User-Centered Virtual Environment Assessment and Design for Cognitive Rehabilitation Applications

by

CALI MICHAEL FIDOPIASTIS
B.S. Biology, University of California, Irvine 1994
B.A. Psychology, University of California, Irvine 1994
M.A. Experimental Psychology (Sociology), University of California, Irvine 1998

ABSTRACT

Virtual environment (VE) design for cognitive rehabilitation necessitates a new methodology to ensure the validity of the resulting rehabilitation assessment. We propose that benchmarking the VE system technology utilizing a user-centered approach should precede the VE construction. Further, user performance baselines should be measured throughout testing as a control for adaptive effects that may confound the metrics chosen to evaluate the rehabilitation treatment. To support these claims we present data obtained from two modules of a user-centered head-mounted display (HMD) assessment battery, specifically resolution visual acuity and stereoacuity.

Resolution visual acuity and stereoacuity assessments provide information about the image quality achieved by an HMD based upon its unique system parameters. When applying a user-centered approach, we were able to quantify limitations in the VE system components (e.g., low microdisplay resolution) and separately point to user characteristics (e.g., changes in dark focus) that may introduce error in the evaluation of VE based rehabilitation protocols.

Based on these results, we provide guidelines for calibrating and benchmarking HMDs. In addition, we discuss potential extensions of the assessment to address higher level usability issues. We intend to test the proposed framework within the Human Experience Modeler (HEM), a testbed created at the University of Central Florida to evaluate technologies that may enhance cognitive rehabilitation effectiveness. Preliminary results of a feasibility pilot study conducted with a memory impaired participant showed that the HEM provides the control and repeatability needed to conduct such technology comparisons. Further, the HEM affords the opportunity to integrate new brain imaging technologies (i.e., functional Near Infrared Imaging) to evaluate brain plasticity associated with VE based cognitive rehabilitation.

This work is dedicated to η προ-γιαγιά μού Cali Rigas whose sage-like understanding of society and its illusory boundaries moved her beyond the stereotypes of her generation and allowed her to steer the course of her life and that of our family's without apology of her gender. I am indebted to her for her resolve to succeed against all odds while remaining compassionate. In your memory, Great Grandmother, I am deeply honored to hold your name.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

xiii

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

|CE| Absolute Value of the Constant Error

2AFC   Two-Alternative Forced Choice

3D   Three-Dimensional

ADL   Activities of Daily Living

AE   Absolute Error

ANOVA   Analysis of Variance

API   Application Programming Interface

ARC   Augmented Reality Center

CAVES   Cave Automatic Virtual Environment

CBDI   Cognitive Behavioral Driver's Inventory

CE   Constant Error

DOF   Depth of Focus

FBO   Frame Buffer Overscan

fMRI   Functional Magnetic Resonance Imaging

FOV   Field of View

HCI   Human-Computer Interaction

HEM   Human Experience Modeler

H-D   Howard-Dolman

HMPD   Head-Mounted Projection Display

HMPD-1   ODALab first generation HMPD

HMPD-2	ODALab second generation HMPD
HWD	Head Worn Display
IPD	Interpupillary Eye Distance
IRB	Institutional Review Board
JND	Just Noticeable Difference
L	Luminance
LASIK	Laser-Assisted *In Situ* Keratomileusis
LCD	Liquid Crystal Display
LGN	Lateral Geniculate Nucleus
MR	Mixed Reality
MRSS	Mixed Reality Software Suite
MTF	Modulating Transfer Function
n	Sample Size
η_e	Effect Size of the ANOVA
η_{ce}	Stereoaccuracy
η_{ve}	Stereoacuity
ODALab	Optics Diagnostics and Applications Lab
OLED	Organic Light Emitting Diode
PDA	Personal Digital Assistant
PEST	Parameter Estimation by Sequential Testing
PROBIT	Probability Unit
PSE	Point of Subjective Equality
QUEST	Quick Estimation by Sequential Testing

RMSE	Root Mean Square Error
rVA	Resolution Visual Acuity
SSQ	Simulator Sickness Questionnaire
TBI	Total Brain Injury
TIR	Total Internal Reflection
TMS	Transcranial Magnetic Stimulation
VA	Visual Acuity
VEPAB	Virtual Environment Assessment Battery
V_r	Variable Error
VGA	Video Graphics Array
V-HD	Virtual Howard-Dolman

CHAPTER 1 INTRODUCTION

Traumatic brain injury (TBI) is an insult to the brain due to external forces (e.g., car accident), internal damage (e.g., gunshot wound), or loss of oxygen (e.g., embolism). The annual incidence of TBI in the United States is 1.5 million Americans per year, with males in the age range of 15 to 24 years at the highest risk. Vehicular accidents are responsible for approximately 44% of TBIs annually (CDC, 2001; Sosin, Sniezek., & Thurman, 1996). Currently, over 5 million Americans live with TBI related disabilities (e.g., cognitive, motor, and affective). Cognitive impairments such as memory, reasoning, judgment, self-awareness prevent most persons with TBI from maintaining functional or financial independence, which places a strain on families as well as healthcare resources (Max, Mackenzie, & Rice, 1991; Rose, Brooks & Rizzo, 2005).

Laws requiring seatbelts and helmets reduced the incidence of TBI over the past few years (National Center for Injury Prevention and Control, 2001; Rutledge & Stutts, 1993). However, the medical rehabilitation field, specifically areas of Clinical Neuropsychology and Cognitive Rehabilitation, has not successfully provided reliable protocols for recovery of cognitive function after a TBI, thus placing the necessity for TBI rehabilitation into question (Ricker, 1998; Carney et. al, 1999). The negative issues surrounding cognitive rehabilitation stem from the expense involved in maintaining rehabilitation teams, which may include a neuropsychologist, an occupational therapist, a physical therapist, and a speech pathologist, as well as the lack of evidence as to the efficacy of treatment plans. More specifically, the cost associated with cognitive retraining could be somewhat justified if transfer of training from the

rehabilitation facility to the patient's home were demonstrated. According to Rizzo and Buckwalter (1997), the advent of virtual environments (VEs) provides a solution for more cost effective and ecologically valid cognitive rehabilitation procedures. However, can VE overcome the historical failures of computer assisted cognitive retraining?

1.1 Cognitive Rehabilitation and Computer-Assisted Cognitive Retraining

It is important to note that cognitive rehabilitation as a coherent field of medicine was not established in the United States until the early 1970's (Boake, 1991). At this time, information processing models based on computers suggested that the brain functioned similar to hardware, software, and data input and output devices. In 1968, Atkinson and Shiffrin suggested that memory consisted of two distinct memory stores, short-term and long-term. In contrast, Craik and Lockhart (1972) posed a "levels of processing" approach to memory where persons engaging in deep, meaningful, information processing experienced more permanent retention of memory items than shallow, sensory kinds of processing. Structure versus processing debates, as well as debates on localization of brain function (modularity) versus holism (distributed processing), continued to qualify the differences between restorative and functional cognitive rehabilitation methodologies (Levin, 1991).

The premise of the restorative approach is that there are discrete brain processes that rehabilitation professionals can identify through neuropsychological testing. Further, these cognitive processes are trainable separately through drills and practice. The information processing metaphor suggests that processing occurs in a hierarchical manner, thus therapists should retrain cognitive functions such as attention before introducing training in daily living

skills. Using the restorative approach, a person with TBI would receive systematic and repetitive

retraining of cognitive processes, and then integrate this learning with motor tasks. Eventually

the patient would receive problem-solving tasks that matched more real life scenarios. The video

game and personal computer industries changed how clinicians administered cognitive retraining

modules and provided a solution to the high cost of TBI rehabilitation (Lynch, 2002).

Presenting training tasks on the computer allowed for more stimulus control, application

consistency, and simpler data collection methods. A person with TBI could sit at a computer

terminal and perform his or her exercises without the assistance of a therapist, potentially cutting

costs. However, the computer software languages available at that time made programming

custom applications almost impossible. Further, home computers were a luxury item not found in

many patients' homes, therefore outpatients continued to receive therapy within rehabilitation

facilities.

In 1982, Lynch evaluated the effectiveness of common Atari video games (i.e., Pac-Man)

for cognitive rehabilitation. Atari introduced their first personal game console in 1977 at the cost

of approximately 200 US dollars (Videotopia, 1998). Video games were a cost effective means

of retraining attention and concentration abilities in the home as well as the therapy center.

Empirical evidence suggested that computer-assisted therapy was effective in retraining some

basic cognitive functions such as directed and divided attention (Lynch, 2002). Unfortunately,

there is a lack of evidence supporting that persons with TBI undergoing computer-assisted

therapy actually transfer their skill to other tasks (Chen, Thomas, Glueckauf, & Bracy, 1997).

Lack of skill transfer from the therapy environment to every day tasks is a general criticism of

the restorative approach (Chase & Eriksson, 1981; O'Conner & Cermack, 1987). Based upon

these results, restorative therapists abandoned or limited their use of computer-assisted retraining

3

applications. The compensatory or functional approach to rehabilitation has shown more success as a rehabilitation paradigm; however, it too has limitations (Sohlberg and Mateer, 2001).

The functional approach to rehabilitation focuses on retraining of activities of daily living (ADL's) such as meal preparation or dressing. The focus is not on changing brain function, but changing behavioral strategies for performing daily tasks. For example, a person experiencing memory problems may use a personal digital assistant (PDA) to store important daily events. Although there is increasing support for successful retraining under the functional approach (Wilson, 2000), opponents contend that the scripted nature of daily living task precludes the flexibility to generalize to new or even similar tasks or contexts (Kirsch, Levine, Lajiness-O'neil, Schnyder, 1992).

Ylvisaker, Hanks, & Johnson-Greene (2002) proposed a more combined approach to cognitive rehabilitation, one based on personalized context and familiarity. The basic assumption of their context-based approach is that persons maintain a domain specific knowledge base within their brain that assists with integrating new information or recalling old. Domain specificity is a more general form of modularity in that different neural areas oversee specific cognitive functions; however, processing is not "encapsulated" or separated from other brain systems. Thus, both declarative and procedural domain-specific knowledge are stored and are available when experiencing the appropriate context.

One could interpret the contextualized approach as suggesting that restorative therapy should only use personalized or familiar stimuli. As well, functional training should only include contexts relating to the individual's own life experience. In actuality, context theory suggests that the distinction between restorative and functional approaches is nonexistent since the functional approach is embedded within that of the restorative. Schultheis & Rizzo (2001) posed a similar

4

argument and further posited that virtual reality technology offers a means to explore transfer of training issues systematically and under more ecologically valid or personally contextualized training conditions. Previous work with driving simulators and persons with TBI showed that computer simulated environments provided an effective testbed for assessing cognitive functioning (Engum, Cron, Hulse, Pendergrass, & Lambert, 1988; Lambert & Engum, 1992).

1.2 Simulated and Virtual Environments as Rehabilitation Tools

Simulation refers to the substitution of real objects or tasks with a model or system representation. Historically, simulation-based training is substituted for real world procedures when the cost of training (e.g., large-scale war games) is exceedingly high or when the learning task may pose dangers to the trainee or others (e.g., surgical training). This training paradigm permits errors and other natural behaviors (e.g., exploration) that occur during the learning processes without risk of harm. A successful application of simulation-based training is the CBDI or Cognitive Behavioral Driver's Inventory (Engum, Cron, Hulse, Pendergrass, & Lambert, 1988).

CBDI utilizes a mixed rehabilitation approach in that TBI participants perform a battery of driving related computerized tests, paper and pencil neuropsychological tests, vision and motoric tasks within a driving simulator. Scores for the test are pass or fail; however, the simulator records errors so that the participant can review his or her performance. The "after action review" whereby participants review their performance is especially important since many person with TBI do not recognize their own impairment. CBDI research suggests that a previous failure of the earlier computer-based training approaches was mainly due to lack of a real world

context (Gourlay, Lun, Lee, & Tay, 2000). These positive aspects of using simulated real world tasks extend to the more interactive virtual rehabilitation environments.

1.2.1 Advantages of VE Training

Burdea & Coiffet (2004) define the virtual environment (VE) as a high-end user computer-interface that involves real-time simulation and interaction through multiple sensory channels (e.g., vision, auditory, tactile, smell, and taste). The appeals of creating virtual environments for cognitive rehabilitation are many. Some positive aspects of VE include the ability to create a personalized environment that matches the disability level of the participant (Wilson, Foreman, & Stanton, 1997), to control experimental design factors, to improve realism through appropriate context, to provide quantifiable feedback for increased safety and self awareness (Schultheis & Rizzo, 2001), and to increase transfer of training (Rose et al., 2000). Further, immersion into virtual environments has the potential to increase a person's engagement in his or her training task, thus potentially facilitating transfer of training transfer (Rizzo and Buckwalter, 1997). VEs also allow researchers to explore brain functioning using brain imaging techniques while participants engage in simulated real world tasks such as driving or navigating through their own home (Baumann et al., 2003).

Current research using brain-imaging techniques (e.g., Transcranial Magnetic Stimulation, TMS) and functional Magnetic Resonance Imaging (fMRI)) supports the occurrence of brain changes after successful stroke motor therapy (Liepert et al., 2000). In addition, studies with brain injured lab animals suggest that exposure to an enriched environment, one with mentally stimulating objects and activities, enhances neurogenesis, the regeneration of brain

neurons (Johansson, 2003). Thus, VE training coupled with brain imaging techniques may augment our understanding of successful transfer of training from a brain reorganization perspective (Rose, Attree, Brooks & Johnson, 1998; Rose, Brooks & Rizzo, 2005).

1.2.2 Transfer of Training from VE to Real World

Success of VE as a rehabilitation-training tool ultimately hinges on the empirical demonstration of general and specific transfer along with generalization of learning to one's home and possibly work environments. Two classic studies of transfer of training using virtual environments both explore motoric skill transfer with "normal" populations. Rose et al. (2000) trained participants on a virtual or real world "steadiness tester", a simple sensorimotor task that trained the participant to move a ring along the path of a nonlinear wire without touching it. The researchers compared VE performance to that of the real world and found that persons who trained in the VE performed similarly or better than those persons trained in the real world. Todorov, Shadmer and Bizzi (1997) demonstrated that a VE trainer displaying only context relevant information (i.e., end effector information) improved performance in table tennis compared to instruction with a professional table tennis coach. Persons with learning disabilities also show positive transfer of training to real world tasks using VE.

Brown, Standen, & Cobb (1998) found that children with learning disabilities performed better in a supermarket after receiving training in a virtual marketplace. Brooks et al. (1999) showed that an amnesic patient, who had been unsuccessful in navigating routes in a hospital unit, learned two routes after VE training. The findings are significant since the participant exhibited new learning although she was unaware that she had learned the routes. An

encouraging explanation is that VE training may have tapped procedural memory processes. Further, post neuropsychological assessment showed "mild improvement on tests of general intellectual abilities, visual perceptual and naming skills but memory and executive functions remained severely impaired" (p.73). Thus, this study as well as others does not explicitly demonstrate transfer of learning of specific cognitive skills from a virtual to a real environment.

Further work is necessary to quantify the degree of similarity between tasks performed in the VE as opposed to those performed in the real world. As well, determining the perceptual or precise conditions that are most relevant to generalization is necessary. More importantly, VE tasks representing ADLs must show a relationship to performance within the home environment. To date there have been no follow-up studies to provide evidence for sustained transfer of training to the home environment.

1.3 Motivation: VE System Design Confounds and System Benchmarks

Although it seems as if we are close to proving the efficacy of VE as a training tool for persons with TBI, one must bear in mind that VE technology is also a basic research tool (Tarr & Warren, 2002). As such, it is imperative that basic researchers ensure that the VE technology provides reliable and valid measures of user performance prior to empirical testing. There are several technological and computer graphics issues that lead to degraded perception in VEs that may confound VE rehabilitation assessments (Fidopiastis, Fuhrman, Meyer, & Rolland, 2004). For example, micro displays within the head-mounted display (HMD) typically limit the user's visual resolution acuity. Further, HMD optical systems with a single imaging plane in many instances degrade the user's ability to utilize depth cue information such as natural

accommodation and convergence mechanisms associated to the human eye (Wann, Rushton, Mon-Williams, 1995). The resultant visual perceptual errors have the potential to distort experimental results, including those obtained from brain imaging.

Stanney (1995) makes the point that task characteristics will affect user outcomes in VE applications. Within the context of human and machine interaction, tasks are subjective units of work that can change the user's performance based on actions needed to initiate, change, or stop system specific states (Christensen, 1993). Task characteristics also determine the level of realism, immersion, and overall detail necessary to achieve the desired training effect. For example, pain management and exposure therapies using VE technologies do not demand high levels of graphical detail or realism to achieve positive behavioral outcomes (Rothbaum & Hodges, 1999; Hoffman, Patterson, Carrougher, & Sharar, 2001). However, rehabilitation protocols for assisting persons with spatial neglect should ensure that distortions of shape and size constancy, as well as depth errors caused by the VE technology are controlled or at least understood prior to therapy and experimentation. Thus, how does one design a confound-free virtual rehabilitation environment utilizing current Human Factors assessment protocols?

Over the past 10 years researchers within the Human-Computer Interaction (HCI) community have proposed varying VE systems design approaches including those that focus on perceptual issues (Wann & Mon-Williams, 1996; Fencott, 1999), usability defined as ease of use and utility (Kaur, Maiden, & Sutcliffe, 1999; Bowman, Gabbard, & Hix, 2002), or combined perceptual and usability models (Stanney, Mourant, Kennedy, 1998; Stanney et al., 2003). Stanney (1995) contended that human sensory and motor physiology may prove a limiting factor in some aspects of VE design. These limitations are accentuated by the sensory and motoric impairments of some person's with TBI. Thus, within cognitive rehabilitation VE applications,

9

user characteristics must be separated from technology limitations (e.g., low display resolution) to ensure meaningful empirical results. Separating the human component from the engineering component may prove difficult (Kalawsky, 1993, p. 43). Melzer and Moffit (1997) tried to address this issue by applying similar user-centered methodologies to head-worn display (HWD) design cycles. As more technology providers (e.g., haptic devices) follow, convergence between the engineering and user requirements will simplify the VE design decision making. However, benchmarks addressing user sensory capabilities across separated and integrated technologies are necessary when evaluating performance of person with TBI. These benchmarks are critical to evaluating negative adaptation effects as well as positive transfer of training.

Currently, usability tests lack appropriate sensory tests (vision, auditory, smell, and touch) to provide accurate benchmarks for VE systems (technology, computer graphics, and user). In this work, we propose a vision test battery that quantifies key components of the human image processing system (e.g., resolution visual acuity and depth perception). The results of such a battery should provide basic and applied vision parameters for the total VE system, which allow for appropriate benchmarks and performance evaluations that control for visual errors (e.g., distorted depth) within cognitive rehabilitation applications. Additionally, the methodology outlined for test design could be extended to other modalities, empowering rehabilitation therapists with a more reliable training tool capable of discerning positive retraining from VE system artifacts.

1.4 Research Summary

The overall thesis of this research is that user-based sensory system testing is necessary for benchmarking VE system performance, especially for applications designed for use with TBI populations. For example, egocentric and exocentric distance judgments are both under and overestimated in virtual environments (VE) compared to the real world. How would this depth error affect the performance of stroke patients exhibiting hemineglect or failure to orient to one side of space (Myers & Bierig, 2000)? Would there be negative adaptation effects as a result of the VE exposure? Without norms and baselines for visual performance within the VE system, the potential for confounds in the results obtained from the VE rehabilitation application are heightened.

Researchers have reported three main factors that contribute to distance judgment errors: head mounted displays (Rolland, Gibson, & Ariely, 1995; Rolland & Fuchs, 2001; Witmer & Sadowski, 1998; Willemsen & Gooch, 2002), computer graphics (Robinett & Rolland, 1992; Hu et al., 2000; Loomis & Knapp, 2003), and human perceptual issues (Rinalducci, 1996; Surdick, Davis, King, & Hodges 1997,). Given these multiple points within the VE system, how does one pinpoint weaknesses in the system setup? Technology and software are now available from the ODALab at the University of Central Florida to create a VE test environment capable of quantifying both near and far field vision errors using a VE systems approach. The resulting vision test battery is a cross platform assessment tool that can determine weaknesses in the VE system design and pinpoint technology that may potentially degrade user performance.

In this work we evaluate VE system performance utilizing two modules of the vision test battery, resolution visual acuity and stereoacuity. We hypothesize that user specified visual

perceptual issues (i.e., dark convergence) may influence visual performance and may contribute to the individual variation seen in previous literature. Additionally, the effects of using a see-through helmet-mounted display for simulation in near space on distance judgments will also be evaluated. One capability of the newest HMD is an adjustable image depth plane of .8, 1.5, and 3 meters. This feature will allow us to study the effects of accommodation and convergence decoupling on depth perception without the constraints of collimated displays. The ultimate goal of such endeavor is to enhance transfer of training from simulation to real world tasks. More specifically, this work lays the foundation for advancing VE rehabilitation protocols for persons with TBI.

Although giving insight to wayfinding and complex tasks of navigation, this work does not specifically address this research domain. Further, although we limit this discussion to completed modules of the test battery, the full battery is outlined. We propose an extension of this work to mixed reality designs within a cognitive rehabilitation scenario in the future works section.

1.5 Dissertation Outline

The rest of the dissertation is organized in the following manner: Chapter 2 provides an overview of previous work in human factors and VE system design, focusing upon human performance issues. The summaries present assumptions, methodology, and outcomes of pertinent work in the field. We extend these design criteria to VE based cognitive rehabilitation applications.

Chapter 3 describes background and concepts related to VE system hardware and software that affect the user's visual perception. Within this chapter an overview of pertinent

features of the human visual system are discussed. As well, VE systems are further qualified by HMD and Computer Graphics related properties. For HMDs, optical imaging will be explained in addition to the importance of eye point location when evaluating system performance (Rolland, Ha, & Fidopiastis, 2004).

Chapter 4 gives an overview of the design of the vision test battery by first describing current vision tests. Noted factors affecting vision (e.g., lighting and dark focus) are discussed in order to provide a context and a foundation for the experiments described in Chapters 5, 6, 7, and 8.

Chapter 5 describes a calibration procedure for the see-through head mounted display used for the experiments outlined in Chapter 7.

Chapter 6 describes the experimental rationale, methods, experimental design, results, and conclusions for an experiment addressing the several questions pertaining to the design of a vision test that assesses resolution visual acuity. In general the main questions asked are: 1) How sensitive and reliable is the vision test at determining the resolution visual acuity of the VE system and 2) do factors such as contrast and other environmental lighting conditions effect the participants resolution visual acuity in the VE.

Chapter 7 describes the experimental rationale, methods, experimental design, results, and conclusions for an experiment addressing the two questions: 1) how adaptable is the test battery to use with different HMDs and 2) does dark focus accommodation affect a VE participant's perception of depth?

Chapter 8 builds on the results of the dark focus experiment adding dark convergence as a predictive metric and utilizing these metrics as potential control parameters in the study of decoupling accommodation and convergence across three image planes (.8, 1.5, and 3 meters).

The three image planes also correspond to the three viewing distances, expressed in millimeters, used in the experiment. Hypotheses, methods, experimental design, results, and conclusions for this experiment are given.

Chapter 9 describes a unique approach to designing mixed reality environments for testing transfer of training with person's with retrograde amnesia. The study also illustrates the need for vision assessment prior to VE retraining for persons with multiple impairments.

Chapter 10 summarizes the results and contribution of the work to the field. In addition, we discuss how this work is extended to mixed reality designs where real and virtual objects are collocated within the VE. Finally, future work utilizing a see-through head mounted display within the mixed reality rehabilitation environment is discussed.

CHAPTER 2 VE SYSTEM DESIGN METHOD APPLIED TO COGNITIVE REHABILITATION

In this section, we provide an overview of previous research in the field of Human-Computer Interaction as it pertains to VE system design. We attempt to extend this work to support cognitive rehabilitation VE designs that require performance baselines. We discuss the approaches, measurements, and guidelines of the chosen design assessments, providing support for the proposed work outlined in this thesis.

2.1 VE System Design: Overview

As described by the International Standard ISO13407, the Human-Centered Design Process for Iterative Systems has four main components (Stedmon & Stone, 2001):

1) Understand and specify context of use involving the users, the computer-based system under design, and the environment within which it will be utilized
2) Determine the functional relationship between users and technology
3) Iterate design solutions including user testing at each stage of design and implementation cycle
4) Seek input from a multidisciplinary team (e.g., software engineers, human factors specialist, ergonomists, etc.) to ensure safety while maximizing performance

Within the user-centered design paradigm, usability evaluation is a key component. According to Preece, Rogers & Sharp (2002), the main goals of usability within the design cycle are to ensure 1) system effectiveness: meets its specified use; 2) efficiency: supports the tasks; 3) safety: protects user from adverse effects, including negative training; and, 4) utility: meets the required goal (p. 12). However, extending usability metrics from graphical user interface design (GUIs) and web-based applications to virtual environment design has not proved satisfactory (Hix & Gabbard, 2002; Bowman, Gabbard & Hix, 2002).

15

Kaur, Maiden, & Sutcliffe (1999) outlined a possible methodology for applying theoretical frameworks to assess interactions within an immersive VE. They created a hypertext tool to assess objects, actions, and system controls within an immersive VE. Figure 1 diagrams the VE design stages determined by empirical evaluation. This general model commonly sited in the literature omits cognitive processing models as well as sensory criteria to understand interactions at a more top level (Kaur, Maiden, & Sutcliffe, 1999). The generality and top level approach may not be appropriate for the needs of designing some VE applications utilized for cognitive rehabilitation, which may require multimodal sensory baselines to assist in determining effects of training as well as mediating adaptation effects.
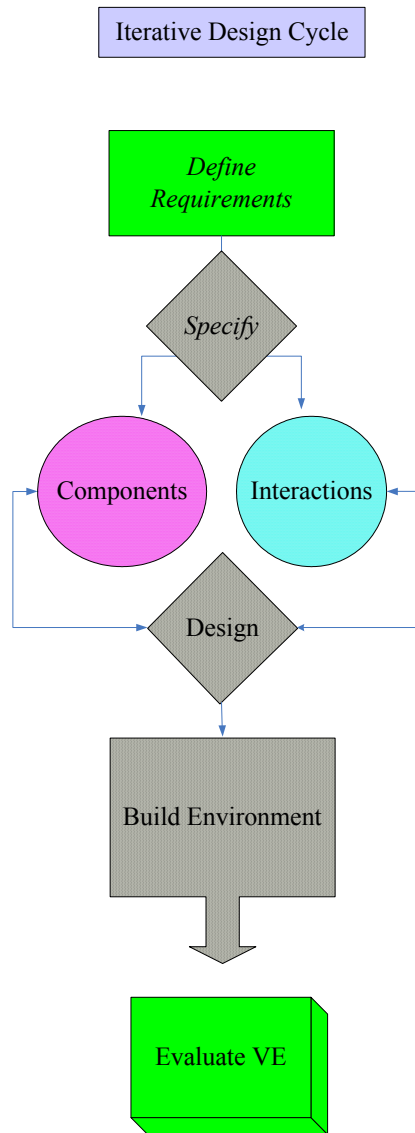
Figure 1 VE system design stages modeled after Kaur, Maiden, and Sutcliffe (1999).

Stanney et al. (2003) designed an automated systematic software tool which integrated all current recommendations for usability engineering as it pertains to VEs. The Multi-criteria Assessment of Usability for Virtual Environments (MAUVE) included design considerations which evaluated multimodal output capabilities as well as cybersickness and aftereffects. Following Gabbard, Hix, & Swan (1999), these criteria were broadly qualified under headings of VE system interface (e.g., technology and software) and VE user interface (e.g., sensory and psychological). Although extendable, the MAUVE system currently cannot evaluate criteria listings for VE designs for cognitive rehabilitation protocols. Also, there are no quantitative metrics evaluating or comparing sensorial throughputs as benchmark criteria for the different VE system interface recommendations.

Bowman, Gabbard, & Hix (2002) proposed an alternative method of usability classification accounting for the VE type (i.e. general testbed or specific application), the necessity of human evaluators, and the testing output (i.e., quantitative or qualitative metrics). The performance results of a testbed usability evaluation are more generic and lend to more universal guidelines for VE design. For each specific application, a sequential evaluation approach is suggested (Gabbard, Hix, & Swan, 1999). This evaluation method employs usability methods such as cognitive walkthrough whereby the user steps through the VE required tasks; formative evaluation where users provide quantitative and qualitative assessments; and comparative evaluation which is a statistical evaluation of two or more VE system configurations.

