.Stanney (Ed.), *Handbook of virtual environments - design implementation and applications* (pp. 277-300). Mahwah, NJ: Lawrence Erlbaum.

Drexler, J., Kennedy, R., Malone, L. (2009). Virtual environment sickness and implications for training. In D. Nicholson, D. Schmorrow, & J. Cohn (Eds.), *Handbook of virtual environments for training and education Vol. 2*, (pp. 219-239). Westport, CT: Praeger.

Eberts, R. E. (1999). *User interface design*. Englewood Cliffs, NJ: Prentice Hall.

Ehrlich, J. (1999). The effect of viewing conditions on visual stress, sickness, and distance estimation in a helmet-mounted display. Ph. D. University of Central Florida.

Ellis, S. R., Bucher, U. J., & Menges, B. M. (1995). The relationship of binocular convergence and errors in judged distance to virtual objects. *Proc. of the International Federation of Automatic Control.*

Ellis, S. R. & Menges, B. M. (1997). Judgments of the distance to nearby virtual objects: Interaction of viewing conditions and accommodative demand. *Presence-Teleoperators and Virtual Environments, 6,* 452-460.

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

Fidopiastis C., Meyer, C., Fuhrman, C., & Rolland, J. (2004a). Methodology for the Interactive Evaluation of Prototype Head-mounted Displays in Virtual Environments: Visual Acuity Metrics. P*resence - Teleoperators and Virtual Environments, Special Issue on Immersive Technology*.

Fidopiastis, C., Meyer, C., Fuhrman, C., & Rolland, J. (2004b). Quantitative assessment of visual acuity in projective head mounted display. S*PIE 17th annual International Symposium on Aerospace / Defense Sensing Simulation and Controls,* Orlando, FL.

Gross, D. (2004). Affordances in the design of virtual environments. Ph. D. University of Central Florida.

Ha, Y., Rolland J.P., Davis Jr., L.D. (2006). Head mounted projection display with wide field of view. *US Patent 7,119,965*.

Hays, R., Vincenzi, D. S. A., & Bradley, S. (1998). *Training effectiveness evaluation of the VESUB technology demonstration system*. Naval Air Warfare Center Training Systems Division [On-line]. Available: http://www.ntsc.navy.mil

Hinckley, K., Pausch, R., Proffitt, D., & Kassell, N. F. (1998). Two-handed virtual manipulation. *ACM Transactions on Computer-Human Interaction, 5*, 260-302.

Hu, H., Gooch, A., Creem-Regehr, S., & Thompson, W. (2002). Visual Cues for perceiving distances from objects to surfaces. *Presence - Teleoperators and Virtual Environments, 11*, 652-664.

Hua, H., Gao, C., & Rolland, J. (2003). A testbed for precise registration, natural occlusion and interaction in an augmented environment using a head-mounted projective display (HMPD). *Proc. of the IEEE Virtual Reality Conference*.

Hua, H., Gao, C., Brown, L., Ahuja, N., & Rolland, J. (2002). Using a head-mounted projective display in interactive augmented environments. *Proc. of the IEEE Virtual Reality Conference*.

Hua, H., Girardot, A., Gao, C., & Rolland, J. P. (2000). Engineering of head-mounted projective displays. *Applied Optics, 39*, 3814-3824.

Hua, H., Gao, C., Biocca, F., & Rolland, J. P. (2001). An ultra-light and compact design and implementation of head-mounted projective displays. *Proc. of the IEEE Virtual Reality Conference*, 175-182.

Inami, M., Kawakami, N., Sekiguchi, D., Yanagida, Y., Maeda, T., & Tachi, S. (2000). Visuo-haptic display using head-mounted projector. *Proc. of the IEEE Virtual Reality Conference*.

Kakeya, H., Isogai, M., Suzuki, K., & Arakawa, Y. (1999). Autostereoscopic display with real-image virtual screen. *SIGGRAPH*.

Kato, H., Billinghurst, M., Weghorst, S., & Furness, T. (1999). *A mixed reality 3D conferencing application*. University of Washington [On-line]. Available: http://www.hitl.washington.edu

Kawara, T., Ohmi, M., & Yoshizawa, T. (1996). Effects on visual functions during tasks of object handling in virtual environment with a head mounted display. *Ergonomics, 39*, 1370-1380.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire (SSQ): An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3(3).* 203-220.

Kennedy, R. S., Dunlap, W. P., Jones, M. ., & Stanney, K. M. (1996). *Screening users of virtual reality systems for after-effects such as motion sickness and balance problems* (Final Rep. No. NSF1-96-4). Arlington, VA: National Science Foundation.