Because of its ease and cost effectiveness, a sequential evaluation approach was chosen as the usability assessment for constructing the mixed reality kitchen reported in detail in Chapter 9. Sensory benchmarks are currently not a component of either usability evaluation.

18

However, researchers have used the VE testbed approach to study human performance aspects of usability, including some perceptual metrics (Lampton et al., 1994; Bowman, Johnson & Hodges, 2001; Griffiths, Sharples & Wilson, 2006).

2.2 Human Performance Metrics and Method

Travis, Watson & Atyeo (1994) argued that user performance criteria should drive VE design to avoid over designing the application with useless and expensive sensory devices (p. 53). Wann and Mon-Williams (1996) additionally argued that the user's perceptual-motor capabilities should provide the foundation for VE system design. These capabilities should be evaluated within the context of user tasks. Wann and Mon-Williams made the point that basic research has not completely elucidated the workings of our sensory organs, including the visual system (p. 833). They successfully argued that the term "virtual reality" was an oxymoron and should be replaced by the term "virtual environments" (see also Kalawsky, 1993, p. 7). Within this construct, veridical representations of some aspects of the real world are unnecessary for VE applications. One component of the real world that should remain consistent is the representation of three-dimensional space.

According to Wann and Mon-Williams (1996), the perception of three-dimensional space is a defining element of VEs along with the potential for interaction (p. 833). Space perception is the ability to visually perceive the spatial layout of our three dimensional world. As such, perception of space is an output of the sensory processing of visual information (Cutting, 1997). Two properties of visual space perception are direction and distance with reference to the observer (Hershenson, 2000). Wann and Mon-Williams (1996) contended that the ultimate goal of correctly represented space was to support navigation and exploration about the space and

manipulation of object contained within the space. Lampton et al. (1994) utilized these
components as the basis for their VE Performance Assessment Battery (VEPAB).

2.2.1 Virtual Environment Assessment Battery (VEPAB)

VEPAB was constructed 1) to determine minimal trials necessary for achieving
competency at interacting within VEs; 2) to identify system component limitations (e.g., effect of
low display resolution on task performance), and 3) to quantify incidences of cybersickness and
negative aftereffects. Participants performed a range of activities providing feedback for task
categories such as vision (e.g., recognition visual acuity), locomotion (i.e., walking or flying),
manipulation (e.g., turning dials), tracking (e.g., moving the input device) and reaction time (i.e.,
target location). The choice of utilizing generic tasks was to maintain baseline performances
longitudinally as new technologies were introduced over time.

Main results showed that for the low resolution (234 horizontal lines by 238 pixels per
eye) Virtual Research Flight Helmet used, participants experience the equivalent of 20/860
resolution acuity, where normal acuity is 20/20 or one arc minute. Also, participants were less
accurate at determining closer distances in the VE than those farther away. This effect was not
noted in real world viewing. Only one person experienced acute sickness symptoms and no after
effects were reported. Performance on locomotion tasks involving turns required many trials to
demonstrate improvement levels capable of supporting more complex tasks such as tactical
decision making. This result supported the authors' claims that novice versus expert differences
in interacting with VEs may hinder training. Several criticisms of this testbed approach are: 1)
the interaction tasks were not chosen within a formal framework such as creating taxonomies of
possible interaction combinations; 2) outside factors such as spatial ability of users' were not

accounted for (Bowman, Johnson, & Hodges, 2001); and, 3) the presentation of tasks were too rigid to allow for clear distinction between novice versus expert users (Griffiths, Sharples, Wilson, 2006).

2.2.2. Structured Approaches

Griffiths, Sharples, & Wilson (2006) developed a performance assessment tool to classify participants as novice or expert VE users. This integrated application tool established baseline metrics on simple navigation interactions such as turning corners and moving backwards and forwards within a maze. Also, the object manipulation task took place on a virtual international space station. These activities are similar to those performed in the VEPAB; however, the performance metrics chosen were meant to function as if in a game with multiple levels of difficulty. The design of such a testing environment required a systematic design cycle (Wilson, Eastgate, & D'Cruz, 2002). Although portable, this application is based upon immersive VE and may not translate well to other types of head-worn displays that require different setting for field of view and interpupillary eye distances. The design cycle approach seems quite extendable to cognitive rehabilitation paradigms.

Bowman, Johnson, & Hodges (2001) applied a strict testbed approach to the assessment of human performance in VEs. Their purpose was to evaluate all possible types of general or universal interactions (e.g., different ways to select an item in 3D space) utilizing a taxonomy of interaction techniques to classify and guide the testing procedures. The top level of the taxonomy hierarchy represented the general manipulation and selection techniques followed by the specific subclasses of the interactions (e.g., object orientation). These distinctions were further qualified by technique of user response (e.g., turn dial). User ability (e.g., spatial ability) and system

characteristics (e.g., binocular viewing) were also taken into account. The response variables were then statistically compared. However, the system characteristics were included as categorical variables and were not further quantified with respect to the user. Baseline performance within and between integrated technologies might not account for individual differences or interactions involving persons with disabilities.

2.3 VE System Design with Clinical Populations

Cost is one prohibitive aspect to the design and the assessment of VEs for cognitive rehabilitation. In addition to the technology (display, tracker, haptics, etc), software, application development, and maintenance are additional expenses. Most applications for cognitive retraining require custom applications (Weiss, Rand, Katz, & Kizony, 2004). However, there is some evidence that when designed correctly base VE platforms can be validly and reliably applied across therapy scenarios (Rizzo et al., 2006).

Rizzo, Buckwalter, & Van der Zaag (2002) first outlined design criteria for VE applications created for clinical populations. Cost-benefit analysis guided by the rule of "elegant simplicity" motivated therapists to consider the necessity of a VE solution in lieu of currently available methods (p. 1047). "Good Fit" assessments determined how well the VE solution presented real world attributes in a more controlled, repeatable manner that allowed for comparable results over treatment effects (p. 1048). This point raises an important issue: VE solutions for cognitive rehabilitation are mostly designed to capture data necessary to evaluate levels of cognitive function or transfer effects pre and post rehabilitation. As such, they are inherently guided by experimental design and scientific principles. Yet, there are no standardized

methodologies for creating virtual rehabilitation environments, which leads to redundancy of systems and platforms.

Lack of standardization also makes comparisons across research endeavors difficult (Moreira da Costa & Vidal de Carvalho, 2000). The use of open source applications is a possible solution; however, understanding which capabilities (e.g., 3D audio) a VE must possess to accomplish successful retraining and transfer is still an open question. Transfer of training and monitoring aftereffects may require a modifiable multimodal testbed, which fits the needs of the various rehabilitation protocols.

While HCI guidelines are important, their goals and methods may change to accommodate for the sensory and perception deficits of this user population. For example, persons with retrograde amnesia may not remember their VE exposure even 10 minutes post training. Questionnaires or verbal report methods may not be appropriate for this population; however, procedural tasks that support motor memory may prove useful. We experienced these design issues when constructing the mixed reality kitchen described in Chapter 9 and propose design guidelines for this population. Thus, VE design for cognitive rehabilitation must match research concerns with the comfort, the safety, and the acceptability of the user. Research and safety constraints may be best satisfied with baseline metrics representing both clinical and unimpaired populations.

2.4 Moving Forward: Sensorial Baselines in the VE Design Cycle

If three-dimensional space representation is the hallmark of VEs (Wann & Mon-Williams, 1996) and if individual differences account for more variability in user performance than system design factors (Kaber, Draper, & Usher, 2002), then why are there no

comprehensive test batteries that attain baselines for different sensory modalities? System characteristics (e.g., low contrast displays) under certain conditions degrade sensory information (e.g., brightness or contrast), thus distorting natural spatial relationships and movements within the VE. This distortion lends to misperception of the VE elements, limiting its usability. Thus, in some instances, beginning usability assessment at such a high level may misrepresent actual user performance. User individual differences such as those found within the heterogeneous TBI population (e.g., vision loss along with memory impairment) work in conjunction with the already perceptually degraded VE. In this case, individual differences and VE system characteristics are not independent factors and may lend to negative aftereffects and cybersickness. For example, Webb and Griffin (2002) found that low visual acuity might be a factor in motion sickness experienced in virtual simulations. Screening participants for low visual acuity may be important for some VE applications.

The Virtual Reality Test Battery for Virtual Reality Users (Howarth & Costello, 1997) was developed for the purpose of quantifying the changes in human perception after exposure to fully immersive VEs. The battery included a visual acuity measure for low (18%) and high (90%) contrast visual acuity. The researchers tested for changes in visual acuity over a 20-minute period of full immersion and found no significant changes. One of the major oculomotor changes observed after VE exposure was that of heterophoria, shifts from normal resting eye position that may create double vision (Howarth, 1999).

Researchers determined that the mismatch between the inter-screen distance (distance between the two displays within a HMD) and the interocular distance (distance between the centers of rotation of two eyes) caused the change in normal resting state position of the two eyes. Manufacturers of one HMD, Virtual i-glasses, worn during the experiment expected this

24

ocular change due to the optical design. However, most HMDs do not offer adequate documentation to determine these types of effects. In Chapter 5, we report on how a vision test can determine issues relating to lens tilt and misalignments of the ocular pieces within a see-through projection display.

Further, we have shown the added value of employing user-centered design methods during the design cycle of HMDs (Fidopiastis et al., 2005). The vision tests created are extendable to the general VE Design cycle in that they can predict user performance as to what the users can see and how well they can see it (Davis, 1997; Robinett & Rolland, 1992). Given vision baselines, one can perform higher level usability tests with the confidence that the 3D virtual space is perceived correctly and appropriately accounts for individual differences of the user. Including such metrics when designing virtual rehabilitation spaces is paramount to the successful evaluation of the remediation techniques employed as well as ensuring the safety of the head-injured participant.

Although we only report on a vision test battery in this work, we propose that sensory testing should occur after the VE scope has been defined and the task analysis is performed. Thus, appropriate technologies can be compared if either a testbed or a specific application usability study is performed. The proposed design flow is outlined in figure 2. In any case, the sensorial baselines are collected prior to administration of higher order tests such as the VEPAB. Then, the framework for the VE can be appropriately constructed followed by iterative testing. The final VE design should provide the validity and reliability required by cognitive rehabilitation protocols and the extensibility to apply to more than one rehabilitation scenario.

In Chapter 3 we outline HMD characteristics that lead to the development of the two vision modules. These technology parameters are discussed with respect to the human visual system. Then, in Chapter 4, we describe how the vision tests were developed.
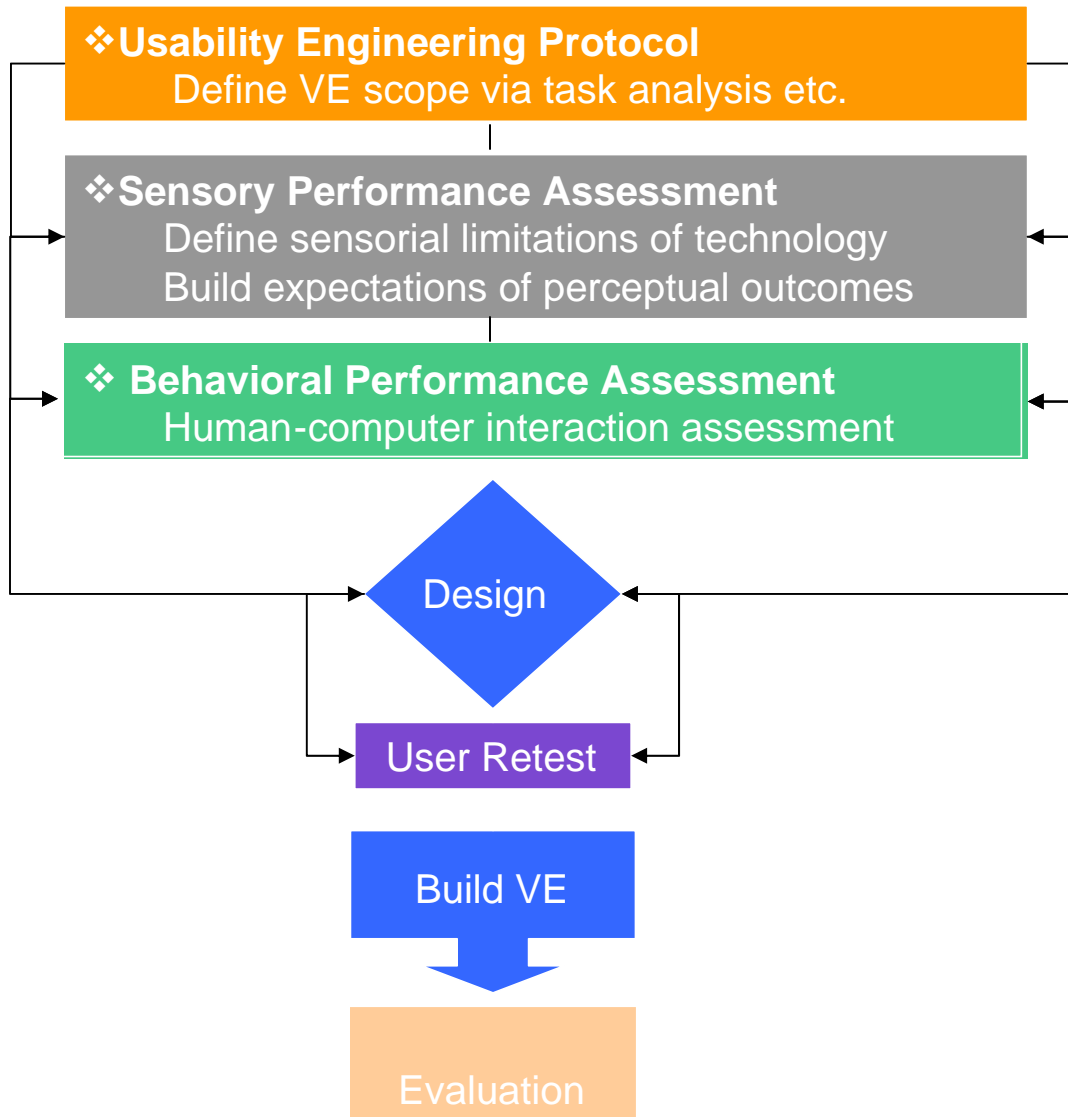
Figure 2 Proposed interactive iterative VE design cycle including sensory performance metrics for establishing baselines within a cognitive rehabilitation VE.

CHAPTER 3 FROM OBJECT SPACE TO VISUAL SPACE: OPTICS, COMPUTER GRAPHICS & PERCEPTION

In a review of the VE applications used in cognitive rehabilitation protocols, Rose, Brooks, & Rizzo (2005) noted that the field is rapidly expanding from solely assessing cognitive abilities such as executive dysfunction (impairments in sequencing and organizing behavior) to include rehabilitation training programs. This extension is now viable given the reduced cost of technology that was once considered as high-end (e.g., dual computer processing units) and advances in graphics technology (e.g., separate graphics card processing units). There are also off-the-shelf software programs such as Vizard made by WorldViz® that allow for easy virtual world construction. However, to extend current specific applications of cognitive assessment to training programs requires a more involved strategy of interactive design. It is imperative to understand the multimodal interactions between the human user and the technology components supporting the VE application.

For example, Duh et al. (2002) outlined scene characteristics that may influence user's perception of the virtual spatial layout. More specifically, the researchers were interested in discerning characteristics of the VE that elicited a sense of presence or the subjective experience of participating within the VE, while physically situated in physical space (Witmer & Singer, 1998). The VE was created utilizing a Kodack DPI 100 projector (1024x768 pixel resolution) and a 3ft dome (180 x 180 deg. FOV). A high resolution fountain scene and a simple radial pattern, both (600 x 600 dpi), were chosen as exemplars of complex and simple scenes. The researchers varied users' field of view (FOV), scene resolution, and scene complexity. Participants' changes in posture were measured from a balance platform.

The results showed that participant's postural stability was negatively affected by increasing FOV. The authors speculated that the increased peripheral stimulation caused the postural disturbances experienced by their participants. This result supports the functional sensitivity hypothesis suggested by Stoffregen (1985) whereby perceived self motion is mediated by the different processing capabilities of the cones in the fovea and the rods in the retinal periphery. Further, more complex scenes presented with a large FOV caused the most postural instabilities in healthy participants. Although these findings must be replicated across display devices, these results are disconcerting for participants who present with vestibular imbalances or visual field cuts due to TBI. Exposure therapy utilizing VE technology is a promising rehabilitation strategy for persons with vestibular impairments (Whitney et al., 2002). However, often therapists are not always aware of such residual effects of the brain assault until the participant is already engaged in a therapy routine. Screening participants prior to VE exposure may assist in limiting negative aftereffects as well as lost data from empirical studies.

3.1 User-Centered HMD Design Cycle

Head-mounted display (HMD) system designers are well aware of the perceptual limitation imposed by the HMD on the VE user. Robinett & Rolland (1992) discussed problematic aspects of HMD optics for idealized observers. Nemire & Ellis (1993) extended Robinett & Rolland's work to account for errors caused during real world viewing conditions. Rolland, Gibson, & Ariely (1995) described procedures to quantify visual errors and propose a system calibration procedure for HMD users. Finally, Melzer and Moffit, (1997) suggested that user-centric design principles should be integrated into the HMD design cycle.

Eggelston (1997) outlined the interdependencies of HMD properties, computer graphics techniques, and their combined effects upon the user's perception of the VE (p. 41). Figure 3 is an adaptation of Eggelston's user performance model. More specifically, the model illustrates how errors in the hardware (HMD optics and display) and software (Computer Graphics) impact the user's ability to correctly perceive the constructed VE space. However, the user also contributes his or her individual differences in perceptual abilities (e.g., spatial processing) to the overall error. Thus, the model must include a user perception to image level interaction.

HMD System Properties

Optics   Display   Graphics

System Level

Scene Effects   _Time Delay_   Object Effects

Image Level

Low Resolution
Viewpoint Mismatch   FOV   Eye Box   Distortions   Structural Errors   Motion Errors   Noise

Scene Perception   Self Perception   Object Perception

User

User Performance
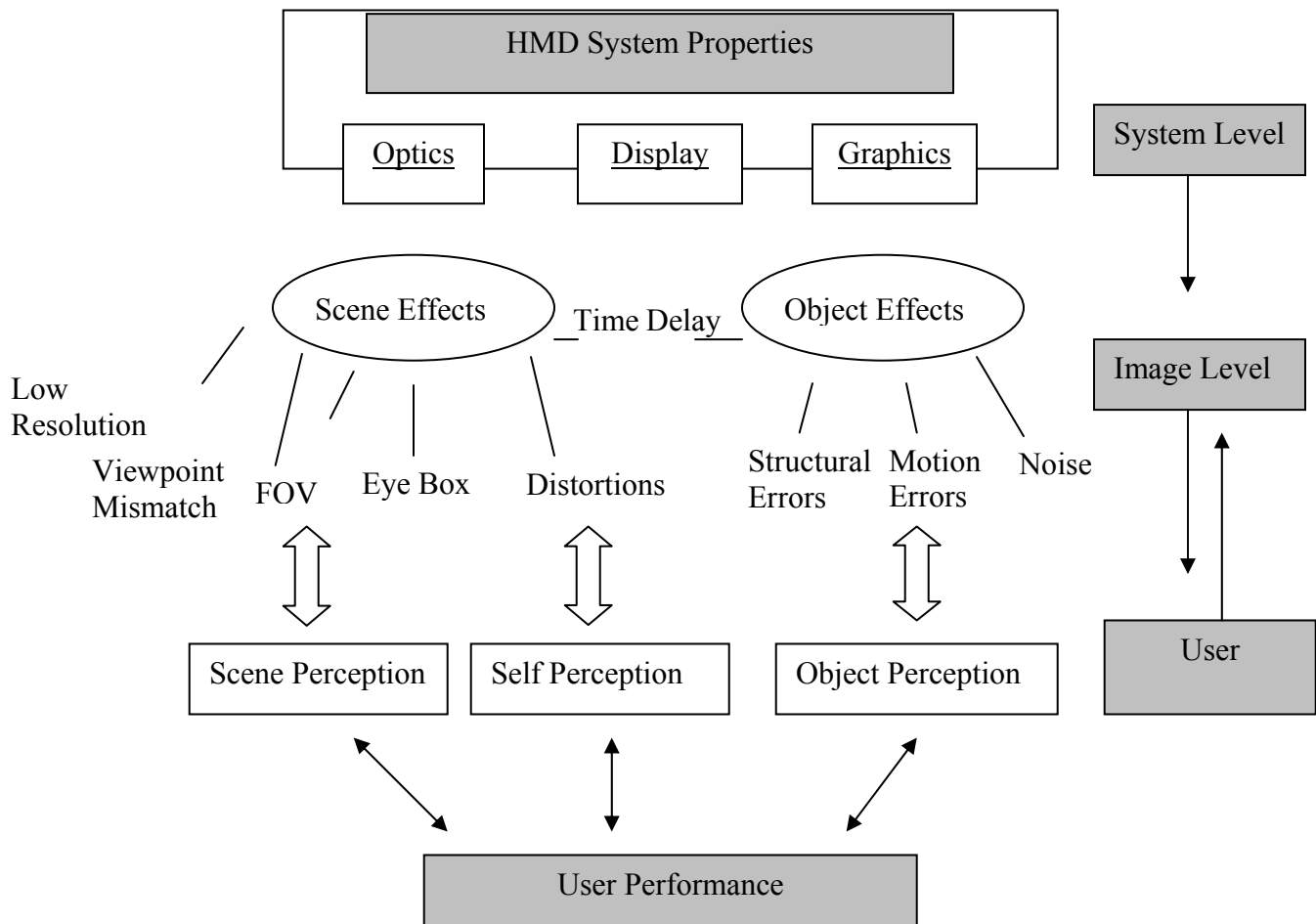
Figure 3 Modified user-centered approach to the head-mounted display design cycle adapted from R.G. Eggelston (1997). User-centered design in the trenches. In J.E. Melzer & K. Moffit (Eds.) *Head Mounted Displays: Designing for the user*.

From the user-centered design model it can be seen that the HMD designer is not only responsible for usability from a user's perspective, but from the software perspective as well. The graphics designers seem more concerned with rendering the three-dimensionality of the VE spatial layout. However, best practices of computer graphics techniques as applied to different types of HMDs would be useful in the design process.

The remainder of this chapter is organized by first describing the evolution of the design cycle for the Head-mounted Projection Display (HMPD) used in these experiments. The design cycle follows the recommendations of the field in assessing the prototype HMD from both the object space (e.g., space measured with rulers) and the visual space (e.g., the psychological space of the user). We then discuss the computer graphics component of the VE system and its interrelationship with the HMD and user. Embedded within each of these discussions are user constraints based on the human visual system. The human visual system will be discussed as it pertains to each section and then will be expanded upon to account for individual differences that are not accounted for by the HMD or software design cycle.

3.2 General Optical Design Issues

HMD designers (Task, 1997) are concerned with limits in HMD parameters such as display resolution (image quality), field of view (information quantity), and contrast (light intensity changes). Such parameters may be quantified by purely engineering means where the eye is replaced by a camera (Speck et. al, 2000) or by user-centered methods. The advantage of the former is speed of measurement, and that of the latter is a more comprehensive evaluation of the system as it applies to task requirements. User-centered testing during the prototype phase of display development has the added benefit of reducing design errors, which in turn tend to

reduced development costs. Additionally, performing user-centered assessments over the full

design cycle, allows for benchmarking, and thus a means of optimizing the VE system as a

whole. Methodologies for applying user-based assessments are difficult since computations are

based upon different characteristics of the human eye (e.g., resolution visual acuity).

3.2.1 HMD Optical Parameters in Object Space

The aim of this section is to introduce the reader to common aspects of the HMD design

cycle and how they relate to the human visual system. Although head-worn displays (HWD)

come in many varieties, we are limiting this discussion to those displays that mount two

miniature visual displays onto the user's head via a headband or helmet (Rolland & Hua, 2005).

A more comprehensive and current review of HWDs can be found in Cakmakci & Rolland

(2006).

Milgram & Kishimo (1994) categorized display systems based upon the Reality-

Virtuality continuum whereby immersive HMDs visualize a computer generated world while

see-through displays allow for visualization of a mix of real world and virtual objects. Although

the VE design may require a mixed reality construct, the concern of typical HMD design is

optimizing object space parameters of the image. As such, the assessments do not account for

aberrations of the human eye or the neural processing that occurs from the retina to the visual

cortex (Cakmakci & Rolland, 2006). Table 1 illustrates the main HMD parameters and their

human visual system correlate.

As Table 1 illustrates, the human eye is a highly adaptable optical instrument that can

adjust automatically to task demands and continue to perform optimally under varying conditions

such as lighting and near or far viewing distances. In contrast, HMDs must be constructed from

materials that are not as flexible as those of the human eye and are thus limited in adaptability to

scope of the task requirements.

Table 1 HMD parameters compared to the human visual system correlate. HMD parameters are adapted from Cakmakci & Rolland (2006) and Rolland & Hua (2005). The HMD parameters also vary dependent on task requirements.

HMD Parameters (Typical Values)	Human Visual System (Typical Values)
Monocular Field of View: 20° to 65° Binocular Field of View: 20° to 110°	Monocular Field of View: 160° horizontal Binocular Field of View: 200° horizontal
Image Quality: 2.0 to 4.1 arc min	Visual Acuity: .8 arc min in fovea
Exit Pupil: 12 mm	Pupil: 2 to 10 mm depends on lighting conditions
Luminance Range: 50 – 1000 cd/m^2	Brightness: Scoptic 10^{-6} to 3 cd/m^2 Photopic > 3 cd/m^2
Image Focus: Fixed	Accommodation: Dynamic and Cornea + Lens have a combined power of 58.6 Diopters
Eye relief: 15 – 50 mm	Eye Glasses imposed ER $\geq$ 17 mm
Optical separation between the L and R lenses , 55 mm to 75 mm based on IPD of user	Interpupillary Eye Distance (IPD): Distance between the two eyes, 55 mm to 75 mm for 95% of Asians, Caucasians, and African Americans

For example, Field of view (FOV), as it relates to the HMD, is the angle subtended by the

virtual image created by the HMD display as viewed by the observer when positioned within the

HMD eyebox, or pupil (Task, 1997, p. 60). For VGA resolution, text-based displays may be

monocular and require a FOV less that 40° to provide sufficient visual acuity (Cakmakci &

Rolland, 2006). In contrast, more advanced tasks (e.g., navigation) may require a FOV of more

than 60° to support the task requirements. The critical properties of HMDs in general that

determine FOV are the size of the HMD imaging element and the focal length of the optics.

More specifically, for a given focal length and a larger imaging element, the FOV

enlarges. Additionally, for a given FOV, a larger the microdisplay requires a larger the focal

length and near elements in the HMD design. A further constraint on the HMD FOV is the size

of the exit pupil. The exit pupil of the system will be inversely proportional to the F-number of

the optics. From an optical design approach, the larger the F-number, the least elements required

to reach a given image quality. Thus, the larger the pupil size, the more difficult the design

becomes.

The exit pupil of the optical system is the entrance pupil of the human eye, which ranges

in diameter (2 to 10 mm) with changing illumination levels in the environment in photopic to

scotopic ranges. In so doing, the pupil limits the light reaching the retina as a protective

mechanism and changes the depth of field or the distance interval where near and far objects stay

in focus without changing optical power (Pedrotti & Pedrotti, 1998, p. 171). Optical power for

the human visual system is obtained through the process of accommodation or the change in lens

shape as the observer fixates upon near or far objects. In optics terms the pupil can be considered

as a true aperture stop (adjustable diameter that limits the amount of light entering an optical

system) since it does not affect the FOV of the eyes (Roorda, 2002, p. 540). The adjustments of

the pupil are automatic and dynamic in response to the changing visual environment.

In order to see the complete range of the HMD FOV, the observer must view the

displayed scene within the diameter of the HMD exit pupil (Task, 1997). The exit pupil must be

sufficiently large to account for eye movements necessary for the eye to focus on relevant parts

of the visual scene. However, an adequate eye clearance and a large exit pupil size are

challenging design constraints since they require a larger diameter optical lens, which typically

requires more optical elements and thus adds bulk to the overall HMD design at a fixed image

quality specification (Rolland & Hua, 2005, p. 6). Figure 4 illustrates the typical HMD FOV as it

relates to the FOV of the human visual system. The image source responsible for HMD

brightness is the microdisplay element. More importantly, the microdisplay also limits the

resolution capabilities of the HMD.
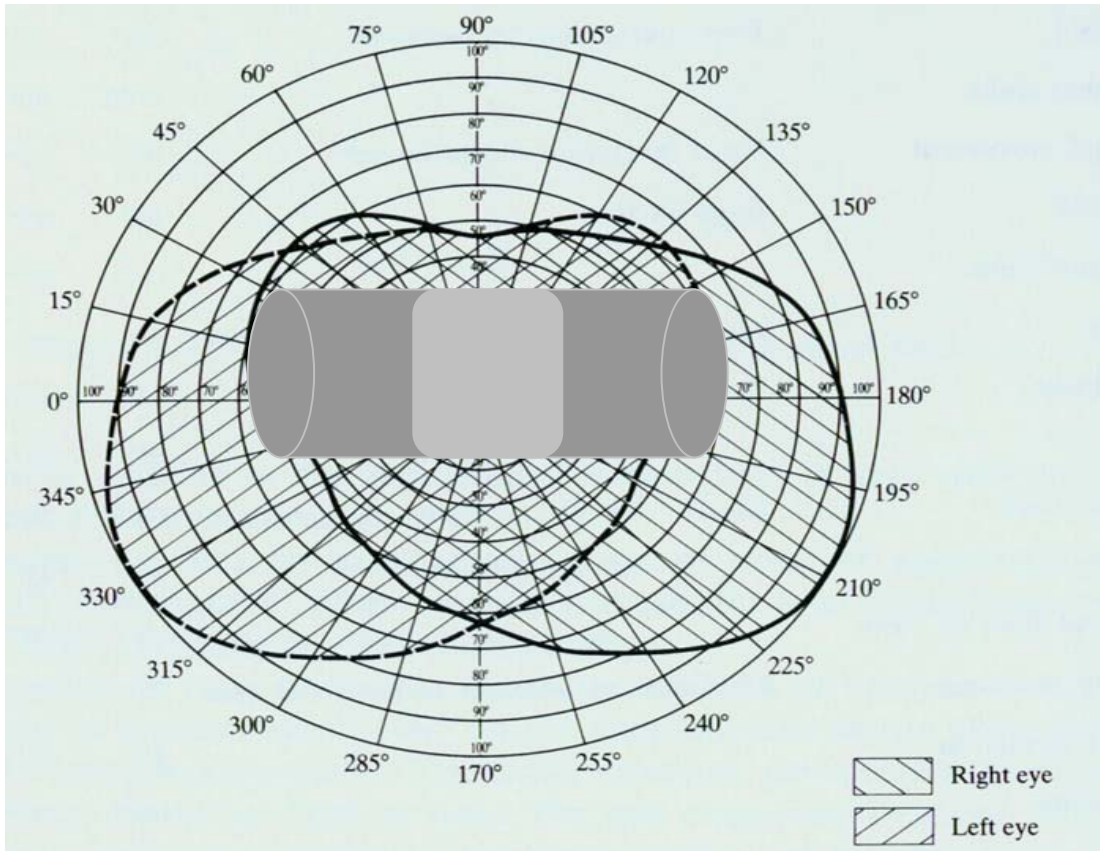
Figure 4 Typical HMD FOV values overlaid are typical FOV values for the human visual
system. Used with permission Rizzo, A.A & Kim, G.J. (2005). A SWOT Analysis of the Field of
Virtual Rehabilitation and Therapy. *Presence 14*(2): 119-146.

Microdisplays are typically less than 1.5 inches in diagonal and despite their size must

provide the brightness and the resolution to meet the vision requirements for the chosen

application (see Rolland and Hua, 2005 for a more complete review of microdisplays). Currently,

the most common microdisplay being considered for the cognitive rehabilitation community is

the OLED or organic light emitting diode. Because the OLEDs are self-emissive they require

less power and can be made without a backlight, which reduces its thickness. The lag between

the input to the pixel and the pixel response is less than 1 ms and provides uniform brightness across the display. The life span of these displays is less than 10,000 hours, which is less than any other microdisplay type. Despite the decreased life span, the allure of utilizing OLEDs is the potential to increase the number of pixels per degree of visual angle, which corresponds to the display resolution. Imaging technology known as photonics displays will likely compete on the market for HMD microdisplays given the ultra high resolution (e.g., 5 µ pixels) and brightness compatible with outdoor illumination under bright sun (Milliez, 2006)

Image quality for the HMD is best quantified by measuring the average subtense of a single pixel in either the horizontal or vertical dimension after being magnified by the optics. This resolution value can be easily compared to that of the human visual system. From Table 1, the average resolution of common microdisplays is 3.5 times less than that of the human eye. In more common Snellen acuity units, a person performing tasks while wearing such a display is operating at 20/70 acuity, as opposed to 20/20 acuity within the real world. Thus, there is also a trade off between HMD FOV and human visual acuity (VA) or the ability to resolve fine detail. More specifically, by increasing the HMD FOV without increasing the number of pixels, the magnified pixel reduces image resolution, thereby decreasing VA of the user. Visual acuity is a visual space parameter that has only recently been added as an assessment parameter in the HMD design cycle.

3.2.2 HMD Optical Parameters in Visual Space

The importance of assessing HMD image quality in visual space was stressed by Ha and Rolland (2002). Because of the novelty of their approach custom designed macros to extend lens design software were created. Their approach aimed to provide exact computations of metrics

36

typically used to quantify defocus in the human image processing in an attempt to provide better guidelines for human factors usability studies. The metrics most relevant to this discussion are astigmatism, accommodation shift, and Modulating Transfer Function (MTF).

Optical aberrations due to defocus such as astigmatism (unequal focusing of light in the perpendicular direction through the display) and accommodation shift (defocusing across the FOV) are optimized taking into account visual acuity and depth of field of the human observer (Ha, & Rolland, 2002). The MTF is a method of evaluating the change in contrast modulation (changes in luminance) due to the effects of optical aberrations over varying sinusoidal frequencies, thus obtaining an overall image quality indicator (Rabbetts, 1998, p. 47). Separate MTFs accounting for the peak spectral sensitivity values for the human system (e.g., 555 nm) and differing pupil sizes were obtained. The MTF for the human visual system (555 nm spectral sensitivity and 3 mm pupil) suggests a 63 cycle/degree spatial frequency cut off (Westheimer, 1986). Thus, scene detail presented below the cut off will not be in focus for the human observer. The ability to compare the HMD image quality with that of the human visual system is crucial to the development of benchmark measures that can guide the VE design process.

3.2.3 Moving Forward: User Visual Performance

Thus far, the HMD assessment techniques discussed optimize the HMD image via computer modeling while accounting for the human visual system via analogous metrics and accepted reference values. Fidopiastis, Fuhrman, Meyer, & Rolland (2005) extended the HMD image optimization procedure to include resolution visual acuity measures of actual observers. The methodology outlined by the authors took into account previous work that defined procedures to accurately render computer generated images as well as calibrate the HMD

(Rolland, Gibson, & Ariely, 1995). The results of these experiments are reported in Chapter 6.
We further extended the work of Rolland et al. (2002) to include a test for stereoaccuracy and
stereoacuity, both of which are measures of the quality of depth perception. The results of these
experiments are presented in Chapters 7 and 8, respectively. The adaptation of the initial work to
include human observers required an understanding of computer graphics techniques and their
interaction with HMD characteristics as well as their effects on the user's visual perception of
the rendered scene. We present an overview of this work in the next section.

3.3 Computer Graphics & Perception

Just as with HMD design, VE design considers the standard limits of the human visual
system (e.g., visual acuity, contrast modulation, and stereoacuity) as minimal user requirements
for optimal viewing of the VE scene (May & Baddock, 2002). However, some authors suggest
that visual errors may be caused more by the graphical techniques used to define the spatial
layout of the VE (Cutting, 1997; Ellis, 1993). Thus, even with a well designed and calibrated
HMD, the VE may not support proper viewing conditions for successful task completion.

In computer graphics, different views from a virtual camera render the perspectives of a
graphically displayed scene using perspective transformations. These calculations rely on
Euclidean geometry and matrix algebra to transform a three-dimensional view of the "world"
into two-dimensions and then, recover the z-axis or depth information to convert the image back
into three dimensions. This transformation procedure foreshortens depth because information is
lost along the z-axis (Foley, Van Dam, Feiner, & Hughes, 1995, p. 671). Cutting (1997)
contended that any graphical system that relied on a camera and lens faces the same depth
problem. For example, pictures taken with a longer lens will compress depth along the z-axis.

VE technology used to visualize computer graphics, such as HMDs, may compound the depth distortions. Robinett & Rolland (1991) recommended that the geometry of the computer generated scene must match the geometry of the HMD, namely the images must account for the relative position of the display screen, first order parameters and distortion, the optics and the eyes of the user. These recommendations are almost never cited in the literature on depth errors and stereoscopic HMD use.

3.3.1 Software Design, HMD Requirements, & Eye Mechanisms

Most literature reviewed suggests that visual cues foundational to depth perception in VE lack realism, and thus possibly contribute to visual errors such as depth misjudgments (Sinai et al., 1999; Loomis & Knapp, 2003). There are several issues with the previous studies that should be noted: 1) computer graphics capabilities are different today than even 3 years ago; 2) there are new graphics techniques to provide more realistic viewing conditions; 3) none of the research documented a calibration procedure for ensuring that the graphics are presented correctly to the user, and 4) HMDs are cross compared without taking into account their performance limitations (e.g., FOV and resolution). As the literature on HMD design suggests, there are software solutions available to render depth more accurately within VEs. The recommendations outlined by Robinett & Rolland (1991) for ensuring that the rendered graphics support the geometry of the HMD are still relevant today.

Of the seven recommendations listed, the easiest to correct is measuring the interpupillary eye distance (IPD) or the distance between the two eyes as measured by the distance between the centers of the pupil. In many older model HMDs, the IPD was set to an averaged standard ranging between 63 and 65 mm. The exit pupil for these displays was either

sufficiently large (12 mm) to allow for eye movement or did not account for IPD differences among users. The IPD is important to vision because the offset position of the two eyes allows for binocular disparity, differences in viewpoint between the two eyes. Differences in image position on the retina can be in the horizontal or the vertical direction. Horizontal binocular disparity gives rise to stereopsis or the perception of three-dimensions (Steinman, Steinman, & Garzia, 2000). The IPD of the user should be set on the HMD and in the software.

A related issue is the depth plane of the virtual image and its effects on the accommodation and the convergence mechanisms of the eye. In many cases the image plane where the virtual world is rendered is set as a fixed distance from the user. This distance could range from action space (within arm's reach) to infinity. When the image plane is set closer to the user, the lenses of the eyes accommodate (change their optical power) and converge (directed nasally or inward). In contrast, for objects rendered at infinity, the eyes are adjusted parallel to one another and the lenses adjust their curvature to a desired optical power. However, computer graphics techniques reduce the virtual cameras representing the user's eyes to single fixed points or eyepoints. Thus, the depth of field for the virtual camera is set at infinity; therefore, all objects in view are in focus. Further, the 2D planes onto which the three-dimensional virtual objects are mapped must be level and perpendicular to cameras to define the straight-ahead direction (Angel, 2003).

Wann, Rushton, & Mon-Williams (1995) suggested that there are several potential issues involved with these rendering constraints. First, if all virtual objects are in focus within the active viewing area of the VE world, then the main cue for accommodation, namely retinal blur is absent. Second, the mechanisms for accommodation and convergence are linked and reciprocally exert influence upon each other as the user's point of fixation changes. These synchronized

40

movements are no longer coupled. Third, as the eyes converge, the centers of the pupils also move. The pupil movement changes the IPD for the user. Since, most HMDs are not equipped with an eye-tracking device, the constraint that a virtual object appears level and perpendicular to the virtual camera does not guarantee that the image will now be centered over the user's pupil or the exit pupil of the HMD. Each of these mismatches between real and VE viewing have the potential to degrade the visual information seen by the user. More specifically, visual acuity and stereoacuity, determinants of image quality for the human eye, may degrade depending on the complexity of the VE scene and task. Robinett & Rolland (1992) recommended that the correct IPD be also set in the software.

However, accounting for accommodation and convergence mechanisms requires new hardware (Rolland, Krueger, & Goon, 2000). Also the development of new HMD hardware and associated algorithms for eye tracking are an active area of research (Vaissie & Rolland, 2002; Hua, Krishnaswamy, & Rolland, 2006). Recently, Rolland, Ha, & Fidopiastis (2004) showed through simulation how a best fixed eye point may be chosen for rendering. VE scenes depends upon whether minimizing the angular errors or depth errors is a task requirement. For reducing angular errors, choosing the pupil as the eyepoint for rendering the VE scene is suggested. Depth errors can be best reduced by choosing the center of rotation of the eyes as the eyepoint. These recommendations only apply to object space and do not account for optical aberrations of the eye. Further, the issue of changing fixation point and thus changing accommodation and convergence over a dynamic range is not addressed.

3.3.2 Challenges for See-through HMDs

Tarr & Warren (2002) contended that the current advancements in VE systems have allowed for basic research to progress in areas such as wayfinding. The advantages of VE systems in basic research are the ability to quantify and qualify the characteristics of the stimuli. These independent variables need not be veridical copies of the real world; however, they must be presented reliably. For example, researchers have tried to create virtual mazes with which to assess rodents on navigation processing. These efforts failed until the visual requirements for the rat were visualized correctly by the VE system. More specifically, the rodents' visual system requires low luminance and resolution images presented over a wide field of view (Hölsche, et al., 2005). VE techniques have also been used to study depth cues used by fruitflies (Schuster, Strauss, & Gotz, 2002). These examples illustrate the need to match the human user's vision characteristics with that of the HMD and the graphically displayed world. However, this constraint is more difficult when utilizing see-through HMDs.

The purpose for performing tasks while wearing a see-through HMD is that the real and the virtual worlds are combined to make up the task space. For example, following the "elegant simplicity" principle, some types of therapy may be best performed utilizing virtual components (e.g., stove burners) along with real world objects (e.g., dials for setting heat) instead of trying to replicate the total rehabilitation setting in a virtual counter part. There are two choices for see-though displays and they are categorized based upon how they merge the real and virtual scene: optical see-through HMDs employ a semi-transparent mirror, while video-see through HMDs use a video camera (see Rolland & Fuchs, 2001 for a complete review). The eyepoint location mismatch seems a greater issue for these HMDs.

Biocca & Rolland (1998) showed that when the eyepoints of the rendered scene do not match the exit pupil of the display significant negative aftereffects occurred, especially for video see-through displays. The negative aftereffects were not sustained more than 30 minutes after VE exposure; however, the adaptation effects during task performance may confound results evaluating the success of transfer of training. Hua, Girardot, Gao, & Rolland (2000) introduced the first miniaturization of the optics for a new type of new head-mounted projection display that reduced visual distortion and allowed for correct occlusion of virtual objects. We utilize the first and second generations of the projection HMD, along with a video-see through and fully immersive display. The characteristics of each display will be detailed in Chapter 5.

3.4 Moving Forward: Establishing Benchmark and Metrics

In this chapter, we reviewed the design cycle of the HMD. There is a clear trend to incorporate the user within the design cycle not just as an extension of object space, but moving toward optimizing for user perception in visual space. This important step allows for a better understanding of human performance and better user safety. In addition, the interdependence among the HMD, computer graphics, and user was discussed. It is important from an experimental design perspective to understand this interplay when assessing outcomes from VE based rehabilitation protocols. Given a calibrated HMD and appropriately rendered graphics, visual errors should be minimized when performing required tasks within the VE. In the next chapter we discuss the construction of computer-based tests, which quantify the visual resolution visual acuity and stereoacuity of the user. These tests provide a step toward obtaining standardized baseline and benchmark metrics of the VE system.

CHAPTER 4 TEST BATTERY DESIGN & DEVELOPMENT: VISUAL ACUITY AND STEREOACUITY ASSESSMENT

Wann and Mon-Williams (1996) argued that VEs should support "salient perceptual criteria" such as binocular vision that allow for the appropriate perception of spatial layout (p. 385). Their contention that VEs design must center upon the perceptual-motor capabilities of the user is an important design criteria for extending VEs to rehabilitation scenarios. Rehabilitation scenarios involving ADLs may necessitate a level of realism beyond generic action spaces such as a grocery store to support transfer of training to the home.

Human performance testing in the visual domain may be difficult since when viewing a complex scene in general the visual abilities of the user are dependent upon the sensory characteristics of the virtual and real objects (e.g., brightness and contrast) as well as the layout of the VE space (Bullimore, Howarth, & Fulton, 1995, p. 804). Further, the human eye is an optical system that is functionally limited much like the HMD in such parameters as display resolution or image quality. Given that the HMD parameters are designed with respect to limitations of the human eye, assessments utilized to determine the functional limitations of the human eye are applicable to testing visual performance when the human eye is coupled with an HMD.

In this chapter, we review the physiology of the human visual system that limits image resolution. We then present an overview of standardized tests that measure these limits. Finally, we present a methodology for applying these assessments (i.e. resolution visual acuity and stereoacuity) to evaluating the HMD together with the other VE system components.

4.1 Physiological Limits to Human Visual Resolution

There are two types of photoreceptors, rods and cones, which contribute to the geometrical organization or mosaic of the human retina (Goss & West, 2002, p. 97). The spacing, location, and function of these receptors determine the limits in resolving power (acuity) of the human eye. Figure 5 shows the projection of the distal stimulus (object location and dimensions in physical space) onto the fovea, an approximately 1 mm in diameter "pit", creating a real and inverted image or proximal stimulus (size and location of object on the retina). In optics, the optical axis passes through the centers of the pupil, 3 mm behind the corneal surface, eye pupil, the posterior lens and is taken as the axis of symmetry from which all calculations for assessing image quality are based (Pedrotti & Pedrotti, 1998, p. 205). However, the visual axis projects through the center of the pupil to the fovea of the retina where the highest concentrations of cone photoreceptors are located (Shiffman, 2000).
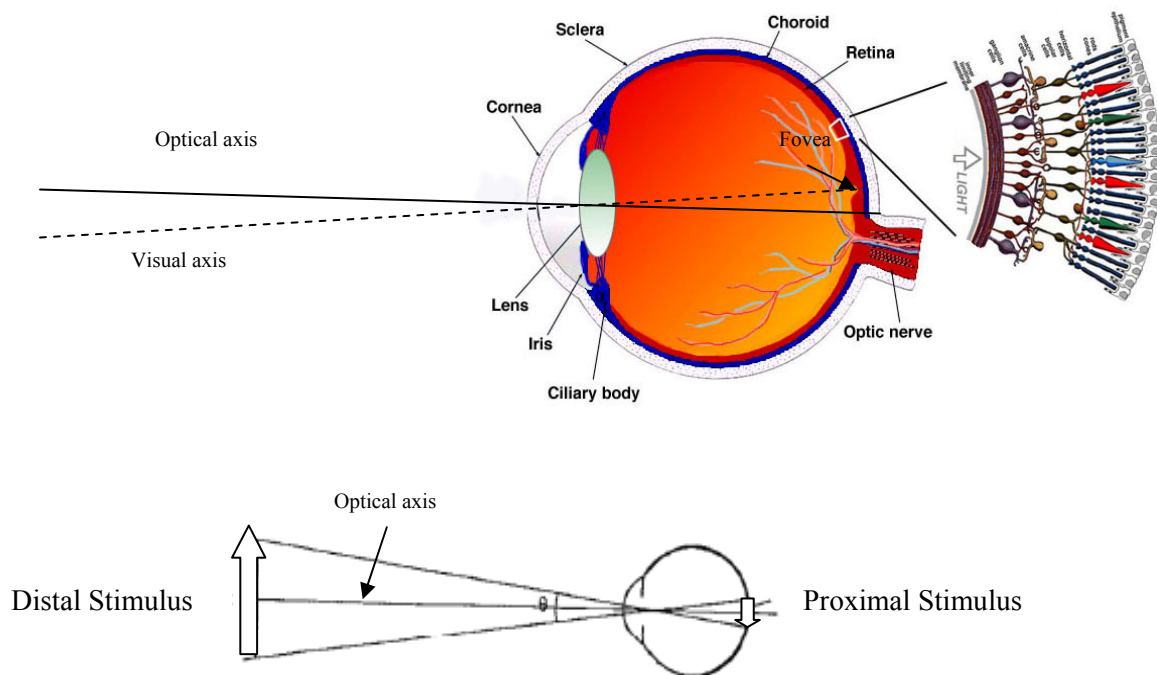
Figure 5 Adapted from Kolb et. al (2006)

The nonuniform density of rods and cones within the retina allows for each type of photoreceptor to perform different functions. For example, there are over 120 million rods located in the retinal periphery. These photoreceptors are neurally connected to summate their responses in an effort to detect low light levels in the environment. Thus, they are more active during scotopic or low lighting conditions, less than 3 cd/m^2 (Bullimore, Howarth, & Fulton, 1995). This many-to-one neural connection decreases the acuity capability of the rods.

In contrast, there are over 5 million cones maximally concentrated toward the foveal region of the retina. Cones maintain approximately a one-to-one connection between a single cone and a single bipolar and ganglion cell (Shiffman, 2000; Roorda, 2002). This neuronal configuration allows the cones to detect fine detail and color in the visual scene. The cones are active during photopic lighting conditions. Figure 6 schematically illustrates the different neurons within the retina along with their interconnections that propagate the neural signal to the optic nerve, terminating in the visual cortex of the brain.
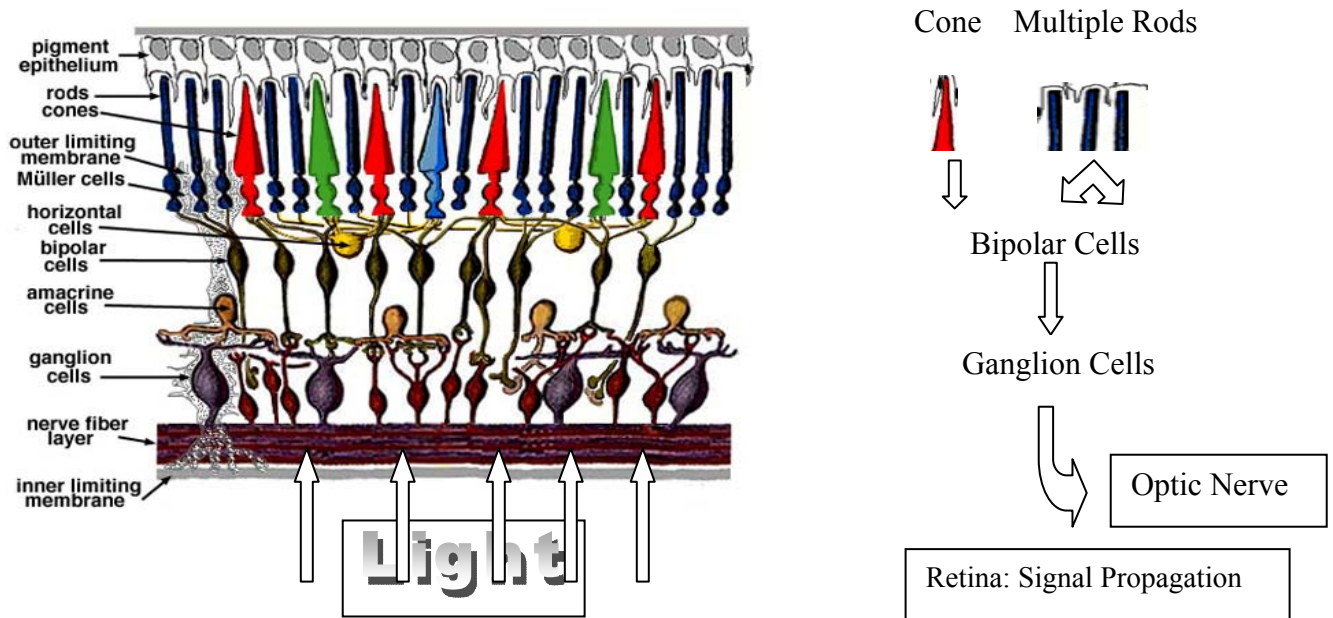
Figure 6 Retina physiology. Adapted from Kolb et al. (2006).

In the periphery, the receptor field size limits the amount of light being detected. Luminance levels below 10^{-6} cd/m^2 are not detectable (Bullimore, Howarth, & Fulton, 1995, p. 806). In the fovea, the distance between the centers of two cones is the limiting factor of acuity. Since this center to center distance is approximately 2 μm and at least 3 cones must be activated to determine the separation between two points of light, this limiting distance is 4 μm (Rabbetts, 1998; Goss & West, 2002). The standard limit of visual acuity is typically reported as 1 arc min. A 4 μm separation distance corresponds approximately to .8 arc minutes, thus the optical limits are well predicted by the physiology of the retina (Goss & West, 2002, p. 97).

There are other physiological factors that affect visual acuity levels of the human eye. For example, under scotopic lighting conditions the pupil size increases to allow more light into the eye. In this case, increasing pupil size also enhances the resolving power of the eye. However, the increase in pupil size during photopic viewing decreases acuity (Rabbetts, 1998, p. 23). The decrease in acuity is caused by spherical aberrations resulting from the inability of the cornea and lens of the eye to properly bring peripheral light rays entering the eye into focus when the pupil dilates beyond 3 mm (Newman, 1970). Additionally, background illumination (amount of light falling on an object) and luminance (intensity of light reflected from a surface) levels affect visual acuity through their perceptual correlates, lightness and brightness (Gilchrist, 1994).

4.1.1 Illuminance, Luminance, and Contrast

Photometric quantities are radiometric measures of light corrected for the spectral efficiency (range) of the human eye. The relevant photometric quantities of light are summarized in Table 2. The rods and cones of the human retina perform optimally under different levels

illuminance (lux) and luminance (cd/m^2). The separate ranges of optimal performance for rod

and cone receptors allow for the visual system to function under highly variable lighting

conditions. For example, under lighting conditions of low illumination, recognition visual acuity

is limited (e.g., reading in the dark) due to scotopic vision. Based on current technology, the

luminance within the HMD must match the environment to maintain contrast.

Table 2 Photometric values, definition, and measurement

Quantity	Definition	Units
Luminous flux	The quantity of radiant flux that creates a visual perception	Lumen (lm)
Luminous intensity	The flux emitted from a point source per unit solid angle	Candela (cd) or lumen steradian^{-1} (lm sr^{-1})
Illuminance	Flux per surface area. Mediates scotopic and photopic vision	Lux or lumen meter^{-2} (lm m^{-2})
Luminance	The luminous flux propagated from a surface point in a particular direction. Mediates brightness and contrast	Candela meter^{-2} (cd m^{-2}) or lumen meter^{-2} steradian^{-1} (lm m^{-2} sr^{-1})
Glare	A light source of higher luminance than anticipated by the visual system, which decreases visual performance	
Reflectance	The ratio of the luminance flux that strikes a surface to the luminance flux reflected off the same surface	

Low luminance (amount of light reflected off an object) causes the degraded visual

performance in the case of reading text in low light. More specifically, the issue is quantal with

respect to reading (i.e., not enough light is reflected from the page) and spatial with respect to the

HMD in that the contrast (difference in luminance between the surround and the projected

image) is lower. The scenario can be best explained by the terms "lightness" or reflectance and

"brightness" or the psychological equivalents of luminance (Walraven et al. 1990; Gilchrist,

1994).

Reflectance is the ratio between luminous flux reflected from a surface and the flux incident to the surface (Fiorentini et al., 1990). Reflectance changes in objects are perceptually perceived as having more or less light on a grayscale or in comparison to something white. Therefore, illuminance and more so the reflectance values control shades and shadows; however, these quantities do not change the perceived contrast. Contrast, as well as the proximity of other bright/dark objects in the vicinity, determines perceived brightness. Perceived contrast, as used in the forthcoming experiments, was calculated using the luminance (L) differences between the target and background as defined by the equation:

$$\frac{L_{Target} - L_{Background}}{L_{Background}} \tag{1}$$

Contrast is thus independent of illumination for high luminance levels (Shapley & Enroth-Cugel, 1984). However, illuminance does effect brightness perception when the neural system fails to maintain contrast constancy (e.g., the invariance of contrast to changing levels of light intensity) at low levels luminance (Heinemann, 1955). Both luminance and illuminance and the derived measure of contrast are necessary variables in assessing visual performance.