Kirkley (2003). *Augmented reality performance assessment battery: Object recognition, distance estimation using optical see-through head-worn displays*. Ph. D. Indiana University.

Kitamura, Y., Tomohiko, H., Toshihiro, M., & Fumio, K. (1999). Virtual chopsticks: object manipulation using multiple exact interactions. *Proc. of the IEEE Virtual Reality Conference.*

Kruger, W., Bohn, C. A., Frohlich, B., Schuth, H., Strauss, W., & Wesche, G. (1995). The responsive workbench - A virtual work-environment. *Computer, 28*, 42-48.

Lampton, D., Knerr, B. W., Goldberg, S. L., Bliss, J. P., Moshell, J. M., & Blau, B. S. (1994). The virtual environment performance assessment battery. *Presence - Teleoperators and Virtual Environments, 3*, 145-157.

Latham, R. (1998). *Robotic device hands real switches and buttons to users of virtual worlds*. http://www.cgsd.com/TOPIT_PR.html [On-line].

Lehner, V. & DeFanti, T. (1997). Distributed virtual reality: supporting remote collaboration in vehicle design. I*EEE Computer Graphics and Applications.*

Leibe, B., Starner, T., Ribarski, W., Wartell, Z., Krum, D., Singletary, B. et al. (2000). The perceptive workbench: Toward spontaneous and natural interaction in semi-immersive virtual environments. *Proc. of the IEEE Virtual Reality Conference*, 13-20.

Lewis, J. (2004, May). In the eye of the beholder. *IEEE Spectrum*.

Lindeman, R. (1999). Bimanual interaction, passive-haptic feedback, 3D widget Representation, and simulated surface constraints for interaction in immersive virtual environments. Ph. D. George Washington University.

Liu, A., Stark, L., & Hirose, M. (1991). Interaction of visual depth cues and viewing parameters during simulated telemanipulation. *Proc. of IEEE International Conference on Robotics and Automation*, 2286-2291.

Lok, B. (2002). Interacting with dynamic real objects in virtual environments. Ph. D. UNC-CH.

Lok, B., Naik, S., Whitton, M., & Brooks, F. P. (2003). Effects of handling real objects and avatar fidelity on cognitive task performance in virtual environments. *Proc. of the IEEE Virtual Reality Conference,* 125-132.

Martins, R., Shaoulov, V., Ha, Y., & Rolland, J.P. (2007). A mobile head-worn projection display. *Optics Express 15*, 14530-14538.

Mendenhall, W. & Sincich, T. (1994). *Statistics for engineering and the sciences*, Upper Saddle River, NJ: Prentice Hall.

Milgram, P. & Colquhoun, H. (1999). A taxonomy of real and virtual world display integration. In Y.Ohta & H. Tamura (Eds.), *Mixed reality - Merging real and virtual worlds* (pp. 5-30). Tokyo, Japan: Ohmsha.

Milgram, P. & Kishino, F. (1994). A taxonomy of mixed reality visual-displays. *IEEE Transactions on Information and Systems, E77D*, 1321-1329.

Mon-Williams, M., Plooy, A., Burgess-Limerick, R., & Wann, J. (1998). Gaze angle: a possible mechanism of visual stress in virtual reality headsets. *Ergonomics, 41*, 280-285.

Mon-Williams, M., Tresilian, J. R., Strang, N. C., Kochhar, P., & Wann, J. P. (1998). Improving vision: neural compensation for optical defocus. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 265, 71-77.

Mon-Williams, M. & Wann, J. P. (1998). Binocular virtual reality displays: When problems do and don't occur. *Human Factors, 40*, 42-49.

MonWilliams, M., Wann, J. P., & Rushton, S. (1993). Binocular vision in a virtual world - visual deficits following the wearing of a head-mounted display. *Ophthalmic and Physiological Optics, 13*, 387-391.

Munro, A., Breaux, R., Patrey, J., & Sheldon, B. (2002). Cognitive aspects of virtual environments design. In K.Stanney (Ed.), *Handbook of virtual environments -*

design implementation and applications (pp. 415-434). Mahwah, NJ: Lawrence Erlbaum.

Murali, S., Lee, K.S., & Rolland, J.P. (2007). Invariant resolution dynamic focus OCM based on liquid crystal lens. *Optics Express 15*(24), 15854-15862 November (2007).