4.1.2 Measuring Illuminance and Luminance

For simulations, illuminance and luminance values can be estimated from a lighting compendium such as the Illuminating Engineering Society's publication, *Lighting Ready Reference* (Kaufman & Christensen, 1989). Otherwise, the values must be measured using an appropriate instrument, illuminance meter or photometer, prior to the experiment. An illuminance meter contains a photocell area in which to collect the photons of falling light. Illuminance measurement depends on the distance one is from the light source and the angle

between the area of measurement and the light source. The further away from the light source, the less light is collected. As well, the photocell should be perpendicular to the light source to obtain the maximum illuminance value; otherwise the value will be reduced by a factor of the cosine of the angle of rotation (Howarth, 1995). Unlike illumination, luminance measures are not dependent on distance. The main factor in measuring luminance is the size of the acceptance angle (light collection area). The photometer should be set perpendicular to the area of interest. The source itself should be larger than the acceptance angle of the meter. Given the VE environment other photometric quantities may be needed. Some of these quantities are listed in Table 2.

Most photometric measures can act as independent variables in the performance metric. For example, a user should experience degraded performance while using a see-through projection display under conditions of high illumination and low luminance. A simple yes/no response in a target detection task over varied light and luminance levels could determine the point where user performance fails. Where user bias is problematic an m-alternative forced choice paradigm would be necessary (Macmillian & Creelman, 1991 p. 125). We discuss these experimental design issues more in-depth in the next sections.

4.2 Standard Visual Acuity Tests for HMD and VE System Performance Testing

For HMD prototyping and VE System assessment, choosing the most appropriate perceptual test is critical for predicting tasked-based human performance while wearing the HMD. Table 3 provides examples of visual acuity tests used by clinicians to determine the status of visual functioning in patients and by vision researchers to determine the limits in acuity under various experimental conditions. Of the visual acuity tests, the Snellen test is the fastest to

perform, and is therefore the most common. The Snellen test results are used mostly for

evaluating tasks where letter recognition is important such as a driver's ability to read or

recognize lettering on traffic signs. However, when evaluating visual perceptual issues regarding

VE, such as the limited resolution of the HMD microdisplays, resolution visual acuity (rVA) or

the ability to resolve small visual angles is a better metric (Bullimore, 1997). Resolution visual

acuity, therefore, provides a metric of evaluating the user's ability to see fine detail within the

VE scenario.

Table 3 Examples of visual acuity tests (Adapted from Shiffman, 2000, p. 99)

Detection	Detect target in visual field	Examples
Vernier or localizing	Identify displacement of two lines in space	Manual vs. Gaze pointing
Resolution (rVA)	Perceive the separation of two distinct elements in space (e.g., resolve gap size in the letter C)	Landolt C
Recognition	Name targets in space (e.g., identify letters)	Snellen
Dynamic	Locate moving targets in space	MultiCAD
Stereoacuity	Align two targets while viewing at differing distances	Howard-Dolman

Given that the perception of three-dimensional space is foundational to VE system

design, evaluating the user's stereoacuity (the ability to judge relative depth from disparity)

within the integrated VE system is also an important metric. Although depth information is

carried by the retinal receptors, limitations in depth discrimination may be caused by the

processing of depth information by the cortex, including regions beyond those of the visual

cortex (DeAngelis, 2000; Neri, 2005). We will discuss these biological issues in the section

pertaining to stereoacuity assessment. It is important to note that given that higher cortical levels

are involved in processing space information, task demands and individual differences in

processing salient VE scene information contribute to human performance error in judging

spatial layout of VEs (Wanger, Ferwerda, & Greenberg, 1992; Cutting & Vishton, 1995). Thus,

stereoacuity baselines may assist in identifying HMD types that more appropriately match tasks requiring binocular vision.

4.3 Psychophysics in Vision Test Design

Reliability, replicability, and cross platform compatibility were important features of the assessment design. Further, baseline metrics should be meaningful when compared to standardized test scores that reflect real world viewing. Thus, we chose vision tests that had an established methodology and real world correlate. Because of the similarity in display constraints, we adapted psychophysical metrics used in optimizing Night Vision Goggles (NVG) such as "frequency seeing" curves to the assessment of prototype HMDs (National Academy of Sciences, 1980; Pinkus & Task, 1998; Fidopiastis, Fuhrman, Myer, & Rolland, 2005). Within the framework of classical threshold theory, the relationship between physical properties of the world such as changing levels of light illumination and human sensory perception are quantified. Although more modern methodologies such as Signal Detection are available, these methods of analysis may be more appropriate for detecting targets in complex and noisy background (Rolland, 1990).

As classically defined, the absolute threshold is the stimulus intensity or magnitude where a positive detection of the stimulus is made 50% of the time (Schiffman, 2000). There are several statistical tests available to evaluate the absolute threshold. Probit, best PEST, and Quest tests have all been used to evaluate visual acuity type data (Finney, 1980; Leberman & Pentland, 1982; Watson, 1983). Adaptive methods such as best PEST and Quest use an adjustable procedure where the individual's response determines the next presentation of the stimuli. There are several computerized tests that measure resolution visual acuity (Bach, 1996) and

52

stereoacuity (Bach, Schmitt, Kromeier, & Kommerell, 2001) utilizing bestPest procedure. Our concern with this type of method was accumulating errors over the large number of trials being conducted by a single subject (Kaernbach, 2001). Pinkus and Task (1998) provide some support for reliability issues using an adaptive method as a threshold estimator. We, therefore, chose a Probit analysis in conjunction with a modified Landolt C test to determine user rVA thresholds.

Three psychophysical methods are typically used to measure stereoacuity thresholds: method of adjustment, method of limits, and method of constant stimuli. Within the method of adjustment paradigm, the observer positions a moveable object offset from a fixed target to be equal with the target. Each offset is randomly chosen and the mean deviations are used to determine stereoacuity. For the method of limits, the observer adjusts the moveable object to be a just noticeable difference (JND) in depth from the fixed target. The ranges of these JNDs are used to estimate stereoacuity. Reading (1983) considers the method of constant stimuli to be the gold standard for evaluating stereo thresholds (p. 113).

The method of constant stimuli requires the user to determine the first position perceived consistently closer than the fixed target and then the position perceived just farther than the target. The range is then partitioned into equal intervals and the moveable rod is placed at each position in random order. The response variable obtained is the percent near responses or percentage of times the observer adjusted the moveable object to be in front of the fixed target. A Probit analysis could be used to determine stereoacuity (Rolland, Gibson, & Ariely, 1995). However, Rolland et al. (2002) investigated the appropriate methodologies for measuring depth perception in a VE, which also controlled for the size and the shape of visually presented stimuli. The study pointed to the method of adjustment as a preferred method over the method of constant stimuli when measuring stereoacuity because with the MOC the observer can learn the

mean size presented within two blocks. If the size is randomly varied to minimize that effect, the data have been shown to be extremely noisy thus requiring more data acquisition and increasing the duration of the experiment. Another concern is lessening the amount of time within the HMD so that any adaptation effects do not affect the results.

In sum, when using psychophysical measures, we can create testable conditions that quantify the relationship between the physical parameters of the VE system (e.g., limited resolution) and the representative visual performance metric such as resolution visual acuity (May & Badcock, 2002; Uttal & Randall, 2001). Based on the knowledge of how the system should perform, we can also make a priori predictions of how the user will perceptually perform (Meister, 1985). Assuming that the HMD is calibrated, vision tests could discern limitations attributed to each separate aspect of the VE system or the system as a whole. The end result is a basic, yet rigorous optimization procedure that lends to a more comprehensive evaluation of the quality of system and user interactions (Melzer and Moffit, 1997).

4.4 Modified Landolt C Visual Acuity Test Overview

The Landolt C visual acuity test is standard for measuring resolution acuity (National Academy of Sciences, 1980). The Landolt C is constructed on a 5 x 5 grid with the gap width of 1 grid unit as shown in figure 7. The gap is randomly presented in one of eight positions. The observer must determine the direction of the gap. From a visual perspective there are two drawbacks in using the Landolt C: 1) uncorrected astigmatisms may make some orientations of the gap easier to see, and 2) guessing may show biased performance at low acuities, usually toward the right (Rabbett, 1998, p. 32).

Figure 7 Standard Landolt C visual acuity test

Landolt broken circle.
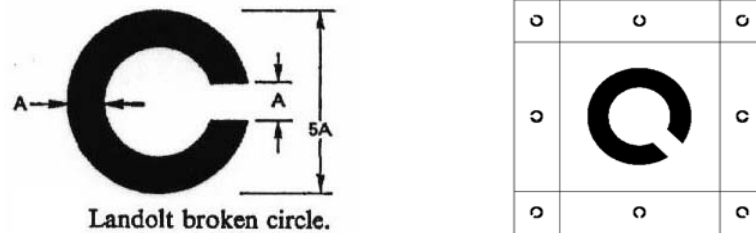
Others have found a "gap-down" effect, where correct responses were lowest for downward facing gaps (Schrauf & Stern, 2001). To control for these potential confounds, all participants were tested for resolution acuity using corrective lenses, if needed. As well, the participants were not constrained by time in responding. Although guessing still occurred, there was no added pressure to perform under a time limit. The more challenging problems were translating the printed version of the test into a software application.

One main issue in converting the Landolt C test into a software application was accounting for anti aliasing effects. Typical stimuli for the Landolt C test, pictured on the left of figure 8, are circular C's. The circular shape provides radial symmetry, which facilitates equal blur along the curve of the C (National Academy of Sciences, 1980, p. 116). However, rendering a circular C utilizing computer graphics techniques requires the use of anti-aliasing methods to remove the "stair-stepping" effect found in low bandwidth displays. Figure 8 shows typical Landolt C stimuli with aliasing and anti aliasing effects present when rendering the curvature of the C.

Figure 8 Example of aliasing and anti-aliasing effects

To remove these effects, a square C was created that retained the 5:1 ratio between the C size and the gap size. To further reduce aliasing and anti-aliasing effects the VA test was changed from an 8 AFC to a 4 AFC. Figure 4 shows the view of the square Landolt C test from the left and right eyes within the HMPD.

Figure 9 Modified Landolt C visual acuity test
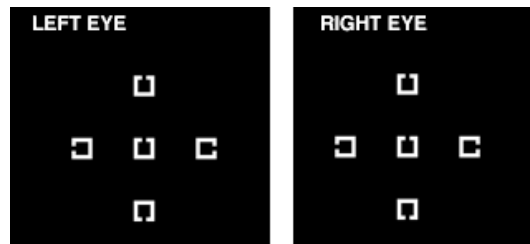
Receptive fields in the human retina are spatially tuned to respond best to specific sizes of stimuli in the visual scene. Visual acuity tests such as Landolt C assess only the smaller receptive fields that are responsible for fine detail processing. Furthermore, the recommended contrast for optotypes used in these tests is .85, which corresponds to high contrast targets (National Academy of Sciences, 1980). While the contrast level of optotypes may be varied there are no standards defined for varying contrast as a function of stimulus size.

The contrast sensitivity function is a mapping of perceived contrast as a function of spatial frequency over all spatial frequency channels. Resolution visual acuity is the limit of this function. Future work will be expanded to investigate how the range of human spatial frequencies is filtered in the VE system. Currently, we focus on the sensory threshold tests for size (i.e., Landolt C) as a function of target contrast over the smaller frequency channels.

4.4.1 Probit Analysis

In Probit analysis the percent correct responses for determining gap orientation are corrected for chance responding, then they are converted to z-scores or normal equivalent deviates (NED). In a 4AFC design, the participant has a 1 in 4 chance (25%) of guessing the correct gap orientation. Alternatively, in a two alternative forced choice design there is a 50% chance of guessing the correct gap direction. The percent correct responses are corrected for chance using the following procedure:

Let P = percent correct trials

PA = percent correct trials corrected for chance

(1) if P < 25% then P = 25%

$$(2) \quad PA = \frac{(Score - Chance)*100\%}{(100\% - Chance)} \tag{2}$$

For our design,

$$(3) \quad PA = \frac{(Score - 25\%)*100\%}{(100\% - 25\%)} \tag{3}$$

The assumption in step one suggests that anything below 25% will approach chance on enough trials, therefore the score is analyzed as 25%. This requisite ensures that any response

under 25% is still accounted for in the final analysis. The general equation for correcting for guessing is given in equation 3. The specific equation used to analyze our dataset is given in equation 3. Next, the corrected scores are converted to normal equivalent deviates (NED) by the following equation:

$$NED = \frac{X - \bar{X}}{\sigma^2} \qquad (4)$$

An NED is the value of a standard normal variable for a Normal distribution with a mean of 0 and standard deviation of 1. The cumulative probability equals the percent correct adjusted for chance. Equation 3 is actually the z-score, a statistic which quantifies the distance a data point is from the mean of the sample measured in standard deviations units. The NED values are used as the dependent variable in a linear regression with gap size (visual acuity) as the independent variable (e.g., $NED = b_0 + b_1*VA$). Curve fitting is accomplished using either a graphical method (plot probit values against the stimulus intensity and read where $p = 0.5$) or the exact method (use an unbiased estimator in an iterative procedure). Finney (1980) suggested using a maximum likelihood estimator with Probit analysis. The slope-intercept (b_0) and the slope value associated with the visual acuity, as computed from the regression analysis, is used to compute the expected NED values. The resulting NED values are converted back into percents and evaluated against the threshold value to determine visual acuity. For a 4 AFC design, the threshold is 62.5%.

4.4.2 Modified Landolt C Assessment for Augmented Reality

The current computer version of the Landolt C test is capable of assessing resolution visual acuity while wearing the HMD or under normal viewing conditions. However, the size of

58

one pixel is different when viewing a computer monitor versus when viewing the same pixel through an HMD. Also, magnification due to optics of the HMD changes the pixel size differently than for normal viewing over viewing distances. Therefore, the software is capable of accepting different resolution and FOV values for different HMDs. The code automatically scales the square Landolt C gap so that its smallest size is equivalent to the size of a pixel as defined by the HMD display parameters and the specified viewing distance. In addition, the IPD for the user is an input variable that allows the rendered stimuli to match the eye position of the observer.

To ensure that the rendered graphics matched the active area of the HMD display and the center of the pupil for the observer, the frame buffer overscan (FBO) was measured. The frame buffer is a memory store for the array of pixels constituting the visual scene (Angel, 2003). The resolution of the frame buffer determines the level of detail seen in the displayed image. FBO occurs because of anomalies in the graphics control device which truncates or stretches the active viewing area to misrepresent the number of viewable pixels (Rolland, Gibson, & Ariely, 1995). To overcome the FBO problem, we needed to code the application using software such as OpenGL that can control where in the frame buffer the rendered image will be appear (Fuhrman, Davis, Fidopiastis, & Rolland, 2006). The type of microdisplay (e.g., LCD) determines the type of FBO. In Chapter 5, we review calibration measures and discuss this topic further.

To create the contrast stimuli, the luminance of the Landolt C was varied by adjusting the binary grey-scale code within the computer program. The background of the display was set at a grey-scale value of 128. From this value, target grey-scale values giving contrasts of .2, .4, .6, .8, and 1 were computed using equation 1. The corresponding target grey-scale values are shown in

Table 4. How the modified Landolt C acuity test was adapted for each HMD assessed is outlined more specifically in the forthcoming experiments.

Table 4 Grayscale values and corresponding contrast

Grayscale Value	Bit Value	Contrast
0.5	128	0
0.6	153	0.2
0.7	179	0.4
0.8	204	0.6
0.9	230	0.8
1	255	1

4.5 Stereoacuity Assessment Overview

Rolland, Gibson, & Ariely (1995) investigated the effects of presenting collimated images (virtual image is located at infinity) to a user performing a depth judgment task between real and virtual objects. The authors calibrated the HMD and controlled for hardware issues such as frame buffer overscan (region of the frame buffer that does not get drawn onto the display screen) to ensure that the optics of the HMD aligned with the graphical display. Without the correction for overscan, a 5 degree error could have been introduced to the results with the microdisplay used. Other controls such as adjusting interpupillary eye distance in hardware ensured that the optics of the display were centered over the pupils of each eye. Adjusting interpupillary distance in software ensured that the objects were rendered at the correct depth.

The optical bench prototype display had a display resolution of 441 x 593, 40 degree diagonal

FOV, and resolution acuity of 3.6 arc minutes.

The stimuli in this experiment were either computer generated generic objects (i.e. a cube

and a cylinder) or real objects of the same type. Three conditions for the objects were tested:

real/virtual, virtual/virtual, or real/real. One object was always fixed (target) at a referent

location, while the other was moved in depth between trials (test). The object depths were

presented in near space at two depths, 0.8 and 1.2 meters, and were separated by 110 mm to

ensure that participants made their depth judgments by comparing the object center of the target

with the object center for the test stimuli. The participants performed a 2AFC task where they

determined whether the test cube was closer or further away than the target cylinder. Stimuli

were presented in a fashion appropriate for the method of constant stimuli assessment.

As expected, Rolland et al (1995) found no depth errors when objects were real/real or

virtual/virtual. However, in the mixed condition (real/virtual), virtual objects appeared further

away than real objects. Also, using optical ray tracing techniques, the authors found that

pincushion distortion in the individual monocular images had the effect of bringing the 3D

virtual object closer to the observer's eyes, in proportion to the amount of distortion. Thus, in the

case of this specific experiment, the actual bias may be slightly higher than the 50 mm measured.

As the authors pointed out, Iavecchia, Iavecchia, & Roscoe (1988) have researched the

effects of presenting collimated images to eyes using an optical device. The results of their

studies suggest that the eye shifts focus to its individual resting point of accommodation (dark

focus)[1]. The resting point of accommodation can be as close as 5 cm to almost 3 meters. Ehrlich

---

[1] Thanks to Dr. Jaschinski-Kruza at the Institut für Arbeitsphysiologie at the Universität Dortmund for assisting me in understanding the dark focus metric.

(1999) measured dark vergence and dark accommodation in VE, she found that dark

accommodation did not change; however dark vergence was shifted inward suggesting that the

observer's focus was closer than the resting set point. In her experiment, the participants

performed a visually directed walking task up to approximately 7.5 meters. The observers

underestimated distances in the VE, and after 15 minutes in the virtual environment, their dark

vergence set points changed. Given the plasticity of the accommodation and convergence eye

mechanisms based on task, the everyday work performed by the participants might have

influence their accommodation and convergence set point prior to entering the experiment

(Tyrell & Leibowitz, 1990). The set points for accommodation and convergence could also be

altered within the experiment. Therefore, measuring these values prior to performing a depth

perception experiment is critical when accounting for potential confounds based on individual

differences.

4.5.1 A Biological Overview of Depth Perception

Because the left and right eyes see two different views of the world, the proximal

stimulus on either retina is offset from one another. This offset is termed binocular disparity and

the resulting depth perception is called stereopsis (Backus, 2000). Table 5 and Figure 10 depict

the processing and neuronal connections from the retina to the cortex of the macaque monkey for

stereopsis. Although the lateral geniculate nucleus (LGN) in the thalamus receives a retinotopic

mapping of the retina, disparity coding does not appear until the visual cortex (DeAngelis, 2000).

Most simple and complex neurons in the striate cortex respond maximally to stimulation that is

within the corresponding retinal field of each eye. The recorded output from such cells is higher

than the intensity of the stimulus delivered to each eye independently suggesting that binocular summation is occurring (Steinman, Steinman, & Garzia, 2000).

Further, binocular neurons are tuned to maximally respond to different levels of disparity. For example, some cells respond best to zero disparity or objects that are at the same distance with respect to the observer's fixation point. Other binocular neurons respond best to objects beyond the fixation point, crossed disparity, or closer than the point of fixation, uncrossed disparity. However, more recent studies of the human visual pathway show that processing of disparity information, and thus depth perception, is not a lower cortical function as once thought (Backus, Fleet, Parker & Heeger, 2001; Parker, 2004).

Table 5 Magnocellular and parvocellular pathways Adapted from Kandel, Shwartz, & Jessell (1995).

Pathway	Perception	Stereopsis	Oculomotor
Magnocellular	Movement, Flicker, Depth, Contrast Sensitivity	Low spatial resolution, 40-36000"	Initiate pursuit & Vergence
Parvocellular Interblob	Shape, Orientation	High spatial resolution, 2-1200"	Maintain vergence
Blob	Color		

Figure 10 Motion, color, form & depth pathways
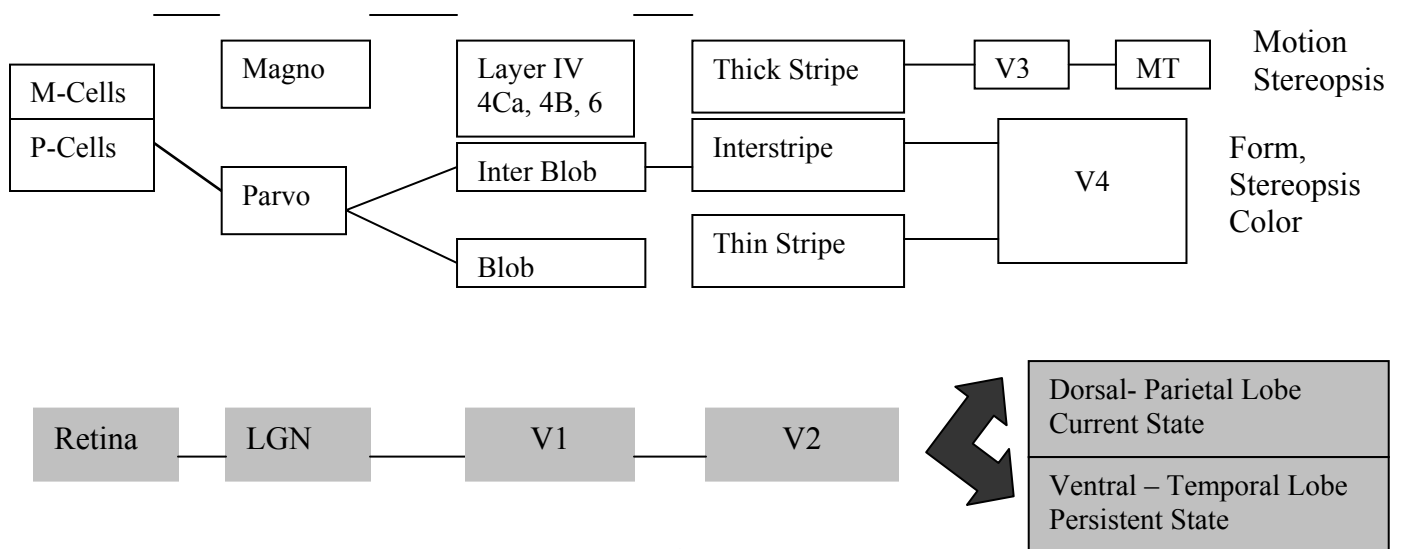
Figure 11 pictorially illustrates the visual processing pathway schematically illustrated in

Figure 10. This visualization shows that stereopsis is processed in two separate pathways, the

ventral stream and the dorsal stream. As shown, the dorsal stream terminates in the parietal lobe

and is thought to monitor visual control of action by providing moment to moment updating on

the location and orientation of objects. In contrast, the ventral stream terminates in the temporal

lobe and provides a more stable representation of objects allowing for object classification and

identification such as face recognition (Goodale & Haffenden, 1998). Functionally, stereopsis

seems to provide the dorsal stream with information needed to moderate vergence control, while

the same information is providing the ventral stream depth information (Neri, 2004).
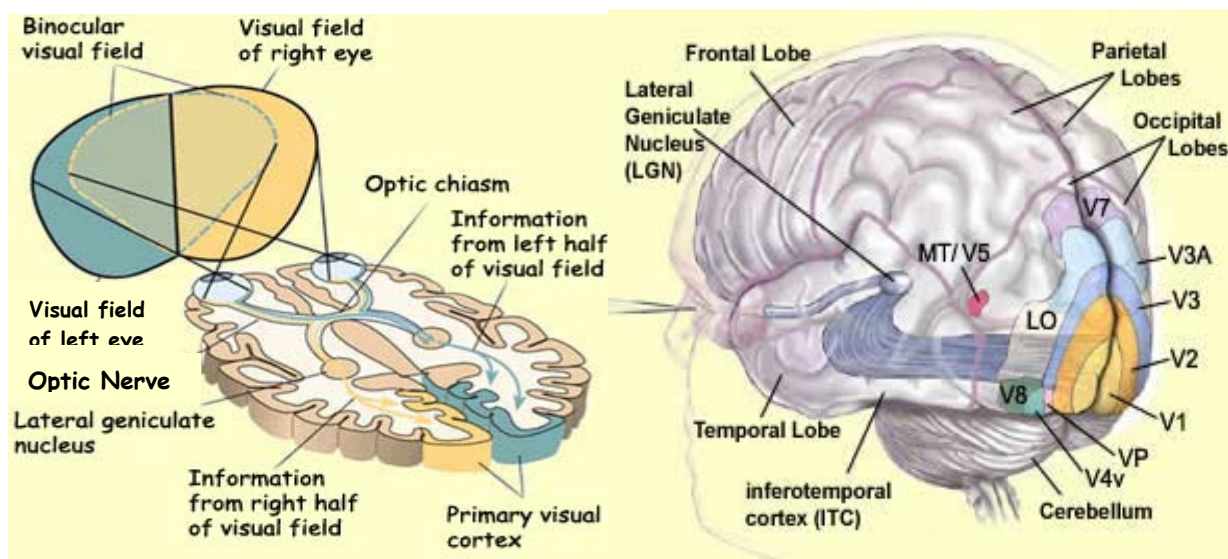
Figure 11 Visual information processing pathway from retina to cortex. Used with permission
from Dubac (2006) The Brain from Top to Bottom.

The functional distinction between the ventral and dorsal streams are important from a

Neuropsychological point of view in that many agnosias (loss of knowledge) such as spatial

agnosia are candidate areas for VE basic and applied research. Persons with agnosia may

experience no lower visual system disturbances, yet they will not form accurate perceptions

64

about the environment because of damage to higher cortical areas such as the temporal lobe. A VE scenario may provide a safe setting where virtual objects are quantified and controlled to accommodate the individual's deficits. However, potential confounds in VE design greatly hinder the validity of basic and applied research outcomes. Another issue regarding the dual function of stereopsis in the visual processing pathway is the locus of stereoacuity limits.

Banks, Gepshtein, and Landy (2004) used a random-dot stereogram paradigm to explore the locus of the stereoresolution (similar to the contrast sensitivity function) degrade. They varied factors pointing to the testing stimulus, the sampling limits at the level of the retina, or the correspondence issues in matching corresponding retinal points in the visual cortex. Their studies showed that for human observers with normal or corrected normal vision stereoresolution limits varied suggesting different contributions of stereo error along the visual pathway. For example, they found that for low dot density stereograms the paucity of information in the stimulus limits the threshold, while for higher dot densities retinal processing is the limiting factor. Further, the size of the retinal areas being matched by the cortical neurons may limit stereoresolution, an area too small may not provide enough information while an area too large may lend to higher ambiguity. Wong, Woods, & Peli (2002) also suggested various factors (cortical and more proximal) that affect stereoacuity.

Wong, Woods, & Peli (2002) used random-dot stereograms as test objects to study observers' near and far stereoacuity. The authors also studied the mismatch between accommodation and convergence mechanisms to detect shifts in stereoacuity since they are intimately tied to appropriate target fixation. The random-dot stimuli were shown with shutter glasses that alternately displayed the stimuli to each eye. In addition, the researchers eliminated many monocular cues to depth perception such as texture and lighting. They found that there was

no difference in near and far stereoacuity for natural viewing. However, in a condition that mimics HMD accommodation and convergence mismatch, the researchers found that stereoacuity was significantly diminished.

The authors further speculated that the mismatch between the observers' knowledge of the object's position and the actual position also degraded stereoacuity. The authors concluded that stereoacuity of VE users should be measured to determine which users may be negatively affected by VE exposure. The authors were skeptical that a random-dot stereogram based stereoacuity measurement could be accurately obtained given the limited resolution of the HMD display. The low resolution of the display would be equivalent to the low density stimuli tested by Banks, Gepshtein, and Landy (2004) and would not be an accurate measure. The proposed stereoacuity metric based upon the Howard-Dolman task may be a more appropriate metric for use with HMDs. We present this assessment in the next section.

4.5.2 Types of Stereoacuity Assessments

Establishing metrics for measuring the depth discrimination threshold have both clinical and practical importance. Saladin (1998) contends that stereopsis is the single best indicator for overall function of the sensory and motoric aspects of the visual system (p. 766). Within the workplace, stereopsis of employees is routinely monitored in such occupations as hazardous waste management, military, and aviation (National Research Council, 1987). Clinical test for stereoacuity can be divided into two categories, real-depth tests and projected-depth tests.

<u>Real-Depth Tests</u>

Within this test paradigm, a real test object is moved in and out of the plane of one or more target objects (Reading, 1983). The metric of the task depends upon the type of

psychophysical assessment employed. For example, when using the method of adjustment

technique, the observer positions a moveable object offset from a fixed target to be aligned with

a fixed target. The resulting measurement is the difference in distance between the test object and

the target (target – nominal), or point of subjective equality (PSD). The PSD value is then used

to estimate the stereoacuity for that observer. The Howard-Dolman two-peg test is a classic

example of a real-depth test.

When performing the Howard-Dolman two-peg test, the participant views two rods seen

against a high contrast background. One rod moves while the other rod remains stationary or

fixed. The task is to move the rod either to the first noticeable offset (e.g., JND) from the

stationary rod or to the depth where the two rods appear at equal depth (e.g., PSD). The Howard-

Dolman apparatus shown in Figure 10 is one example of a Howard-Dolman testing instrument.

Figure 12 Howard-Dolman apparatus
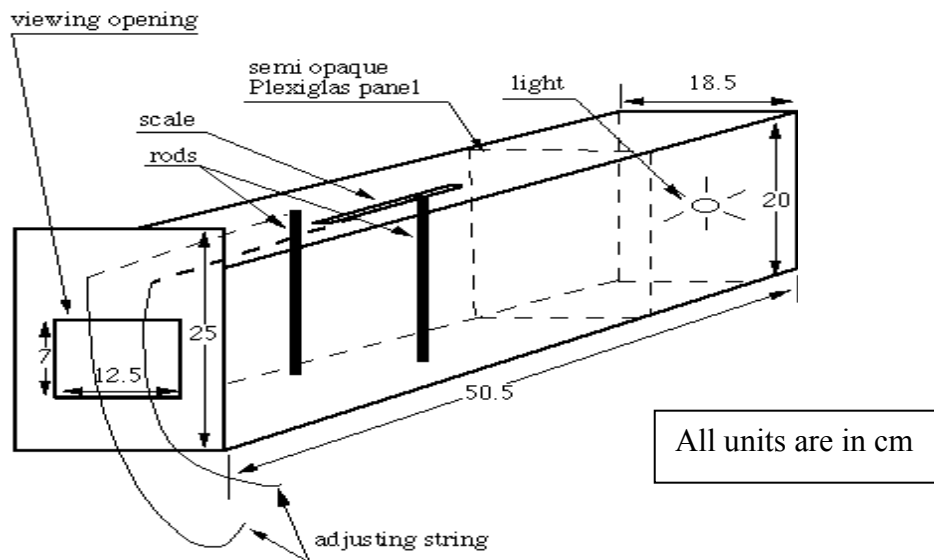
The Howard-Dolman apparatus completely encloses the two-rod mechanism in order to

maintain uniform lighting across the two rods. Additionally, the participants view the rods

through a 7 cm x 12.5 cm aperture to reduce other cues in the environment that may assist in

performing the task. The front face of the Howard-Dolman apparatus with 120 lux illumination

is shown in Figure 13 a. The typical viewing distances for the task are between 4 to 6 meters

(Saladin, 1998, p. 743). The participant adjusts the rods by a string-and-pulley mechanism as

pictured in Figure 13 b. The scale shown in Figure 12 is graduated in centimeters and the pointer

indicates the disparity between the moveable rod and the target.

Figure 13 (a.) Howard-Dolman rods and pulley system with lighting (b.) Participant performing
Howard-Dolman task.

a.   b.

Projected-Depth Tests

Stereograms, such as those used by Wong, Woods, & Peli (2002), are an example of a

projected-depth test. Stereograms require stimuli to be presented through a stereoscopic device

such as liquid crystal shutter glasses. Random-dot stereograms are computer generated left and

right eye views that are time sequenced to appear independently at a rate that the human system

cannot detect (Burdea & Coiffet, 2003). The result is a 3D view of the virtual scene. Other

stereograms employ vectography, a 3D photographic process of printing stereo images onto

polarized material (Reading, 1983, p. 180).

The Titmus Stereo Test shown in Figure 14 is a more standard projected-depth test. There are three types of stereograms represented: the Stereo Fly test, the Circle test, and the Animal Stereo test. The stereograms range in disparity. For example, the Circle test has 9 diamonds with four circles inside. One of the circles in the set of four has a disparity which could be singularly valued from 40 to 400 arcsec. The Titmus Stereo test requires the observer to wear glasses with polarizing filters. The stereograms are typically viewed from approximated arm's length or within 40 cm. The participant reports which circle is out of the plane in relation to the other three.

Figure 14 Titmus stereo test, Photo from Bernell, http://www.bernell.com/
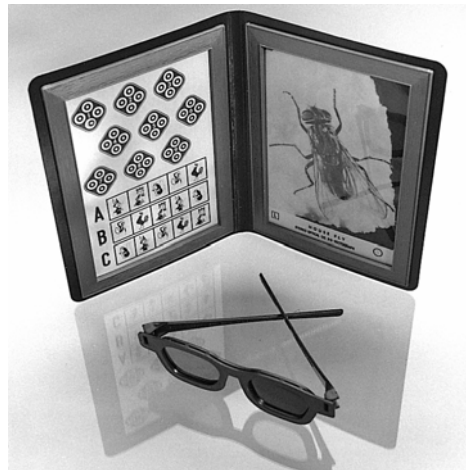
Potential Confounds to Stereo Tests: Depth Cues

Thus far, we have been describing stereopsis as a foundational perceptual aspect of human depth perception. However, the degree of binocular disparity provides more precise information about an object's distance from the observer or from other objects in the visual scene

(Wickens & Hollands, 2000, p. 141). Although distance perception is a very powerful indicator of depth, to accurately perceive depth other cues from the environment are needed.

Depth cues are typically classified as primary, those cues which are physiological, and secondary, those which are learned. Primary depth cues include convergence, accommodation, and binocular disparity. Primary depth cues provide depth information over short distances since beyond 6 meters there is no convergence. Artists typically employ secondary depth cues to create depth in drawings or paintings. These pictorial cues include static information such as an objects size, interposition or occlusion, lighting and shadow, and linear perspective. There are also kinetic depth cues such as motion parallax where far objects appear to move slower and in the same direction as the moving observer while closer objects appear to move faster and in the opposite direction. Because many secondary depth cues may be present in the testing environment, these cues should be controlled for when measuring stereoacuity. More specifically, the cue of binocular disparity should be the only depth cue present in the testing environment.

The Howard-Dolman apparatus was adapted from Howard (1919) who recognized that the vision tests for fighter pilot in WWI were confounded by extraneous depth cues. The original Howard Depth test pictured in Figure 13 presented the observer with a view of two rods, each 1 cm in diameter and 26 cm in length, through a screen and silhouetted against a background of uniform illumination. The head was fixed and the person's eyes were centered with respect to the aperture window. The testing distance was 6 m. In addition the rods were placed such that only the center portions of the rods were visible.

Figure 15 Howard depth test (Howard, 1919)
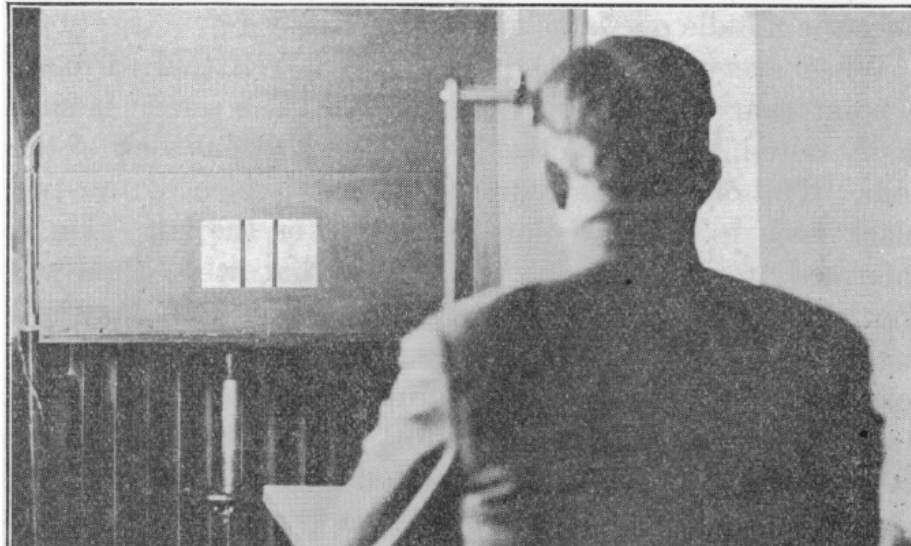
The experimenter predetermined the disparity level and the participant responded by stating which rod, left or right, was closer to him. This procedure was repeated 20 times per disparity level: thus, the Howard Depth test employed the method of constant stimuli. Stereoacuity was measured as the lowest disparity level where the participant could correctly identify the object as nearer or farther to him 75% of the time.

Stereoacuity measured over 106 pilots within Howard's study varied from 1.8 to 132.2 arc seconds. To better understand individual differences among the pilots, Howard ranked them based upon their stereoacuity levels and then systematically evaluated all vision tests performed in conjunction with the depth perception test (e.g., resolution visual acuity). Howard found that IPD and resolution visual acuity had an affect on the stereoacuity for that observer. More specifically, pilots with a larger IPD and better visual acuity achieved stereoacuity levels less than 8 arc seconds. This result is an example of how individual differences may affect user performance on visual assessments. Thus, it is important to collect multiple metrics to assist in

explaining these differences. Passing stereoacuity for pilots was assessed as a 3 cm averaged error as measured from the viewing distance of 6 meters. The drawback to the Howard Depth test was the duration of the testing time, which made it impractical as a clinical test. The Howard-Dolman apparatus was created to decrease testing time.

There are several secondary depth cues available to the observer when performing the Howard-Dolman Depth test such as motion parallax due to head motion. The depth cue of apparent size or the change in size of the proximal stimulus on the retina with changing distance, along with accommodation, is also present. Because of these confounds, many early depth perception researchers questioned the validity of the Howard-Dolman test in measuring stereoacuity (Warren, 1940). Others argued that the Howard-Dolman task was relevant for assessing stereoacuity for tasks which predominantly required stereopsis with the support of minor primary and secondary depth cues (Sloane & Gallagher, 1945; Sloan & Atlman, 1954). More recently, Larson (1985) suggested a return to the original Howard Depth task with modifications to support clinical use. Projected-depth tests are becoming more clinically dominant since they provide a level of control over monocular depth cues while lessening test duration.

Of the projected-depth tests mentioned only the random-dot stereogram can eliminate all other depth cues and provide the most accurate measure of stereopsis (Howard & Rogers, 1995). However, as Wong et al. (2002) point out, the use of random-dot stereograms with HMDs is not currently possible. In addition, while assessing different types of HMDs, introducing shutter glasses may not be a prudent experimental design decision. The modified virtual Howard-Dolman task (V-HD task) developed by Rolland, Gibson, & Ariely (1995) and later improved upon by Rolland, Meyer, Arthur, & Rinalducci (2002) qualifies in general as a projected-depth

test. At this time, the assessment provides the best metric for measuring stereoacuity with regard to VE system assessment.

4.5.3 Modified Virtual Howard-Dolman Depth Test for Augmented Reality

The V-HD task controls for the familiar size cues by presenting generic objects which have no real world correlation (Rolland, Gibson, & Ariely, 1995). Thus, there is no expectation of size when simultaneously viewing both objects. However, aspects of the graphics such as lighting may provide weak cues to depth. Motion parallax was controlled for by using a chinrest as Howard (1919) had done. A chinrest could be employed because the HMD used was a bench prototype. The HMDs worn in the forthcoming experiments were functioning prototypes or off-the-shelf products. Given the weight of the HMDs and their respective cabling systems, it was not appropriate to restrain the head from moving. Also, we were interested in the errors generated when the observer was allowed some free movement of their head.

Secondary depth cue systems, accommodation and the convergence, are stimulated unnaturally while viewing collimated images (Roscoe, 1991). Rolland et al (2002) adjusted for conflicts between accommodation and convergence mechanisms by placing the monocular optical images at the target or nominal depth where the 3D virtual objects were rendered. In the experiments discussed in Chapter 7, three types of HMDs were tested: optical see-through, video see-through, and fully immersive. Because the optical see-through display was built in the ODALab, we had full control over setting the depth plane for the microdisplay. Even though the HMPD-1 was designed to project a collimated image, we were able to adjust the settings of the optics to project the 3D test object to match the graphically rendered objects distance. This adjustment could not be done with off-the-shelf HMDs. Further, the second generation HMPD

(HMPD-2) worn in the experiments presented in Chapter 8 was specifically optimized to project images at the same distance as the rendered objects. Properties of each HMD will be discussed further in the next chapter,

To account for the potential differences in accommodation and convergence when performing tasks while wearing each respective HMD, several vision tests were added as part of the procedure. More specifically, prior to testing participants were screened using the Titmus Stereo Test. All three tests were used to confirm that the participants' stereoacuity was at least 40 arc seconds prior to the start of the experiment. Further, the Howard-Dolman peg test using the Howard Dolman apparatus was performed before and after VE testing to monitor changes in visual performance over the course of the experiment. In addition, dark vergence and dark focus were measured before and after the participant performed the V-HD task to assess accommodation and convergence changes that may result from the VE exposure. These procedures are further detailed in Chapter 7.

Virtual Howard-Dolman Stimuli

Rolland et al. (2002) showed that for the HMPD-1 system the optimal test stimuli for determining stereoacuity was a small cube or octahedron of the same volume (diameter equal to 4 degrees of visual angle) presented with a long thin cylinder (diameter equal to 1 degree of visual angle). An example of the stimuli presented to the right eye is shown in Figure 16. In each experiment, the octahedron was the moveable object and the cylinder remained fixed at the nominal depth. However, the nominal depth changed according to the lens setting of the HMPD.

74

Figure 16 Virtual Howard-Dolman task octahedron and cylinder stimuli
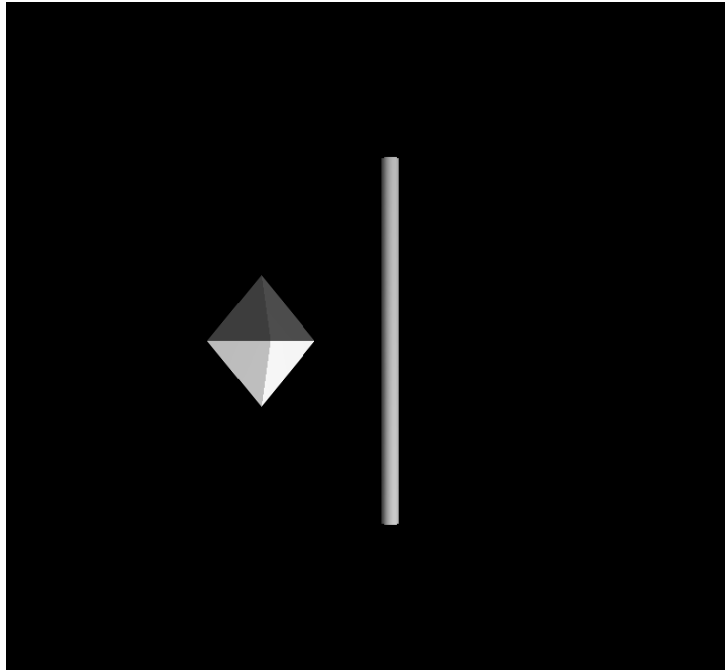
The change in viewing depth also changed the size of the virtual objects as well as the

IPD of the observer. The software had to account for the changes in object size with respect to

the variable IPD over the test viewing depths. Besides setting the lighting parameters, the most

difficult aspects of designing the test were choosing the correct input device and the optimal

resolution for the input device with which to move the octahedron. To establish the appropriate

resolution for the input device, the resolvable depth for the HMD was calculated. The resolvable

depth also establishes a predictive metric for determining the precision (variable error) of the

HMD at each of the respective viewing distances.

The depth of focus (DOF) is the regional extent about the image plane where the

projected image remains clearly focused without adjusting optical power. The depth of field

represents the DOF in object space. There is also a range of distances in front and in back of the

image plane where a single pixel size is just resolvable. This distance is considered the

resolvable depth for the HMD. Rolland, Krueger, & Goon (2000) give the equations for

determining resolvable depth for an HMD. We review a simplified version of the calculations.

Figure 17 depicts the geometry overview for determining DOF. The distance between

points a and b represents the horizontal FOV for each eye across the projected image plane. The

distance D is the distance from the image plane to the pupil of the eye. Three pixels are

represented in the center of the binocular overlap region.

Figure 17 Geometrical overview for determining resolvable depth
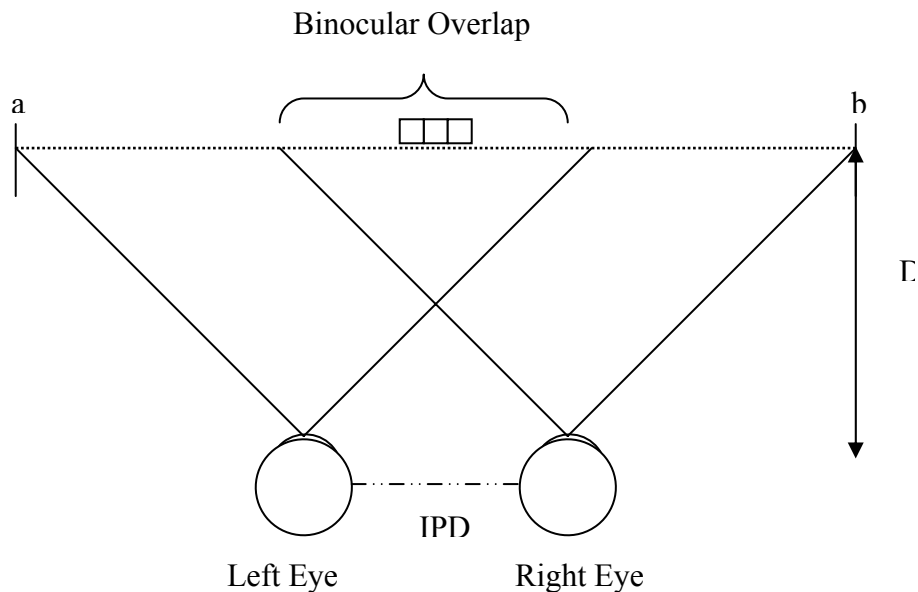
Figure 18 shows a simplified version of Figure 17 for the left eye, where L is the length

of the image plane and θ is the FOV for the left eye. For the octahedron to remain in focus when

it is moved off the image plane, its position cannot exceed the limits imposed by the DOF for the

HMD, which is a function of the numerical aperture (NA) of the optics (NA equal n' sinα' with

n'=1 in air and α' and the magnification). The magnification defines the size of a pixel projected

onto the image plane at each respective viewing distance. As presented in the previous section,

the resolution of an HMD is equivalent to the size of a pixel in arc minutes after being magnified

by the optics. The size of a pixel in pixels per degrees is given by equation 7, where M is the

horizontal FOV in degrees.

Figure 18 Calculating the size of a pixel
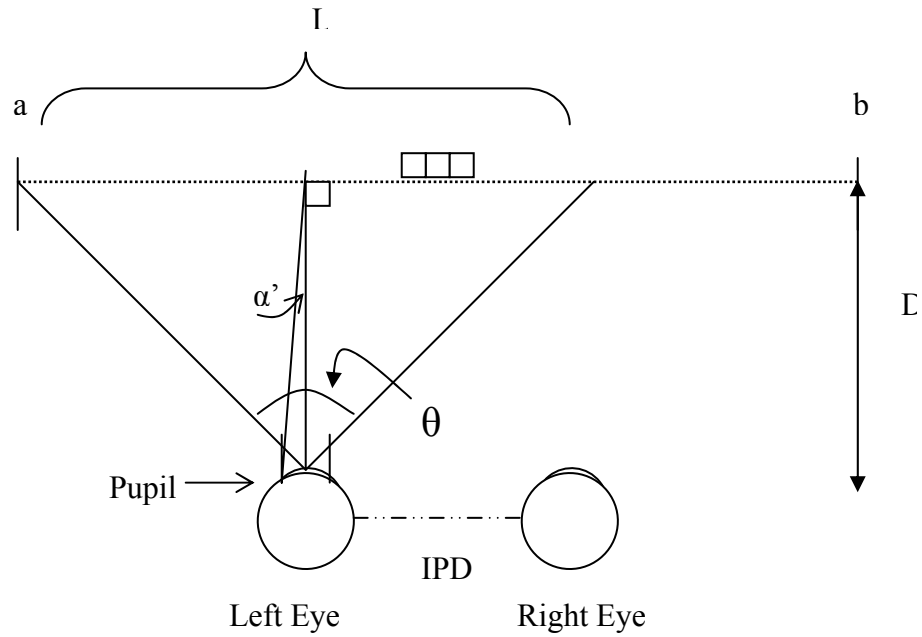
$$FOV = \theta \tag{5}$$

$$Tan\theta / 2 = \frac{L}{2D} \tag{6}$$

$$Pixel\ Size\ (\rho) = \frac{L}{M}, \text{ with microdisplay of M x N pixels} \tag{7}$$

To determine the resolvable depth for the HMPD, the distal (D_2) and proximal (D_1) positions, with respect to the fixation plane, were calculated as shown in equations 8-12. The depictions for D_1 and D_2 are shown in Figures 19 and 20, respectively. Both projections used to define D_1 and D_2 pass through the centers of adjacent pixels. This pixel spacing is taken as the limiting resolution of the microdisplay. For square pixels, this distance should equal one pixel.

Figure 19 Calculating D_1

$$Tan\ \beta = \frac{\frac{IPD}{2} - \rho}{D} \tag{8}$$

$$D_1 = \frac{\frac{IPD}{2}}{Tan\ \beta} \tag{9}$$

78
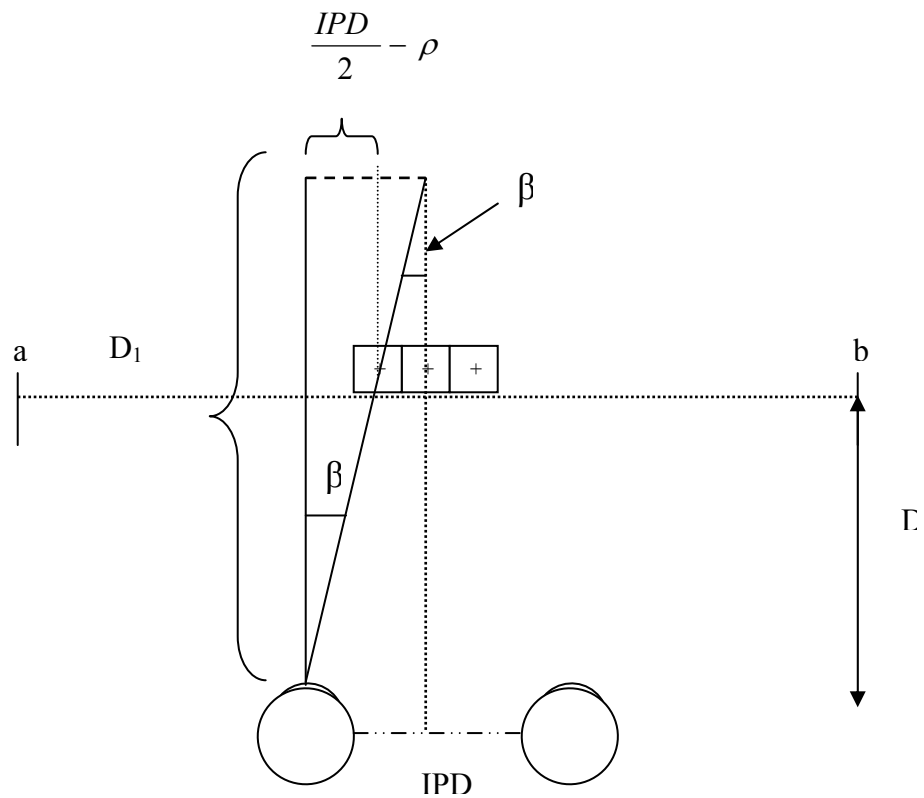
Figure 20 Calculating D$_2$

$$Tan\ \gamma = \frac{\dfrac{IPD}{2} + \rho}{D} \qquad (10)$$
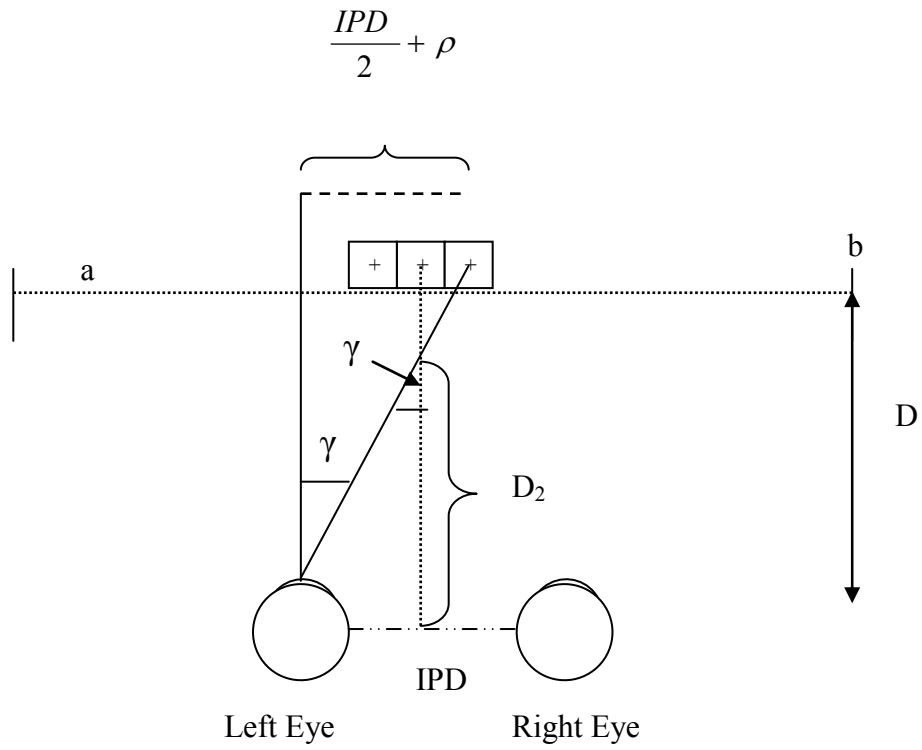
$$D_2 = \frac{\dfrac{IPD}{2}}{Tan\ \gamma} \qquad (11)$$

$$\Delta D = D_1 - D_2 \qquad (12)$$

As an example, the resulting ΔD values in mm for the second generation HMPD (HMPD-2) are given in Table 6. These values pose the theoretical limits of stereoacuity (as computed with VE) for the user while performing tasks while wearing this display. Thus, we expect average variable errors to be within 1.50 to 14.98 mm for a viewing distance of 800 mm, based on pixel or sub-pixel resolution, 5.33 to 53.30 mm for the 1500 mm viewing distance, and 21.45 to 214.79 mm for the 3000 mm viewing distance. Resolution values for the input device and the results of the study are detailed in Chapter 8.

Table 6 Predicted stereoacuity and resolvable depth in mm for the HMPD-2 as a function of pixel size and viewing depth.

Pixel Size	Depth (mm)	Resolvable Depth (mm)
1	800	14.98
	1500	53.30
	3000	214.79
1/2	800	7.49
	1500	26.64
	3000	107.29
1/4	800	3.74
	1500	13.32
	3000	53.63
1/10	800	1.50
	1500	5.33
	3000	21.45

Predicting Measures of Stereoacuity, Resolvable Depth, and Other Response Variables

Stereopsis is not based on absolute disparity or the distance of a single object from the egocenter of the observer (Steinman, Steinman, & Garzia, 2000, p. 187). Absolute disparity is the cue for vergence and guides eye movements toward the to-be-fixated object (Howard &

Rogers, 1995). Thus, the Howard-Dolman Depth test measures relative distance, which is a comparison of the distance between two or more objects in the visual field and is the cue for stereopsis. In essence, the observer can discriminate objects that are closer from those that are farther away relative to some fixation point based upon relative disparities encoded at the level of the retina, yet processed in the visual cortex.