Murali, S., Thompson, K.P., and Rolland, J.P. (2009). Three-dimensional adaptive microscopy using embedded liquid lens. *Optics Letters, 34*(2), 145–147 (2009).

NASA (1981). NASA technical memorandum 8127B: Characteristics of flight simulator: visual systems. Moffett Field, CA: NASA.

NATO RTO (2003). Report of the Research and Technology Organization (RTO) Human Factors and Medicine (HFM) panel, technical report 18, virtual reality: state of military research and applications in member countries Neuilly-sur-Seine Cedex, France: NATO RTO.

Popescu, G., Burdea, G., & Trefftz, H. (2002). Multimodal interaction modeling. In K.Stanney (Ed.), *Handbook of virtual environments - Design implementation and applications* (pp. 435-454). Mahwah, NJ: Lawrence Erlbaum.

Robinett, W. & Rolland, J. (1994). A computational model for the stereoscopic optics of a head-mounted display. In R. Earnshaw, M. Gigante, & H. Jones (Eds.), *Virtual reality systems* (pp. 51-76). San Diego, CA: Academic Press.

Rolland, J.P. (2000). Wide angle, off-axis, see-through head-mounted display. *Optical Engineering - Special Issue on Pushing the Envelop in Optical Design Software, 39* (7), 1760-1767.

Rolland, J.P., Biocca, F., Hamza-Lup, F., Ha, Y., & Martins, R. (2005). Development of head-mounted projection displays for distributed, collaborative augmented reality applications. *Presence: SI Immersive Projection Technology, 14*(5), 528-549 (2005).

Rolland J.P., Ha, Y. & Fidopiastis, C. (2004). Albertian errors in head-mounted displays: choice of eye points location for a near or far field task visualization", *J. Opt. Soc. Am. A, 21*(6), 901-912.

Rolland, J. & Fuchs, H. (2000). Optical versus video see-through head-mounted displays in medical visualization. *Presence - Teleoperators and Virtual Environments, 9*, 287-309.

Rolland, J., Meyer, C., Davis, L., Hamza-Lup, F, & Norfleet, J. (2002a). Merging augmented reality and anatomically correct 3D models in the development of a training tool for endotracheal intubation. *Proc. of the IEEE International Symposium on Biomedical Imaging*, 895-898.

Rolland, J., Parsons, J., Poizat, D., & Hancock, D. (1998). Conformal optics for 3D visualization. *SPIE, 3482*, 760-764.

Rolland, J. P., Gibson, W., & Ariely, D. (1995). Towards quantifying depth and size perception in virtual environments. *Presence - Teleoperators and Virtual Environments, 4*, 24-49.

Rolland, J. P., Krueger, M. W., & Goon, A. (2000). Multifocal planes head-mounted displays. *Applied Optics, 39*, 3209-3215.

Rolland, J. P., Meyer, C., Arthur, K., & Rinalducci, E. (2002b). Method of adjustments versus method of constant stimuli in the quantification of accuracy and precision of rendered depth in head-mounted displays. *Presence-Teleoperators and Virtual Environments, 11*, 610-625.

Rolland, J. P., Wright, D. L., & Kancherla, A. R. (1997). Towards a novel augmented-reality tool to visualize dynamic 3-D anatomy. *Stud. Health Technol. Inform., 39*, 337-348.

Rolland, J.P., A. Yoshida, A., Davis, L., & Reif, J. H. (1998). High resolution inset head-mounted display. *Applied Optics, 37*(19), 4183-4193.

Rushton, S., MonWilliams, M., & Wann, J. P. (1994). Binocular vision in a Bi-Ocular world - New-generation head-mounted displays avoid causing visual deficit. *Displays, 15*, 255-260.

Santhanam, A., Willoughby, T., Kaya, I., Shah, A., Meeks, S.L., Rolland, J.P., & Kupelian, P. (2008). A display framework for visualizing real-time 3D lung tumor dosimetry. *IEEE Journal of Display Technology, Special issue on Medical Displays, 4*(4): p. 473-482.

Schaffer, R., Cullen, S., Cohn, J., & Stanney, K. M. (2003). A personal LCAC simulator supporting a hierarchy of training requirements. *Proc. of Interservice/Industry Training, Simulation and Education Conference.*

Schiefele, J. (2000). Realization and evaluation of virtual cockpit simulation and virtual flight simulation. Thesis, Technical University Darmstadt.

Shibata, T. (2002). Head mounted display. *Displays, 23*, 57-64.