Cumming and Parker (1999) showed that absolute disparities are processed by area V1 in the visual cortex of a Macaque monkey. As discussed previously, the cortical locus of stereopsis is still under investigation. Understanding the visual processing stream, allows for better interpretation of experimental results since, in this instance, we are answering an applied and not basic research question. As Larson (1985) pointed out, there are multiple methods of analyzing Howard-Dolman type data suggesting that the test itself is flawed. However, given that the test was developed before knowledge of binocular disparity neurons were confirmed, allowing the testing methodology to evolve is reasonable if the test itself is reliable or consistent over measurements.

Weymouth & Hirsch (1945) reported on the reliabilities of different depth tests, which included both the Howard Depth test and the Howard-Dolman test. The test-retest correlation for the Howard Depth test was .88, while the correlation for the Howard-Dolman test was .62. For the Howard-Dolman task, the authors reassessed the reliability after changing the scoring procedure from assessing absolute error (AE) calculated by equation 13 to that of calculating the root mean square error shown (RMSE) in equation 14. The result was to increase reliability of the metric to .86.

$$\Delta d_{AE} = \frac{\Sigma |X|}{N} \tag{13}$$

$$\Delta d_{ERMS} = \sqrt{\frac{\Sigma X^2}{N}} \tag{14}$$

Given that the response variable is the distance error in mm between the two rods, equations 13 and 14 can be written as:

$$AE = \frac{\sum_{i=1}^{n} |\chi - T|}{n} \tag{15}$$

$$RMSE = \sqrt{\frac{\sum_{i=n}^{n} (\chi - T)^2}{n}} \tag{16}$$

Where χ is the participants' response, T is the fixed distance of the target rod, and N is the number of repetitions. Schutz & Roy (1973) defined both scores, AE and RMSE, as composite metrics of both constant and variable errors. Constant error (CE) is the participants' signed measurement error or bias in aligning the two rods. The mathematical expression for CE is given in equation 17. Variable error (VE_r) as shown in equation 18 is the measure of dispersion about the participants' own mean error ($\overline{X}$), thus it is a measure of precision.

$$CE = \frac{\sum_{i=1}^{n} (\chi - T)}{n} \tag{17}$$

$$VE_r = \sqrt{\frac{\sum_{i=n}^{n} (\chi_i - \overline{X})^2}{n}} \tag{18}$$

In a summary article on the proper use for the CE and VE_r metrics, Schutz (1977) recommended that the CE should not be used as a dependent measure in inferential statistics tests

such as t-tests. Instead, the absolute value of the CE (|CE|) along with VE_r should be used as response variables to compare for between or within group treatment effects. Howard & Rogers (1995) term the CE with regard to the depth discrimination threshold as stereoaccuracy (p. 150). Ogle (1962) cautioned against using the CE to assess stereoacuity since individual differences in refraction errors or even ocular dominance may skew these results. In addition to evaluating the VE, Ogle recommended calculating the equivalence disparity angle (η) as a means to assess stereoacuity from Howard-Dolman like depth tests.

Calculating the equivalence disparity angle (η)

Binocular correspondence theory suggests that there is a mapping of the binocular area of the visual field onto each retina of the eye (Reading, 1983). As discussed and shown in Figure 11, the retina of each eye maps topographically to the lateral geniculate nucleus and follows a common pathway to the visual cortex. Along the pathway the monocular images are fused creating a single percept of the objects in the visual field. Small horizontal retinal disparities created by non corresponding retinal points give rise to stereopsis (Patterson, 1992).

Dolman (1919) designed his depth test to measure angular deviations in depth as measured from the point of fixation. At the observer's fixation point, there is zero disparity (Steinman, Steinman, & Garzia, 2000). Consequently, all objects equidistant with the point of fixation are perceived as being at the same depth. That is to say that the relative distance between objects appears to be the same. The horopter is a theoretical construct that represents a mapping of objects in the visual field with zero disparities or of corresponding points along the retina (Patterson, 1992). This concept is instructive since it provides a geometrical method for predicting the limits of stereopsis.

83

For the Howard-Dolman task the fixation target sets the horopter for that particular viewing depth. The moveable rod can either appear in front of or behind the point of fixation, (e.g., as crossed or uncrossed disparities, respectively). Figure 21 shows the geometry for determining the equivalence disparity angle (η). Within this framework, the angle Φ is the angle of convergence for the far object. Similarly, the angle ε represents the angle of convergence for the near object.

Figure 21 Geometry for determining the equivalence disparity angle (η)
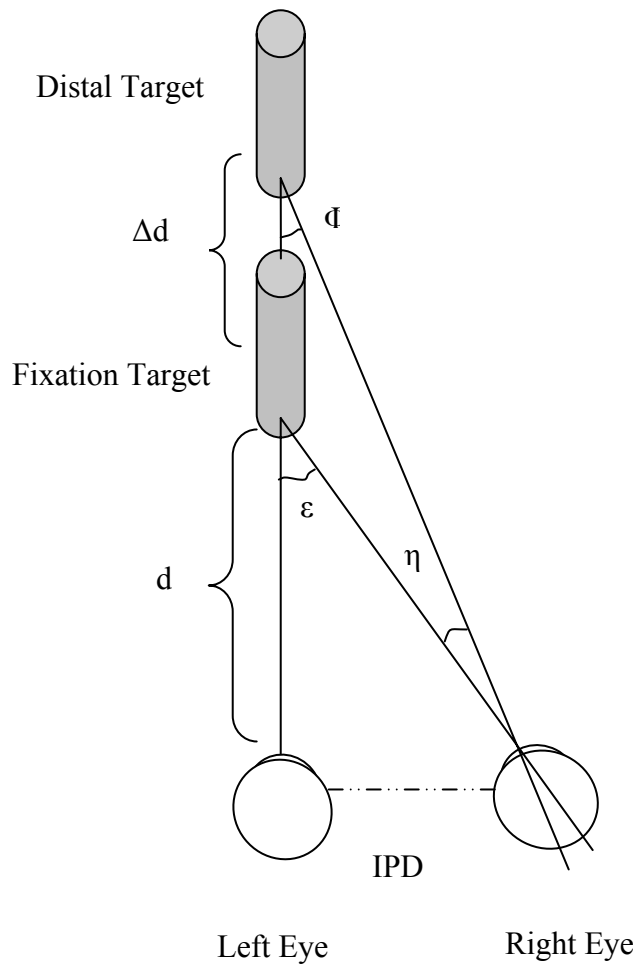
$$\varepsilon - \phi \approx \frac{IPD\ \Delta d}{2d^2}, \qquad (19)$$

and,

$$\text{angle}\,(\varepsilon - \phi) = \eta \qquad\qquad (20)$$

Since ε, Φ, and η are small, then the sine, tangent, and angle in radians were taken to be equal. If the IPD and distance b are in the same units, then the angle can be converted to radians. The estimate for η in arc seconds by employing a conversion factor from radian to arc seconds becomes

$$\eta = \frac{IPD\Delta b}{b^2}(206265) \qquad\qquad (21)$$

In the forthcoming experiments, we calculated CE, |CE|, VE, and η as response variables from both the Howard-Dolman and modified Virtual Howard-Dolman task. In the next chapter we describe the HMD used in each experiment and calibration procedures for each type of display.

CHAPTER 5 CALIBRATION PROCEDURES FOR HMPD AND VE SYSTEM SETUP

5.1 Introduction

Many off-the-shelf HMDs are improperly or inadequately documented. When reading technical specifications many important parameters of the display such as the microdisplay active area size is either not given or only partially specified. At the very least, specifications for any type of HMD should include: FOV for the x and y dimensions, size of the microdisplay, active area of the display in both the x and y dimensions, and pixel size or the equivalent so that these parameters may be computed. These parameters are necessary not only for software coding purposes, but also for calibration procedures. HMD manufactures should also include information about the optics alignment (e.g., whether the optical axis is convergent or divergent) to ensure proper use of the displays.

The purpose of this chapter is not to review the field of calibration for HMDs, but to report on the calibration methods that we employed to ensure that test stimuli for both the modified Landolt C and the Virtual Howard-Dolman test stimuli were properly viewed. We also discuss the HMDs worn during the experiments and the VE system setup.

5.2 Overview of HMDs

Four HMDs were used during various experiments reported on in later chapters. Two HMDs were first and second generation prototype head-mounted projection displays (HMPDs) whose optics was developed in the ODALab at the College of Optics and Photonics at the

University of Central Florida (Rolland et al., 2004). The mechanical alignment of the ODALab

HMPDs was performed by industrial partners. We also included an immersive display, VR6

(Virtual Research Systems), and a video see-through HMD, the Canon COASTAR™. The

HMDs are shown in Figure 22. Table 9 shows the parameters for each display as they are

presented in their documentation.

Figure 22 The upper left panel shows the VR6. The upper right panel shows the Canon Coastar. The bottom two panels show the ODALab prototype displays, on the left the HMPD-1 and on the right the HMPD-2, respectively.

Table 7 Technology specifications for HMDs used in experiments.

HMD Type	Display Type	FOV (Degree)		Resolution (Pixels)		Display Size mm	Focus Plane mm	IPD mm
		H	V	H	V			
HMPD-1	LCD	41	31	640	480	27 x 20	Infinity	55-75
800		40	31				800	
1500		41	31				1500	
3000		42	32				3000	
HMPD-2								
800	OLED	33	25	800	600	12 x 9 mm	800	55-72
1500		34	26				1500	
3000		34	26				3000	
Canon	LCD	57	37	640	480		2000	63
VR6	LCD	48	36	640	480	1.3 x 2.59	914	52-74

Both HMPDs have adjustable focal planes; however, only the optics of the HMPD-2 was designed and optimized for projecting images to the depths of 800, 1500, and 3000 mm. The adjustment for changing the depth plane for either HMPD was a manual adjustment. However, for the HMPD-1 we had to open the casing to adjust the optics each time we changed viewing depths. The adjustable focal plane aide in eliminating the accommodation and convergence decoupling that occurs in most HMDs. For example, the focal plane for the VR6 is set at 914 mm, which is approximately 3 feet. Most imagery for this HMD will be set beyond this depth. Because of the limited adjustability of the Canon and the VR6, we only tested these HMDs at the 1500 mm viewing distance.

Table 10 gives the resolvable depth predictions for each HMD at the 1500 mm viewing distance. As discussed in the previous chapter, the resolvable depth allows us to make predictions on user performance when aligning the octahedron with the cylinder at the 1500 mm viewing distance. More specifically, at each pixel and subpixel resolution we can predict the

user's precision or variable error per HMD. For example, the ability for a user to discriminate front and back positions of the octahedron with respect to the cylinder should be compromised in both the Canon and the VR6. More specifically, the step sizes between each subpixel resolution are larger than that of the HMPD-1. Thus, the participants should have more variability in there responses because of the decreased discriminability of front and back positions. However, other aspects of the viewing environment such as positioning of the stimuli with respect to the focal plane of the HMD may improve performance.

Table 8 Predicted resolvable depth in mm for the HMPD-1, Canon, and VR6 as a function of pixel size.

	Resolvable Depth for 1500 mm (mm)		
Pixel Size	Canon	VR6	HMPD-1
1	95.54	78.32	65.93
1/2	47.73	39.14	32.95
1/4	23.86	19.57	16.47
1/10	9.54	7.83	6.59

Calibrating the HMDs prior to performing the assessments is a necessary step to ensure accuracy of the testing method. The HMPD-1 had undergone engineering based calibration prior to the start of the experiment. The main issue with this display was overscan of the frame buffer as discussed in Chapter 4. The HMPD-2 was a true prototype display in that the mechanisms for adjusting the depth planes and IPD were still under design. This HMD posed new challenges when performing the experiments discussed in Chapter 7. The off-the-shelf HMD models were not easy to calibrate. The best we could do was to ensure that the software accounted for the FOV values listed in Table 9 when running the experiments which were included in their owner's manual. However, this condition is similar to real world experience where the HMD for

first time users is essentially a black box with poor documentation. Once the HMD is best aligned, software can be used for fine tuning the calibration. HMD calibration is an active area of research for which there is currently no off-the-shelf solution (Tuceryan & Navab, 2000; Owen, Zhou, Tang, & Xiao, 2004).

5.3 First Generation HMPD-1 and VE System

Hua, Girardot, Gao, & Rolland (2000) introduced the first miniaturization of the optics for a new type of new head-mounted projection display that reduced visual distortion and allowed for correct occlusion of virtual objects. The resulting prototype of this work is shown in Figure 20. Figure 21 illustrates the mechanism of how the HMPD functions.

The HMPD consists of a pair of small projection lenses, a beam splitter, a pair of micro displays, and retroreflective material strategically placed in the environment. The computer generated images are projected through the lens and retro-reflected back through the pupils of the eye. Through the aide of the beam splitter the exit pupil of the projection optics is coupled with the pupil so that the image can be seen. The beam splitter for both HMPDs is set at a 45 degree angle. The use of projection optics instead of eyepieces allows for a larger FOV and less optical distortion. The use of retroreflective material also distinguishes this type of technology from other projection based stereoscopic displays such as CAVES, Cave Automatic Virtual Environment (Cruz-Neira, Sadin, DeFanti, 1993).

Unlike the CAVE, which uses multiple back projection screens and projectors to simulate a collaborative virtual workspace, the HMPD setup allows for independent viewpoints of the same space without cross talk from multiple users sharing a common viewpoint. In addition, the

retroreflective material allows for correct occlusion when observing both real and virtual objects (Hua, Girardot, Gao, & Rolland, 2000).

Projection Lens

Retroreflective Screen

Projected Image

Eye

F'

Beam Splitter

Intermediary Image
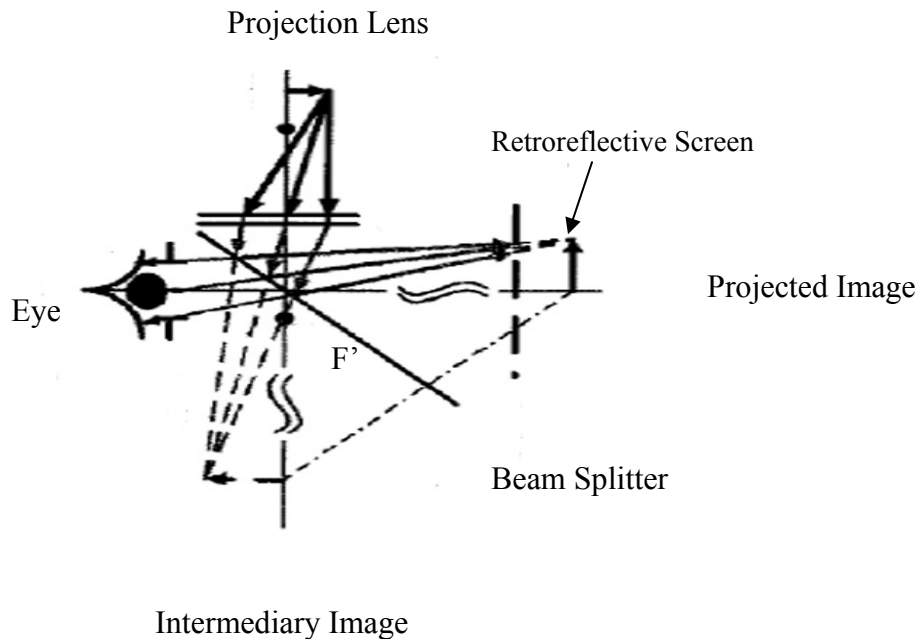
Figure 23 HMPD imaging pathway. Adapted from Hua, Gao, & Rolland (2000)

5.3.1 Retroreflective Material and the Augmented Reality Center (ARC)

The main property of the retro-reflective material is that light rays hitting its surface at any angle are reflected back onto themselves. Because of this property, the image features such as shape and location are independent of retroreflective screen placement. Currently, there are two types of retroreflective material used in the ODALab VE system, Scotchlite[TM] 3M Fabric Silver (micro-bead array) and Scotchlite[TM] 3M Film Silver (corner-cubed array).

Each type of retroreflective material has a different underlying array structure (i.e., micro-beaded or corner-cubes) as shown in Figure 22. The beaded material is an off-the-shelf

material that acts through specular reflection, while the cubed material was custom specified and utilizes total internal reflection (TIR). A simple definition of TIR is that incoming light is bent in such a manner upon hitting the surface of the material that the entire ray is sent back into the direction from which it came. Both materials are optimized for safety applications, not imaging. However, an optical structure operating on TIR properties is expected to yield higher reflectance, thus appears brighter. The point-spread functions for each material confirm that the cubed material achieved not only higher levels of brightness across the image plane, but also higher resolution (Martins & Rolland, 2003).

Figure 24 Array structures of the retro-reflective materials. From Capdevielle, Martins, & Rolland (2003).
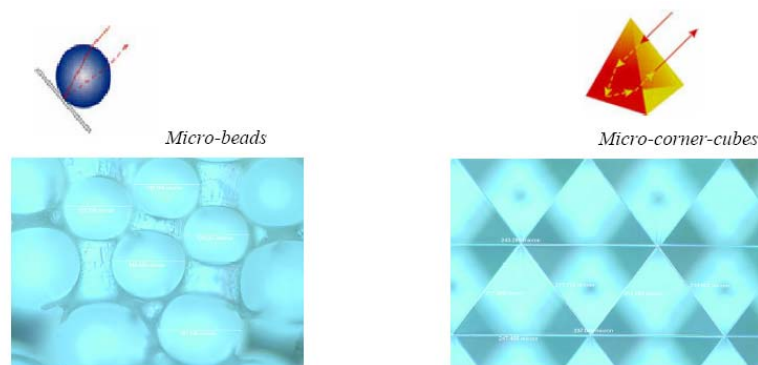
The retroreflective material, either type, aligns the Augmented Reality Visualization Center (ARC) shown in Figure 25a. The ARC is a prototype multi-modal Augmented Reality System with 3D visual, 3D audio and haptic capabilities configured as a quasi-circle of 4.57 m (15 ft.) diameter (Davis, Rolland, Hamza-Lup, & Ha, 2003). The first four experiments reported were completed in the ARC. Due to a change in research venues a modified version of the ARC also shown in Figure 25b was used for the stereoacuity assessment of the HMPD-2. Thus, one aspect of this research was to ensure test replicability between research venues, which has implications for performing the vision tests with other VE systems.
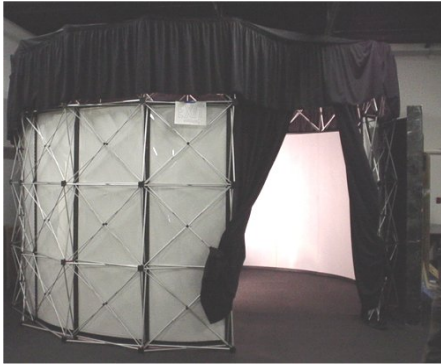
a.   b.

Figure 25 (a) The ARC   (b) The quarter ARC

3.3.2 Calibrating the HMPD-1

As shown in Figure 26, the control box for the HMPD drives the microdisplay. Its size necessitates larger cables, thus limiting mobility of the user. Because this is also a prototype version, there is no standard to maintain consistency between control boxes. However, as with any controller for microdisplays there are errors in the electronics as part of the trade-off for miniaturizing these devices. Although one control box is needed to drive the left and right microdisplays, each display and each control box needed to be tested separately.

93
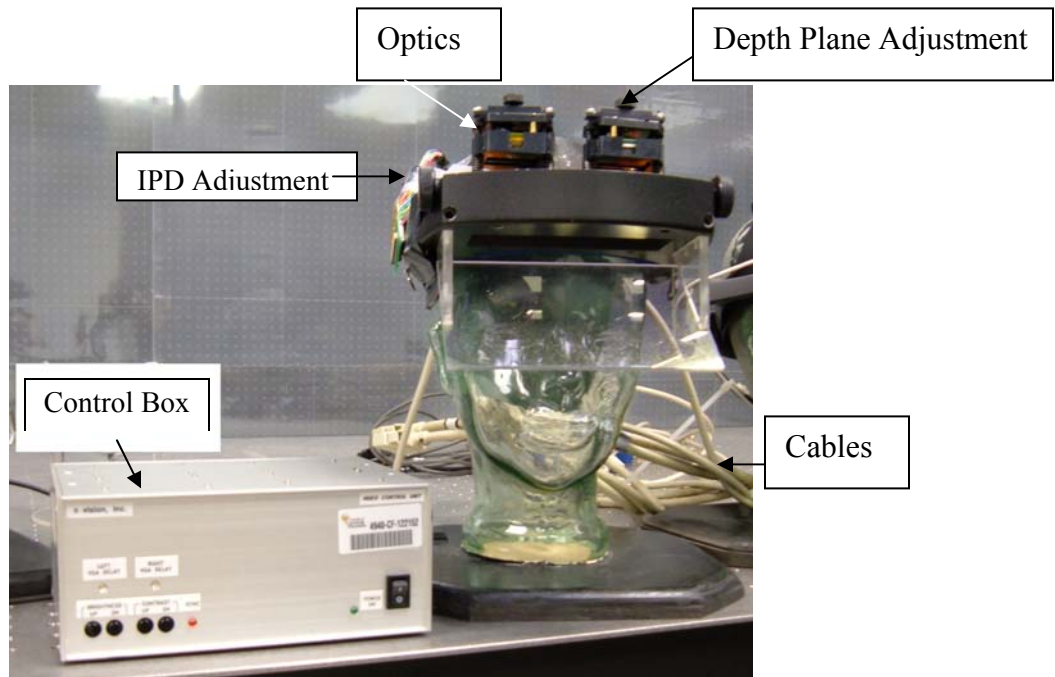
Figure 26 HMPD-1 system components

As previously described, the FBO is an error generated by the display device controller whereby some parts of the active area of the screen are not displayed. To measure the FBO for each control box we used the functionality of OpenGL, low level Application Programming Interface (API) standard in VE design. Figure 27 shows the OpenGL visualization pipeline. At each level within the pipeline, OpenGL offers different functions to control each aspect of the imaging process. At the level of the frame buffer graphical operations diagramed are confined to the viewable area of the buffer using the glViewport function. The function allows us to determine the displayable area accounting for the FBO.
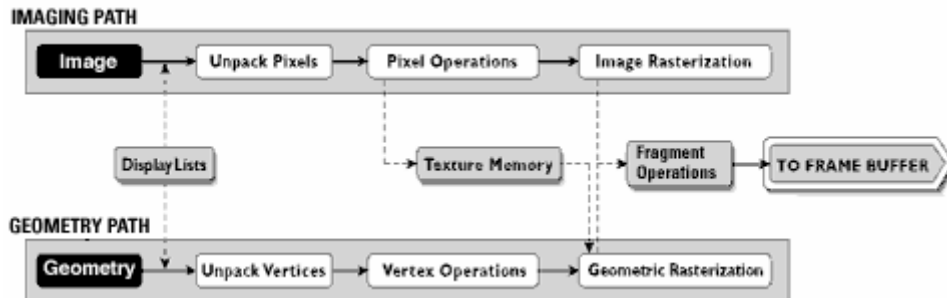
Figure 27 OpenGL visualization pathway from OpenGL, http://www.opengl.org

Figure 28 shows the relationship between the graphics frame buffer and the viewable area of the display along with a screen shot of the FBO application. Table 9 describes the parameters used in the calculation. To obtain this area, the X and Y coordinates of the viewport must be located. The outer edge of the frame buffer is the starting point for the measurement. Finding this point on the X axis is the initial reference point. This initial point determines the initial Y reference point. Next, the coordinates that define the pixels that first come into view in the active viewing area are obtained. The lines shown in the application are one pixel width, thus, moving line over one space covers the range of the field of view one pixel width at a time. The final X and Y positions are subtracted from the initial X and Y positions to determine the number of pixel lines not drawn to the active area of the display. The pixel positions as defined in Table 9 are outputted to the screen. This measurement must be obtained for both the left and the right viewports. Then, the new field of view for the virtual reality system is then recalculated to coincide with the FBO.
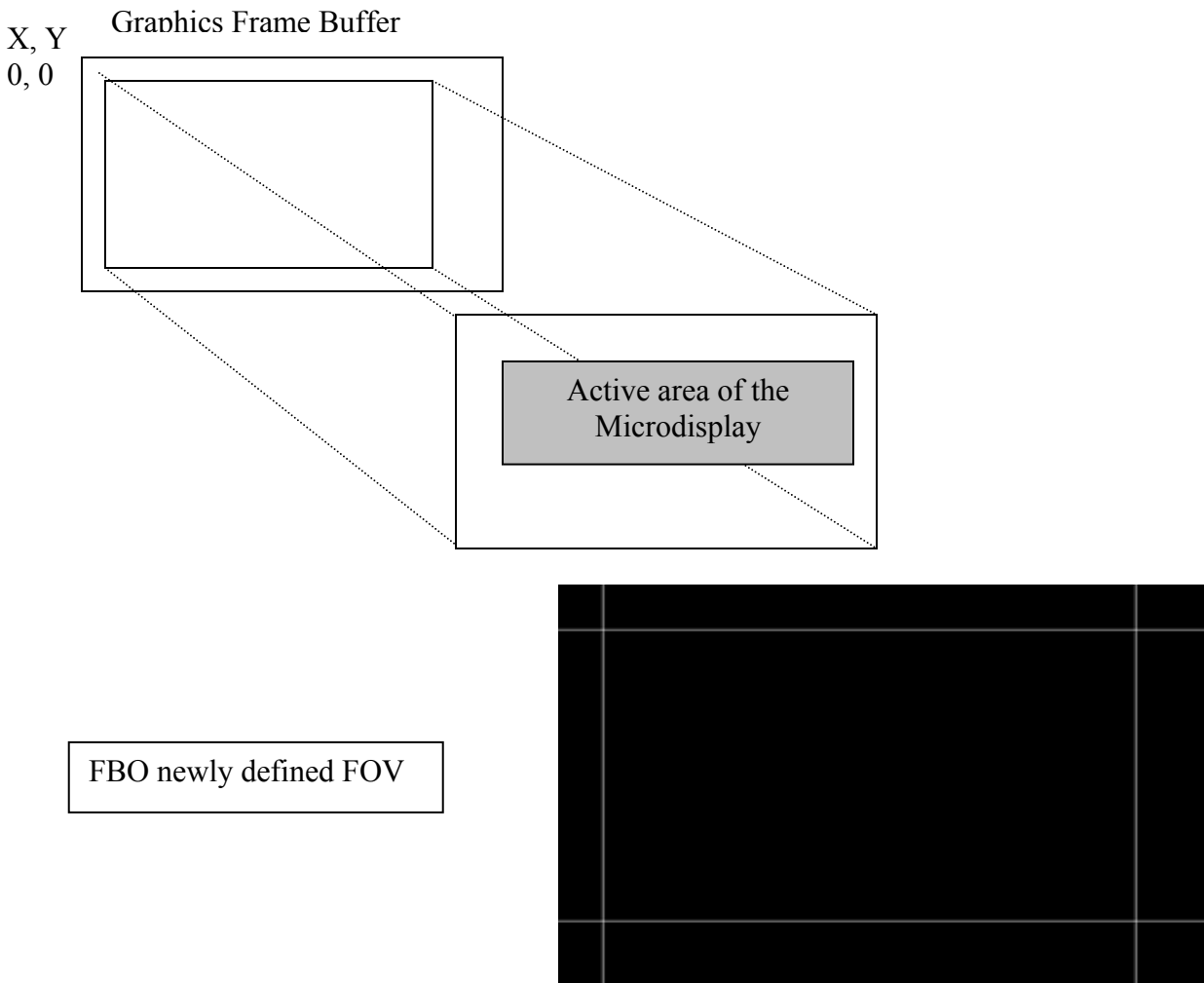
Figure 28 Frame buffer overscan measurement procedure

Table 9 Parameter calculations for FBO[2]

void glViewport (GLint X_i, GLint Y_i , GLsizei width, GLsizei height)
Parameters
X_i = the point where the buffer is first transcribed on the X-axis
X_f = the final point that the buffer transcribed into the buffer on the X-axis
Y_i = the point where the buffer is first transcribed on the Y-axis
Y_f = the final point that the buffer transcribed into the buffer on the Y-axis
Displayable Width in pixels = $(X_f - X_i) + 1$
Displayable Height in pixels = $(Y_f - Y_i) + 1$

---

[2] All code for the FBO measurement was created by Kit Fuhrman of the DARE ODALab coding team.

Table 10 Results of the FBO procedure

| | | X_i | X_f | Y_i | Y_f | FBO Values | |
						$(Y_f - Y_i)$ + 1	$(X_f - X_i)$ + 1
Box 263	Right	1	631	0	472	473	631
	Left	2	639	1	479	479	639
Box 264	Right	0	639	0	473	474	640
	Left	1	639	0	474	475	639

As shown in Table 12 the FBO for each eye measured from the same control box was different. In addition, the values for the left and right eyes changed when a new control box was attached to the system. More specifically, the FOV specified by the microdisplay is 640 x 480. When using control box 263, the FOV for the right eye was actually 631 x 473. In contrast, the values for the left eye were 639 x 479, which matches the documented FOV for the display. Since both microdisplays have a different center point for the screen, the information rendered to the display will not be exactly registered for each eye when viewed by the user. Although the HMPD was optimized for distortion, other HMDs do not have this feature. Compounded with distortion of the HMD lens system, errors in depth rendering caused by the FBO together with the alignment of the viewport centers on the optics, may result. However, as Robinett and Rolland (1992) pointed out, accounting for FBO in the VE software design is a relatively easy way to ensure that the HMD and the graphics work together to give the best possible imagery.

Performing the FBO procedure using the HMPD-1 was easier compared to that of the HMPD-2. In both cases the microdisplays were directly accessible and a separate calibrated monitor was not needed. However, the HMPD-2 uses an OLED as apposed to an LCD. Because of the burn-in rate of the OLED, the active area of the display randomly shifts to ensure that a single pixel is not constantly being turned on and off with use. This leads to a randomly moving

FBO. Given the nature of our task, taking an average of these FBO measures may not be appropriate since the shift can be between 1 and 5 pixel widths. How best to measure the FBO in some HMDs (e.g., VR6) is yet to be determined.

At the very least, the appropriate FOV values for both the X, and Y directions should be used to define the viewport of the microdisplay in the graphics. These values will allow the software to better estimate the size of a pixel for that HMD. Although this step is part of the OpenGL rendering specifications, other API's may only require the resolution of the display as an input variable. The VE designer should be aware of the possible visual errors that this higher level approach may induce. However, calibration of the HMD should ultimately not be the job of the end user. The HMD optomechanical assembly team is best able to measure the FBO of the microdisplay before assembly. Thus, before the microdisplay is mounted to the optics, the center of the viewport defined based on the FBO should be mapped to the central axis of the optics. With such alignment, the FOV is by construction symmetric with respect to the eyepoint. The Manuals would then report these parameters for optimal use of the device.

5.4 Second Generation Prototype HMPD-2

The HMPD-2 is the second generation HMPD-1. The optics and projection system covered a 42° diagonal FOV as apposed to a 52° for the HMPD-1. The compactness of HMPD-2 reduces the overall size and weight. As shown in Figure 29, the most notable changes were also the resolution (800 x 600), microdisplay type (OLED), and adjustable depth plane. The control boxes are now split into miniature boxes using eMagin's microdisplay technology, one for each eye. The miniature control boxes increase the participant's mobility by reducing the bulk and thickness of the cables that connect to the display. The one drawback to this display was that it

was a quickly assembled prototype. As such some of the features (i.e., the IPD adjustor) were not as stable as we expected and necessitated recalibrating the HMPD frequently.
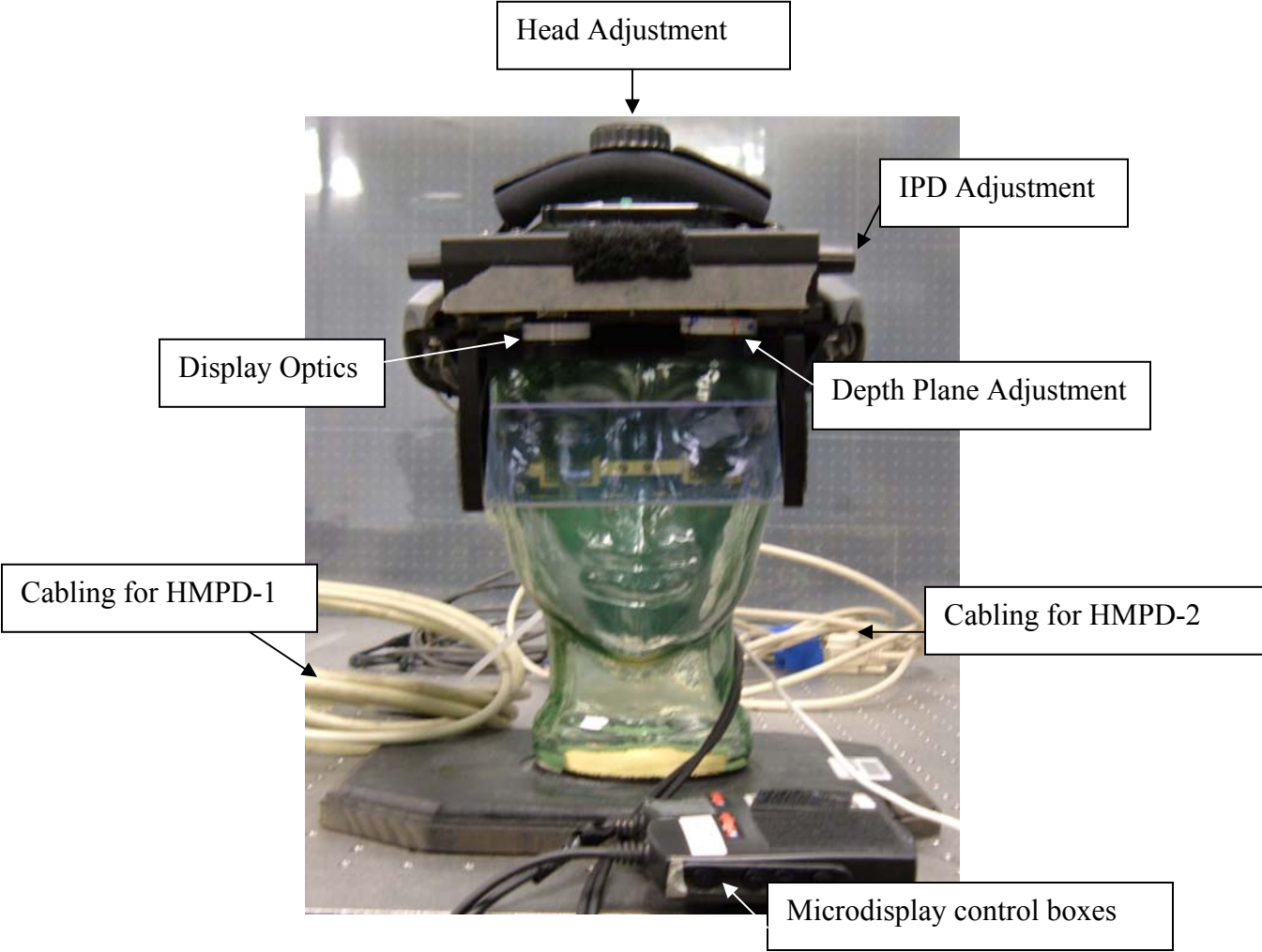
Figure 29 HMPD-2 system components

5.4.1 Assessing Lens Misalignment with Human Observers

When we first tested participants with the Virtual Howard-Dolman task (V-HD task) while wearing the HMPD-2, they were unable to fuse the image. Further, the participants reported a tilting of the virtual objects with respect to one another. Analysis of the resulting depth errors from the V-HD task showed that the mean responses were well beyond the range predicted in section 4.5.3. The results of the depth assessment suggested that one or both of the lenses within the HMPD-2 were tilted with respect to the image plane as diagrammed in Figure 30.

(a)   (b)
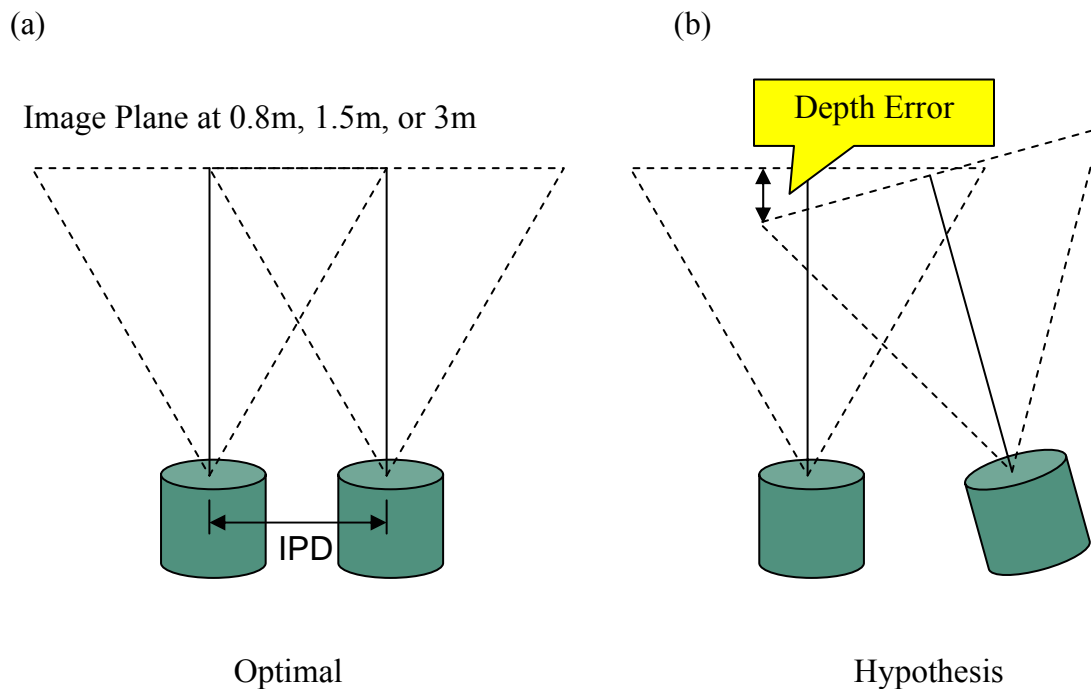
Figure 30 (a) Correct alignment according to specification (b) Suspected alignment based upon assessment.

We were able to test this prediction by measuring the parallax error between both eyes when participants viewed a stationary target with each eye separately and from each of the respective viewing distances. The stereoscopic stimuli were rendered to the display taking into account the IPD of the observer. Left and right eye views of a scene were generated by rendering the stimuli from each eyepoint separated by IPD. Given that the HMPD-2 has two separate channels for generating images to each eye, measurements can be taken by closing one eye channel at a time. By marking the location of a single point seen by one eye and then marking the location of that point as seen by the other eye, should return the IPD of the user if the lenses are parallel. Using this method we tested whether the lenses were set parallel as specified by the optics design when they were aligned by the mechanical design team. The lenses had been tilted at mechanical assembly and were realigned by the optics team and perceptual scientists.

The ARC is a semi-circular structure, thus to minimize the possible effects from the surface being curved a stand approximately .91 m wide and tall was created out of Plexiglas.[3] The material was chosen because it was lightweight and provided a smooth surface with which to affix the retroreflective material. The stand was angled at 90 degrees and was leveled so that the retroreflective material was parallel to the observers' line of sight.

A bull's eye was computer generated and rendered to the display with the corresponding IPD of each observer. We measured the mean angle of deviation imposed by the IPD for each observer. Figure 31 illustrates the setup for this assessment. A double cross grid was constructed on transparent vellum so that it could be written on and would not distort the image from the HMD.

(a)   (b)
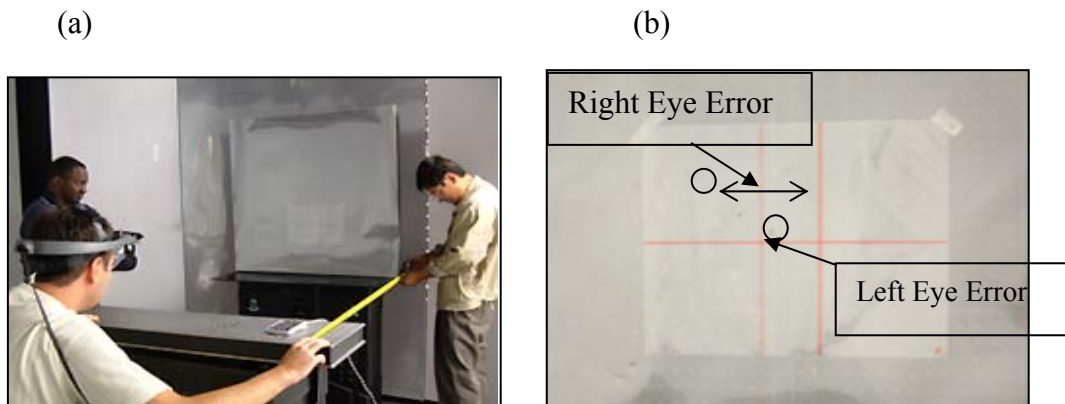
Figure 31 (a) Experimental setup for determining the angle of the lens tilt (b) Left and right eye deviation from respective IPD position

Three observers performed the assessment. A viewing distance was chosen and the observers took turns performing the assessment. The test was performed until each of the three depths had been tested. The procedure for each participant was to close one eye channel of the HMPD-2 and align the reference hair of the Bull's eye and then they aligned the opposite eye. The final position of the Bull's Eye was marked on the vellum as shown in Figure 31b.

The results of the parallax test showed that when the image plane was set to .8 meters there was a 5 cm (3.57 degrees) error between the projected focus target and the perceived image. Additionally, the perceived image was shifted to the left of the target suggesting that one lens was tilting inward while the other lens remained parallel. These results were repeated over the 1.5 and 3.0 meter depth planes thereby confirming the misalignment of the two lenses and are presented in Table 11. This procedure provided a solution to the issue of lens tilt raised by

---

[3] All machine work was conducted by Richard Zotti at the College of Optics CREOL-FPCE at UCF

Wann, Rushton, & Mon-Williams (1995), who suggested that optical see-through head-mounted displays should be tested for lens tilt prior to running experiments.

Table 11 Misalignment error of the two lenses

Image Plane (mm)	Error (cm)	Angle Error (degrees)
800	5	3.57
1500	8.5	3.24
3000	15.5	2.96

The results of the experiment suggested that the lenses were adjusted to converge as opposed to their specified optical design requirement of being parallel. However, the assessment also suggests that one lens was tilted horizontally off axis. The resulting errors were between 5 and 15 cm. The HMPD-2 needed to be calibrated prior to running the assessments. Since there were no methods to calibrate the display, a methodology was proposed and tested.

5.4.2 Calibrating the HMPD-2

We confirmed the results from the observer study with a laser setup. Figure 32 shows the initial setup. The HMPD-2 was flipped, so that the optics were pointing upward. An apparatus was setup to hold a penlight laser perpendicular to the optics face and was set 6 meters from the stand. A 1 inch by 1 inch window mirror was placed over the left eye optic barrel.

Figure 32 Initial laser based alignment setup

We ensured that the beam of the laser pointer was perpendicular to the mirror by positioning the laser so that the beam reflected back onto itself or in the center of its "lens". As shown in Figure 33, we used a paper "deflector" to make sure that the laser was not wildly pointing at an angle. The setup was placed on a horizontal slide so that we could move the setup to the other eye without disturbing the alignment of the laser
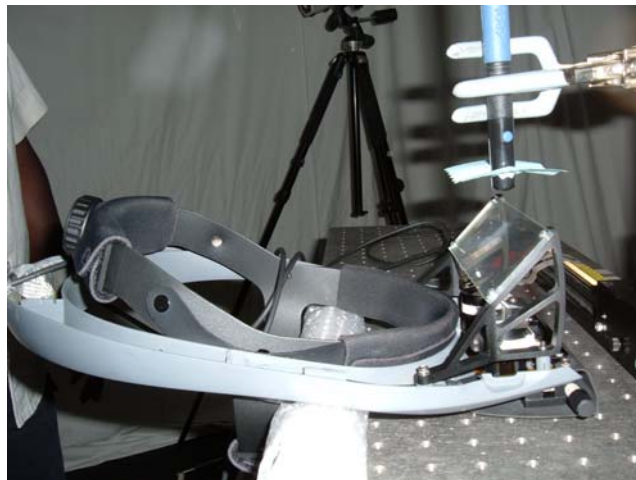
Figure 33 Deflector and laser pointer

The laser was then turned on and the projected point location was marked onto a piece of vellum that was previously used to check the image alignment during the Bull's-eye test. Once the point was made, we moved the laser setup toward the right eye and then re-centered it over the left eye. The laser was turned on and the point marked. This point fell exactly over the previous point, thus we concluded that translating the laser setup would not disturb the laser's perpendicular alignment. After marking the point, the mirror was moved to the right lens. Hereafter, the left eye position was used as the reference point.
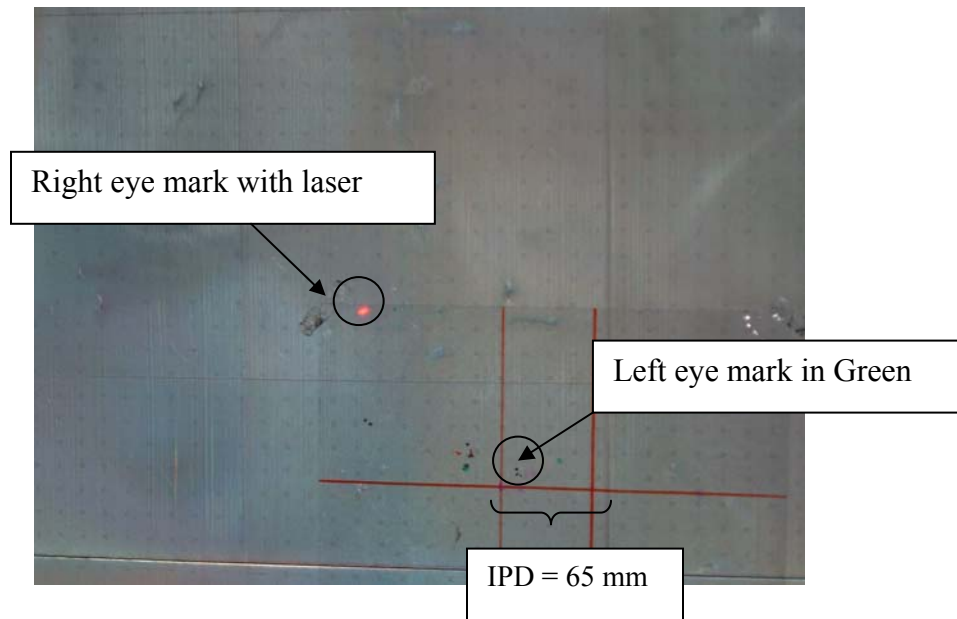
Figure 34 Laser error point representing alignment of the right eye and marker representing the left eye measurement

When we centered the lens under the laser and reflected the beam off the mirror to the retroreflective material, the resulting point appeared above the horizontal alignment of the left mark and displaced 13.4 cm to the left. The HMPD-2 head mount was set to .065 m or 6.5 cm suggesting a huge offset in the position of the right lens.

To realign the lenses the optical barrels had to be loosened with a hex wrench and manually moved. Figure 35 shows the instructions on how to align the optics. With this initial setup we were able to confirm the results of the experiment with human observers. However, the setup was not robust enough to allow for accurate alignment of the lenses. Thus, an alternative setup was devised and is pictured in Figure 32.

Loosen these screws to rotate the optics about the axis parallel to the screws to change the convergence
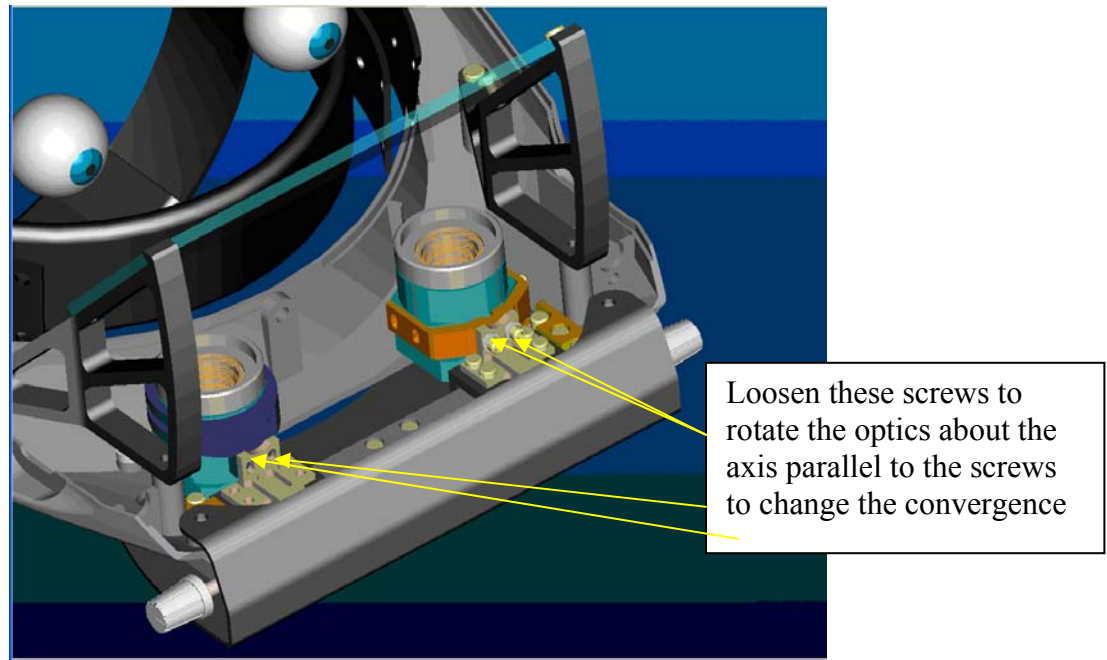
Figure 35 Alignment mechanism for both left and right eyes of the HMPD-2

A low level infrared laser was setup as pictured in Figure 36. The laser was set at almost a meter above the optics board. The board was made level with respect to the floor and the stand affix with retroreflective material was leveled with respect to the laser apparatus. Once the system was leveled, the vellum sheet with the double cross grid was adjusted so that the center line was parallel with respect to the laser level. This step was included so that the target stayed parallel with respect to the floor and perpendicular to the laser beam.
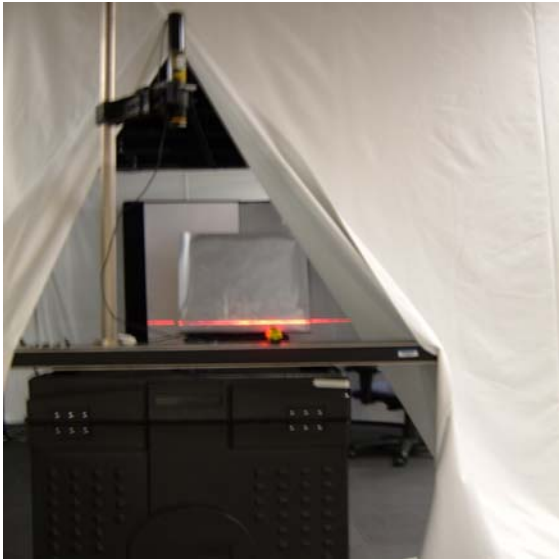
Figure 36 Final laser setup and alignment

The laser beam was set on a moving platform and could translate in the X direction, although it remained fixed with respect to its Y position. To ensure that the laser was aligned with the optics board, a mirror was placed on the board and a beam was projected off its surface as shown in Figure 37a. The resulting beam reflected back to the laser. If the laser were perpendicular to the table then the laser beam would follow the same path back to its source, which is located in the center of the barrel as shown in Figure 37b. We adjusted the laser until the path of the reflected beam aligned closely with the source. Once the laser was set, we placed the HMPD-2 level and directly under the source. The mirror was placed on one of the lenses as shown in Figure 37c. The laser beam was then reflected off the mirror. Because of the beam splitter, the laser beam was directed toward the retroreflective material as shown in Figure 37 d.
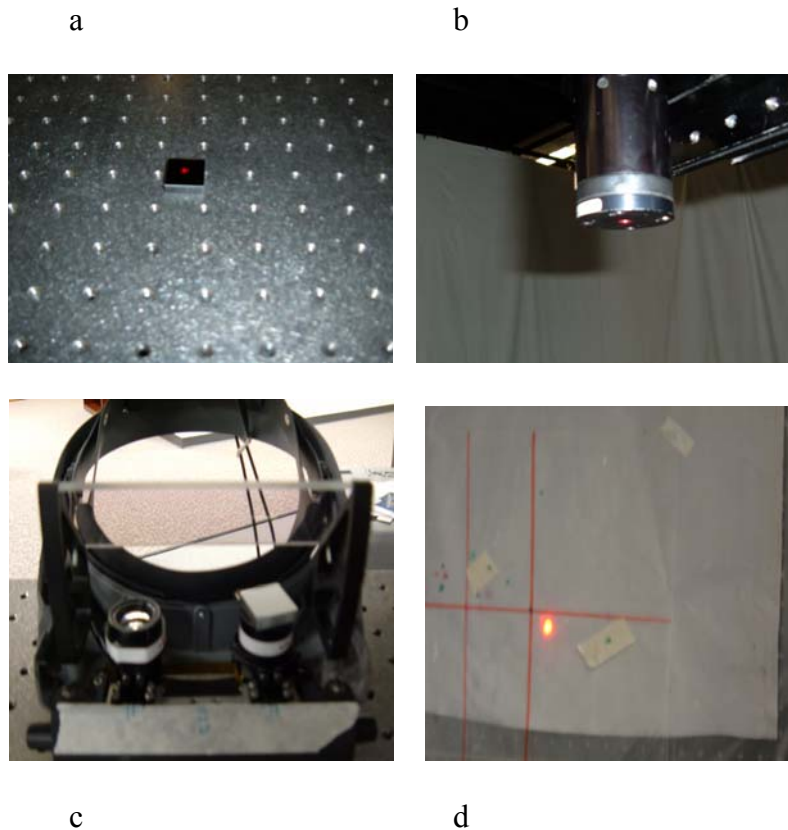
Figure 37 (a) Ensuring that the laser is perpendicular to the table (b) laser beam reflected off the mirror (c) mirror placed on lens and laser beam directed toward retroreflective material (d) laser beam reflected from the mirror onto the retroreflective material.

The left and right optics were adjusted separately. The optics were considered aligned when the laser beam location was as close as possible to the IPD point for that respective eye. This procedure took approximately three hours to perform. As well, the displays which were set inside the HMPD were not anchored mechanically. If one display disconnected, then the casing had to be opened to reconnect it. This issue became problematic because the left and right eye lenses were set to move independently when using the IPD adjuster. Thus, one could adjust one eye without moving the other. However, the track that the optics was set upon was unstable once the housing was removed. Therefore, if the housing is removed then the HMPD-2 would need to

be recalibrated. It should be noted that once the calibration procedure is complete, the HMPD-2 is stable and reliable in the field.

After the results of the first experiment were completed, the mechanical design team was contacted regarding the alignment of the lenses. The team confirmed that the optics had been set to converge at an image distance of 2 meters, which was not consistent with the optics specifications. We are unsure how customers of off-the-shelf products are able to confirm the parameters of the HMD without better documentation and baseline comparisons. On a personal note, I would never spend over 5,000 dollars on a stove for which I had to determine the heat setting by trial and error. Similarly, I would never buy a high end automobile, which is the cost of some of the HMDs on the market, if I had to constantly adjust the engine before driving it. This calls for the optomechanical design team to communicate closely with the optics and perception scientists in how they assemble the final product.

5.4.3 Setting the Image Planes for the HMPD-2

In addition to the lens alignment, the image plane distances had to be calibrated. This calibration is only performed when the barrels are removed from the system. The barrels were removed twice during the course of the experiments because of issues with the OLED. After a period of use they developed black dots in the FOV. Given that the microdisplay was glued to the optics barrel, the barrels themselves had to be removed and shipped out of lab for removal of the display.

The expected image plane size for each viewing depth can be determined from the active viewing area of the display, the focal length of the HMPD, and the FOV in both the X and Y directions. The parameters above were set in an optical design software package called Code V.

109

The software used the parameter values in a ray tracing, tracing a ray of light from the eye back through the image plane into object space. In essence, the code matches the image height to the display height. The resulting image heights are given in Table 14.