Shilling, R. & Shinn-Cunningham, B. (2002). Virtual auditory display. In K.Stanney (Ed.), *Handbook of virtual environments - design implementation and applications* (pp. 65-92). Mahwah, NJ: Lawrence Erlbaum.

Shneiderman, B. (1998). *Designing the user interface (3rd Edition)*. Reading, MA: Addison Wesley Longman, Inc.

Shneiderman, B. & Maes, P. (1997). Direct manipulation vs. interface agents. *Interactions, 4*, 42-61.

Southard, D. (1997). Designing HMD systems for stereoscopic vision. In J. Melterz & K. Moffitt (Eds.), *Head mounted displays - Designing for the user* (pp. 253-283). New York, NY: McGraw-Hill.

Stoffregen, T. A., Draper, M. H., Kennedy R.S., & Compton, D. (2002). Vestibular adaptation and aftereffects. In K.Stanney (Ed.), *Handbook of virtual environments - design implementation and applications* (Mahwah, NJ: Lawrence Erlbaum.

Surdick, R. T., Davis, E. T., King, R. A., & Hodges, L. F. (1997). The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. *Presence-Teleoperators and Virtual Environments, 6*, 513-531.

Urey, H., Nestorovic, N., Ng, B., & Gross, A. (1999). Optics designs and system MTF for laser scanning displays. *SPIE, 3689*, 238-248.

Valyus, N. (1966). *Stereoscopy*. London: Focal Press.

van Koesveld, J. J. M., Tetteroo, G. W. M., & de Graaf, E. J. R. (2003). Use of head-mounted display in transanal endoscopic microsurgery. *Surgical Endoscopy and Other Interventional Techniques, 17*, 943-946.

Viirre, E., Pryor, H., & Nagata, S. F. T. (1998). The virtual retinal display: A new technology for virtual reality and augmented vision in medicine. *Proc. of Medicine Meets Virtual Reality*.

Vince, J. (1995). *Virtual reality systems*. New York, NY: ACM Press.

von Wiegand, T. E., Schloerb, D. W., & Sachtler, W. L. (1999). Virtual workbench: Near-field virtual environment system with applications. *Presence-Teleoperators and Virtual Environments, 8*, 492-519.

Wagner, A., Ploder, O., Enislidis, G., & Truppe, M. (1996). Image guided surgery. *International Journal of Oral and Maxillofacial Surgery, 25*, 147-151.

Wagner, A., Rasse, M., Millesi, W., & Ewers, R. (1997). Virtual reality for orthognathic surgery: The augmented reality environment concept. *Journal of Oral and Maxillofacial Surgery, 55*, 456-462.

Wann, J. & Mon-Williams, M. (2002). Measurement of visual aftereffects following virtual environment exposure. In K. Stanney (Ed.), *Handbook of virtual environments - design implementation and applications* (pp. 731-750). Mahwah, NJ: Lawrence Erlbaum.

Wann, J. P., Rushton, S., & Mon-Williams, M. (1995). Natural problems for stereoscopic depth-perception in virtual environments. *Vision Research, 35*, 2731-2736.

Wanschitz, F., Birkfellner, W., Figl, M., Patruta, S., Wagner, A., Watzinger, F. et al. (2002). Computer-enhanced stereoscopic vision in a head-mounted display for oral implant surgery. *Clinical Oral Implants Research, 13*, 610-616.

Wendt, M., Sauer, F., Khamene, A., Bascle, B., Vogt, S., & Wacker, F. K. (2003). A head-mounted display system for augmented reality: Initial evaluation for interventional MRI. Rofo-Fortschritte Auf dem Gebiet der Rontgenstrahlen und der Bildgebenden Verfahren, 175, 418-421.

White, M. A., Arena, C., Newton, R., & Hopper, J. (2003). VIRTE: A training system for the future...today! *Proc. of Interservice/Industry Training, Simulation and Education Conference*.

White, W. E., Wharton, L., Kotick, D., & Anschuetz, E. (2003). PC-based microSims in a distributed military simulation environment. *Proc. of Interservice/Industry Training, Simulation and Education Conference*.

Wright, D. L., Rolland, J. P., & Kancherla, A. R. (1995). Using virtual reality to teach radiographic positioning. *Radiol.Technol., 66*, 233-238.

Yeo, T. T., Kockro, R. A., Serra, L., Chua, G. G., Ng, H., Lee, E. et al. (1999). The KRDL virtual workbench: A virtual reality environment for visualisation, planning and simulation of epilepsy surgery. *Epilepsia, 40*, 52.