Table 12 Image plane width, height, and diagonal measures based upon viewing distance

Screen distance(m)	Screen Width (mm)	Screen Height (mm)	Screen diag. (mm)
0.8	479.7	360.4	600.0
1.5	911.6	684.8	1140.2
3	1837.5	1380.4	2298.2

To adjust the settings for the HMPD-2, three observers were needed. The general procedure is for a single observer to stand at one of the viewing depths. With the HMPD adjusted for their IPD, they are to view a projected image of a high contrast stimulus (e.g., text) through one eye at a time. They then turned the adjustment knob for that lens until the stimulus appeared clear. Text was especially helpful since he or she could read the text when the HMPD-2 was brought into proper focus. Each observer must confirm the setting for that level of viewing distance for each eye. The barrels of the optics were then marked with an indicator for each depth.

CHAPTER 6 RESOLUTION VISUAL ACUITY: LIGHTING, CONTRAST, & RETROREFLECTIVE MATERIAL AFFECTS

6.1 Introduction

Our goal in developing the HMD Performance Assessment Battery is to provide VE system designers with a user-centered methodology for systematically optimizing each aspect of the virtual environment (VE) from the HMD, computer graphics, to interactive devices (Fidopiastis, Meyer, Fuhrman, & Rolland, 2003). Given this approach, we hope to minimize problems associated with integrating display technology with end user task specifications. For example, when performing tasks using a wide field-of-view HMD, the user's resolution visual acuity is typically reduced due to the limited image resolution of the miniature displays used in the HMD. Quantifying the decrement in rVA would allow the designer to make better decisions on what types of graphics to display or even to determine if the HMD is appropriate for the task. However, other aspects of the VE setup may degrade user performance in addition to or as the limiting factor. Such is the case with the retroreflective material used with the HMPD.

In the ARC display, the optics of the HMPD project a separate left and right image into the augmented VE using a 50/50 beam splitter to reflect light off the retroreflective material. Because the retroreflective material reflects the light in the opposite direction of its incidence, stereo images are returned to the eyes of the user. The limit in user resolution acuity is imposed by the limited resolution of the microdisplay in current HMDs. One method of overcoming the limitations of the microdisplay is to augment the properties of the retroreflective material (e.g., increased brightness).

Currently, there are two types of retroreflective material used in the ARC display, Scotchlite[TM] 3M Fabric Silver (Beaded) and Scotchlite[TM] 3M Film Silver (Cubed). The material for the Scotchlite[TM] 3M Film Silver (Cubed) should reflect more light, thereby enhancing the apparent brightness of the projected image. This quality should increase the light entering the user's eye, thus improve resolution visual acuity. In the forthcoming experiments, we assessed users' rVA with the computerized version of the Landolt C (referred to as the optotype) visual acuity test while viewing the virtual optotypes with either the Scotchlite[TM] 3M Fabric Silver (Beaded) or Scotchlite[TM] 3M Film Silver (Cubed) in the environment. Both high contrast and low contrast stimuli were used to compare differences between the two materials given differences in their ability to reflect light.

As pictured in Figure 38, the ARC display is a circular structure with retroreflective panels covering the interior. The panels are .91 meters wide and 3 meters high. The ceiling and entrance of the ARC is covered with a heavy black cloth to block extraneous light from entering the VE. To assess differences in retroreflective material performance when lighting conditions changed, we tested rVA under three different light levels. Because of the relatively low level of illumination in the HMPD, only mesopic and scotopic light ranges were used in the ARC. These light levels roughly correspond to illuminance IES Lighting Handbook (1987) for lighting spaces for optimal user performance listed in Table 13.

Table 13 Illuminance categories and values IES Lighting Ready Reference (1989), p. 87

Performance on Tasks of high contrast or large size	200 to 500 Lux (18.58 to 46.45 footcandle)
Occasional working space	100 to 200 Lux (9.29 to 18.58 footcandle)
Simple orientation for short temporary visits	50 to 100 Lux (4.64 to 9.29 footcandle)

6.2 High Contrast rVA Assessment Experiment 1

6.2.1 Hypotheses for High Contrast rVA Assessment Experiment

Given the current VE system setup, we can make the following predictions:

Hypothesis 1: Based on the HMPD-1 display size, and FOV, we estimated a maximum visual acuity of 4.1 minutes of arc.

Hypothesis 2: Given the increased brightness of the Scotchlite^{TM} 3M Film Silver (Cubed) material and subjective visual observation of brightness differences between the two materials, we expected slightly higher rVA scores with Scotchlite^{TM} 3M Film Silver (Cubed) material in the environment.

Hypothesis 3: As the light levels in the ARC increase, the contrast of the optotype decreases, this may lead to a decrease in performance on the rVA.

6.2.2 Method High Contrast Experiment

Participants

This study was approved by the University of Central Florida Institutional Review Board (IRB). The protocol is listed in Appendix X. Twelve participants (11 men and 1 woman, mean age = 28years, SD = 5.70) performed the Landolt C Visual Acuity test under 3 different light levels presented in Table 13 under real world viewing and augmented reality viewing (e.g., using the HMPD) conditions. Half the participants viewed the cubed material while performing the

Landolt C visual acuity test wearing the HMPD-1. The other six participants performed the assessment while viewing the beaded material. Each participant was either corrected for or had 20/20 vision. Glasses or contacts were worn during each part of the experiment.

VE System

We used a first generation see-through projection head mounted display (HMPD) with a 52 degree FOV. The microdisplays for the HMD were off-the-shelf Liquid Crystal Displays with a VGA resolution of 640x480. Software ran on a computer system with Linux RedHat 7.2 OS and a dual processor graphics card. The LCD monitor was a Dell 17 inch flat screen. Within the ARC, we replaced two of the, Scotchlite^{TM} 3M Fabric Silver (Beaded) panels with Scotchlite^{TM} 3M Film Silver (Cubed) panels.

Figure 38 VE system: ARC, HMPD, and computer setup

Modified Landolt C Computer versus Augmented Reality Version

Six different gap sizes representing the different levels of visual acuity in minutes of arc were presented. The stimuli were the same in each condition; however, the visual angle corresponding to each size (arc minutes) of the Landolt C changed from the computer to the

augmented reality condition. Table 14 below presents the change in visual angle for each

environment. To accommodate for these differences, the participants were seated 1.85 meters

from the LCD monitor in the computer version and were 2 meters from the retroreflective

material in the augmented reality condition.

Table 14 Visual acuity conversion values

Computer (Arc Minute)	Snellen Conversion (Ft and Meters)	Augmented Reality (Arc Minute)	Snellen Conversion (Ft and Meters)
1	20/20 or 4/4	4.1	20/82 or 4/16.4
2	20/40 or 4/8	8.2	20/164 or 4/32.8
3	20/60 or 4/12	12.4	20/248 or 4/65.6
4	20/80 or 4/16	16.5	20/330 or 4/66
5	20/100 or 4/20	20.6	20/412 or 4/82.4
6	20/120 or 4/24	24.7	20/494 or 4/98.8

Lighting setup

Although other light configurations can be used (e.g., light reflected off the ceiling), we

chose to suspend the light from the ceiling of the ARC to better control the effects of glare and

extraneous light that may bounce off the floor onto the retroreflective material. Three 13 Watt 4-

pin PI fluorescent lamps (780 lumens, Warm 2700K) were mounted on a 51.5 cm by 59.5 cm

gator board. The board was attached to the ceiling in the center of the ARC as illustrated in

Figure 39. A diffuser of sheer black material and length .72 meters was attached to the board.

The lighting levels were set using a Minolta T-10 illuminance meter. Illumination ranged from

High (225 -209 Lux), Medium (140 to 132 Lux), and Low (61 to 53 Lux) to give an average 80

Lux difference between each light level. These light levels are presented in Table 13.

Figure 39 (a) Lighting and computer setup (b) Adjusting the HMD for a participant

Experimental Design

Participants were randomly placed in a retroreflective material condition (beaded or cubed). The three light levels were changed for each participant according to a partially counterbalanced design determined prior to the experiment. Within each test block, each Landolt C gap size was presented in random order. As well, there were five random presentations of each direction (up, down, left, and right) per level of visual acuity. Test blocks were run first with the computer version of the Landolt C test under real world viewing conditions, and then within the HMPD-1. Participants completed one test block per lighting condition for a total of 3 blocks per viewing environment. Within each block, they experienced each gap size 20 times. There were a total of 720 responses per participant. Once the responses were converted to percent correct and, then converted to NED values, a Probit analysis using light level, visual acuity, and gap direction as independent variables was completed.

<u>Procedure</u>

The initial lighting level was set for each participant prior to their entering the ARC. Participants were first seated 1.85 meters from the computer monitor. The interpupillary distance (IPD) of each subject was measured using a pupillometer. This value was entered into the Landolt C test to set the correct position for the left and right eye display of the stimuli on the LCD monitor. The IPD on the HMPD was also adjusted to center the optics and the image on the eyes of the observer in the augmented reality condition. This initial setup allowed the participant to visually adapt to the lighting conditions within the ARC prior to the start of the experiment. Once correctly positioned and the IPD entered, the participants were given a keyboard, and instructed to use the arrow keys to indicate the direction of the gap in the Landolt C stimuli.

Each time the participant key pressed their response, a beep would precede the appearance of the next stimuli. The test stimuli would remain on the screen until the participant indicated a response. There was no time limit in responding. The participant continued to respond until the program indicated that the test was complete. The software program recorded each response in a tab delimited file format that was exported to MS Excel. A combined data file was created for each subject across each lighting and viewing condition (computer or HMPD-1).

Once the subject had performed the computer version of the VA test, they were asked to stand on a mark 2 meters from the retroreflective material. The HMPD was adjusted for IPD and head size as shown in Figure 39b. Once the rVA test was completed within the HMPD-1, the subject was asked to sit again while the experimenter changed the light level. The procedure was repeated until data was collected for all three light levels.

6.2.3 High Contrast Visual Acuity Experimental Results

In all experiments, the ANOVAs were performed using SPSS 11.5, the Probit analysis was done using STATGRAPHICS plus 5.1, and all graphs were constructed using Minitab 14. Each statistical software package performs the respective analyses using slightly different formulas and assumptions. Software was chosen based on exactness of procedures. For example, STATGRAPHICS plus 5.1 provides a Probit analysis across multiple independent variables using a maximum likelihood regression, whereas SPSS does not.

The results showed that there was no difference in performance on the real world viewing condition for both groups in terms of visual acuity (M_{Cubed} = 98.64, SD = 3.22; M_{Beaded} = 99.07, SD = 1.96). For visual acuity corresponding to 20/20 vision (1 arc minute), the participants in each group performed well above the 62.5% threshold (M_{Cubed} = 93.70, SD = 4.82; M_{Beaded} = 97.04, SD = 2.89). Although performance was poorer in the HMPD-1 for the smallest visual acuity (4.1 arc minute), the proportion of correct responses was still above the 62.5% threshold (M_{Cubed} = 88.89, SD = 11.91; M_{Beaded} = 81.48, SD = 14.66). These results are graphically displayed in Figure 41 a and b.

a.

b.
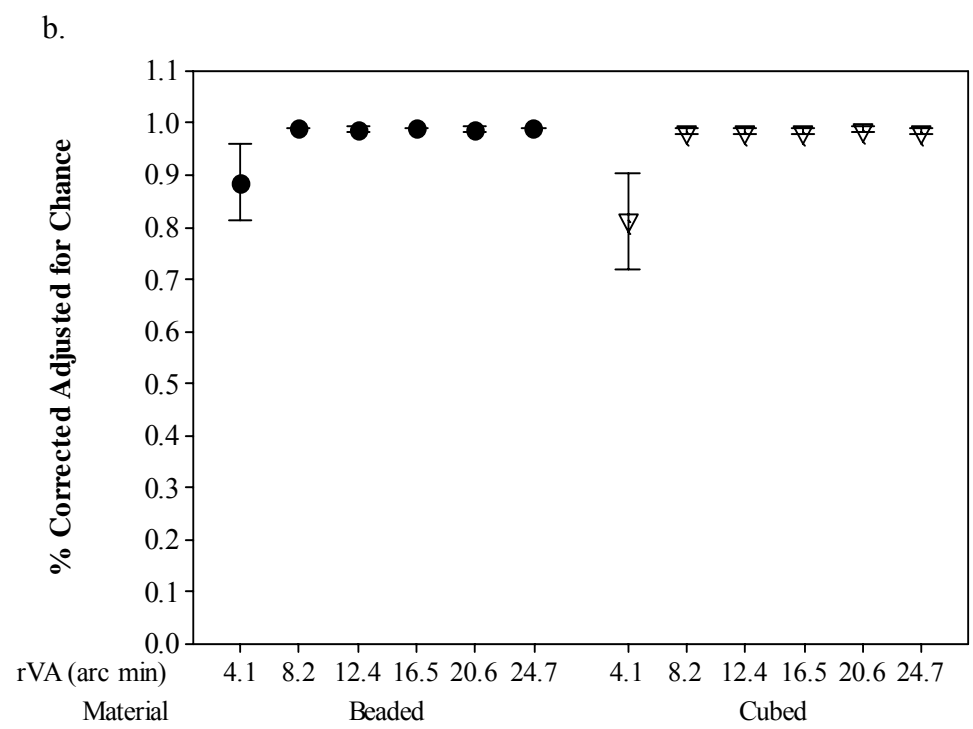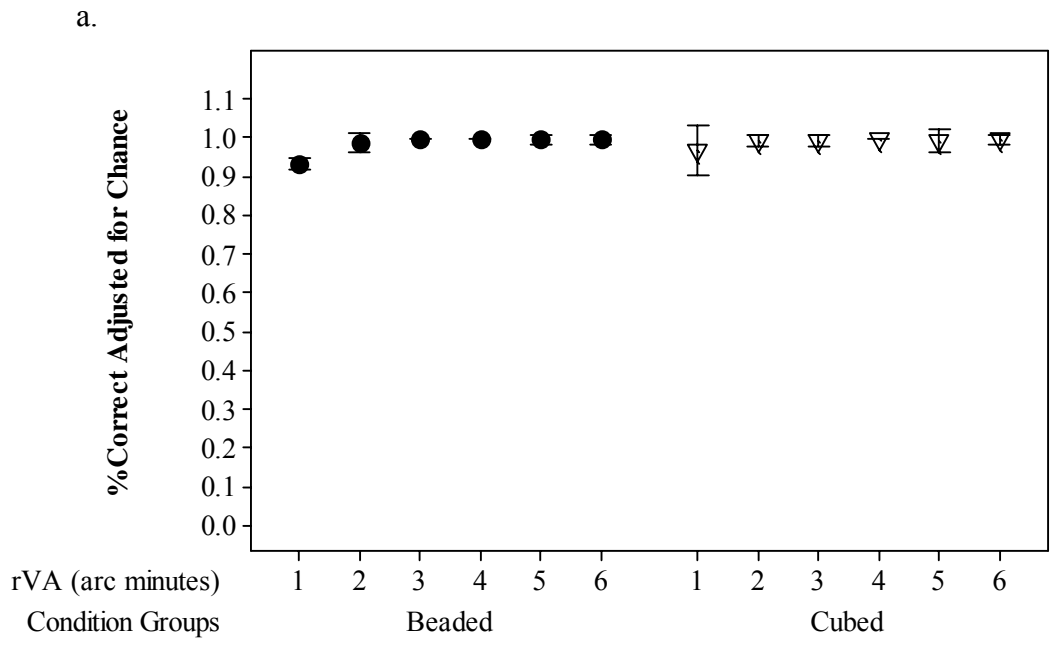
Figure 40 (a) Real world viewing of rVA assessment showing no change in responses over testing (b) Modified virtual Landolt C results show rVA limit of HMD.

A mixed ANOVA (2 material types x 6 gap sizes x 3 light levels), where material type is a between groups measure and gap size and light levels are within groups measures, was used to analyze the percent correct responses adjusted for chance. The ANOVA showed that there was no significant effect of room lighting on rVA while testing in HMPD-1 or while viewing either type of retroreflective material $F(2, 144)= 4.39$, $p = .139$, $\eta = .217$. There was a main effect of gap size, $F(5, 50) = 6.83$, $p = .001$, $\eta = .406$. As figure 39 shows, rVA performance declined when reaching the pixel size limit of the display. Also, retroreflective material type did not affect rVA performance.

As expected, participants exhibited gap direction errors at the 4.1 arc minute level when viewing the optotypes on either of the different materials regardless of lighting conditions. However, certain gap directions were harder to identify depending upon material type. Those who performed the rVA assessment viewing the beaded retroreflective material were only 65.9% correct at determining a gap facing upward. Accuracy at identifying leftward facing Landolt Cs was at approximately 73%. Participants were above 90% correct at identifying the other two directions. In contrast, those participants who viewed the cubed material were 86.7% correct at identifying the upward facing gap. While the participants were only 75.8% correct at identifying a gap facing downward. The participants were above 95% accurate at determining the other two directions. For participants in both groups, the right facing optotypes were the most correctly identified (98.5% for the beaded group and 100% for the cubed). For other gap sizes, the participants were over 90% accurate at identifying each gap direction. All percentages reported were adjusted for chance.
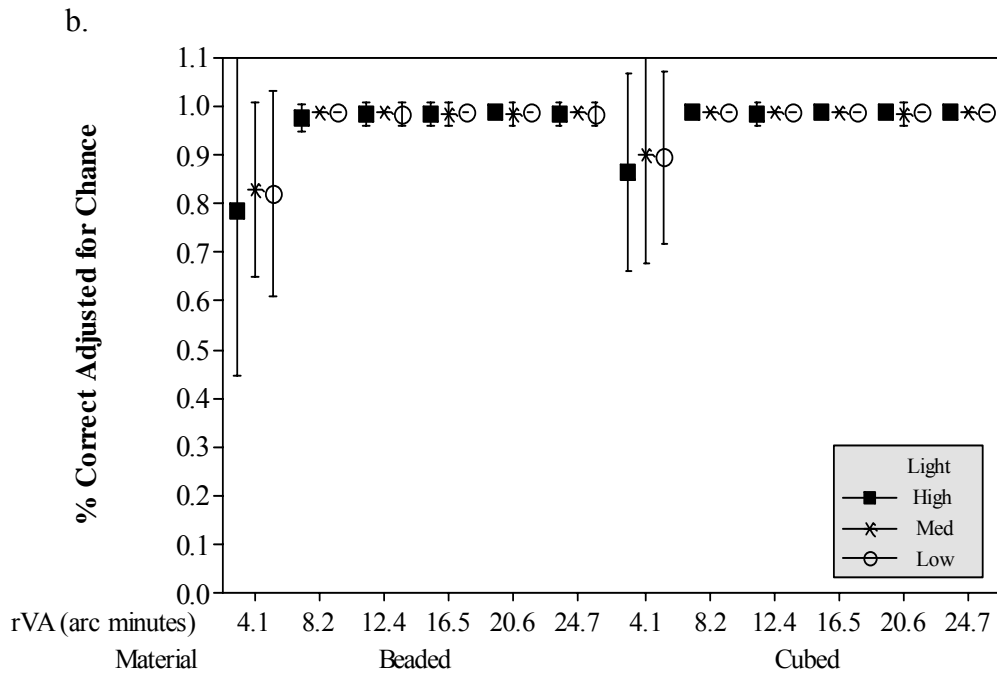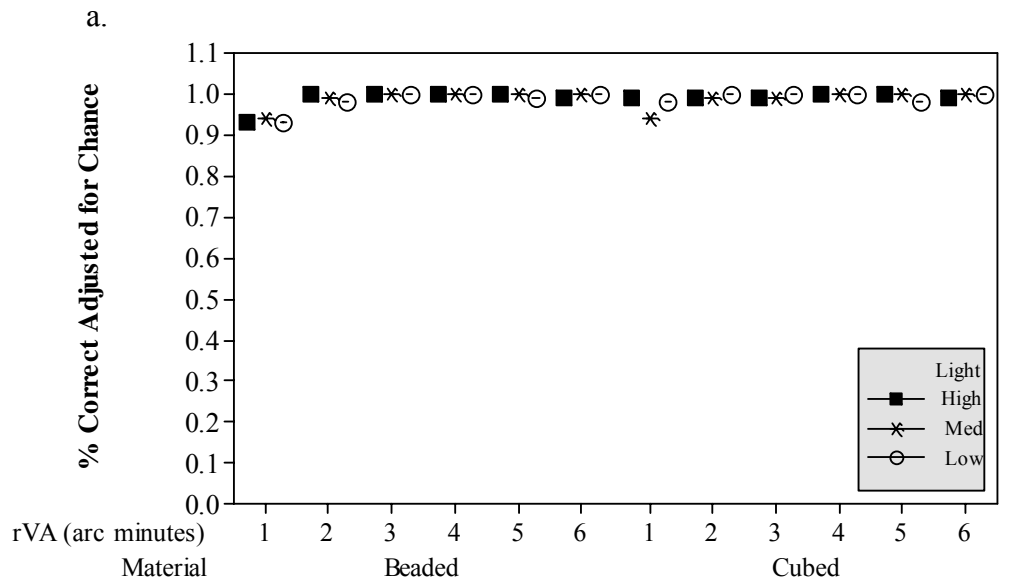
Figure 41 Percent correct corrected for chance across material, viewing, and lighting (a) real world viewing (b) HMPD-1 performance under different lighting conditions

6.2.4 Conclusion High Contrast Experiment

The results of the experiment confirmed that the visual acuity limit using the first generation see-through HMD was 4.1 arc minutes. This rVA limit is imposed by the microdisplay, and not the retroreflective material. However, as we are able to use higher microdisplay resolutions, the retroreflective material may become the systems limiting factor. Participants performed no differently when viewing stimuli projected from either retroreflective material. Thus, given high contrast visual stimuli either material will support the user's ability to see details in the scene up to the limit of the HMPD-1.

The experiment also showed that given the constraint in lighting imposed by the HMPD, there is no benefit to visual acuity in changing light levels within a scotopic or mesopic range. Although the rVA values were not improved by the lighting conditions, they were not degraded either. This result suggests that participants may be able to perform tasks in personal space (slightly beyond arms reach) that require natural viewing while also viewing a 3D scenario (e.g., occupational therapy instructions) projected into action space (beyond arms reach to 30 m). The ARC system may provide an optimal training environment for different types of rehabilitation (e.g., physical) where some interactions are difficult to replicate in a virtual world. Given that we can reduce the ARC to a single panel of retroreflective material, the system may be easily setup in the patient's home where the patient can view the therapist remotely while performing the training tasks using real world objects. Although more studies are needed to explore the efficacy of such a VE setup within the field of rehabilitation, this research lends support to such applications.

With respect to the HMD, lighting conditions within the ARC may also become less constrained as HMDs become less reliant on the contrast created between low levels of

illumination and the projected image. The next experiment will explore the effects of contrast given the current ARC setup. In addition, errors in reporting gap direction using the modified Landolt C should be explored further. Visual errors may be introduced by the pattern of the beads or corner cubes within each type of material. These patterns may become more apparent as the resolution of the display increases.

6.3 Low Contrast rVA Assessment Experiment 2

In real world viewing conditions, all information is not always presented with high contrast and fidelity. We often experience fluctuations in environmental lighting that degrades the visibility of objects, such as viewing car brake lights at twilight. As with see-through HMDs, low contrast between the displayed image and the surrounding environment degrades image detectability beyond that of the human visual system. We know from experiment 1 that all users with 20/20 corrected vision can reliably identify high contrast targets at the resolution limit of the HMPD-1 when viewing optotypes on each retroreflective material. Next, we used the rVA assessment to evaluate the projection head-mounted display for changes in user visual acuity as a function of target contrast. Participants performed the visual acuity test while viewing either retroreflective material to evaluate the head-mounted display system.

The goals of this experiment are three-fold. First, we are interested in determining limitations in user performance while wearing the prototype HMPD-1. This evaluation assists in determining upgrades or improvements in the next phases of the design cycle. Second, we would like to evaluate the VE system components to better understand weaknesses from a human interaction perspective in the current system setup. Given a calibrated HMPD and an understanding of the retroreflective material properties, the rVA assessment should aid in

determining properties of the VE system which are limited by the HMPD-1 or the projection material separately. Thus, future tests could discern limitations attributed to each individual component of the VE system or the system as a whole. Third, we are developing a means to obtain baselines and quantify individual differences in performance when interacting within a VE system. The assessment must be sensitive enough to quantify different levels of individual performance. This capability will allow us to screen participants prior to performing experiments as well as to understand modifications to the VE design needed to support user interaction.

In this experiment we held the lighting constant to evaluate perceived contrast separately from changing illumination levels in the ARC. Luminance of the Landolt C was varied by adjusting the binary grey-scale code within the computer program as reported in Chapter 4. The background of the display was set at a grey-scale value of 128. From this value, target grey-scale values giving contrasts of .2, .4, .6, .8, and 1 were computed. As with the previous experiment, participants performed the rVA assessment under real world viewing conditions as well as in the HMPD-1.

6.3.1 Hypotheses for Low Contrast rVA Assessment Experiment

Given the current VE system setup, we can make the following predictions:

Hypothesis 1: The addition of the contrast stimuli will allow for the screening of participants who are not capable of discerning low contrast targets in general. Thus, we are better able to control for individual differences within an experimental design.

Hypothesis 2: The rVA assessment will reliably measure rVA individual user performance over multiple days.

Hypothesis 3: The point spread function for each material suggests that participants

viewing the changing contrast stimuli projected onto the Scotchlite[TM] 3M Film Silver

(Cubed) material will attain higher rVA scores. Thus, the HMPD-1 will not be the limiting

factor to rVA acuity.

6.3.2 Method Low Contrast Experiment

<u>Participants</u>

Six participants (all male, mean age = 30.5, SD = 7.09) performed the complete experiment.

Each participant was either corrected for or had 20/20 vision. Glasses or contacts were worn

during each part of the experiment.

<u>VE System</u>

Participants performed the augmented reality version of the assessment while wearing the

same HMPD-1 discussed in Experiment 1. The microdisplays for the HMD were off-the-shelf

Liquid Crystal Displays with a VGA resolution of 640x480. Software ran on a computer system

with Linux RedHat 7.2 OS and a dual processor graphics card. The LCD monitor was a Dell 17

inch flat screen. Within the ARC, we replaced two if the, Scotchlite[TM] 3M Fabric Silver

(Beaded) panels with Scotchlite[TM] 3M Film Silver (Cubed) panels.

Contrast Modified Landolt C Assessment

Six different gap sizes representing the different levels of visual acuity in minutes of arc were presented, as with Experiment 1. The same adjustments for viewing angle were applied in this experiment. The participants we seated 1.85 meters from the LCD monitor in the computer version and were 2 meters from the retroreflective material in the augmented reality condition. Each contrast level was presented randomly. To control for carryover effects, all levels of gap size were randomly presented within a contrast block.

Experimental Design

Participants performed the contrast rVA test over a two-day period. They were randomly placed in retroreflective material condition (beaded or cubed) on Day 1, and performed the contrast rVA test on the opposite material the following day. There were 4 random presentations of each direction (up, down, left, and right) per level of visual acuity per contrast level. Test blocks were run first with the computer version of the Landolt C test, and then within the HMPD-1. Participants completed one test block per contrast for a total of 5 testing blocks. There were a total of 480 responses per participant for each viewing condition. Probit analysis using contrast, visual acuity, and gap direction as independent variables was completed. Additionally, a within subjects repeated measures ANOVA was performed on the percent correct responses adjusted for chance.

126

The 12 participants from Experiment 1 performed the contrast version of the Landolt C rVA test under constant low light levels with either the monitor or HMPD-1 as the display type. Participants were screened for contrast rVA above chance on the monitor version of the rVA test. More specifically, participants who failed to reach threshold for identifying gap directions at the 1 arc minute level with the lowest contrast were not included in Experiment 2. Of the 12 participants, only 6 participants qualified to perform the complete experiment. Each participant who performed Experiment 2 was either corrected for or had 20/20 vision. Glasses or contacts were worn during each part of the experiment.

The participants performed the contrast rVA test over a two-day period. On Day 1 participants were randomly placed in a retroreflective material condition (beaded or cubed). Two days later they performed the contrast rVA test again viewing the opposite material.

There were four random presentations of each direction (up, down, left, and right) per level of visual acuity per contrast level. Contrast was calculated by adjusting the binary gray-scale code within the computer program. After setting the background of the display to a gray-scale value of 128, target grey-scale values giving contrasts of .2, .4, .6, .8, and 1 were computed. A Probit analysis using contrast, visual acuity, and gap direction as independent variables was completed to determine the rVA threshold for each contrast level. Additionally, a within subjects repeated measures ANOVA was performed on the percent correct responses adjusted for chance.

6.3.5 Low Contrast Visual Acuity Experimental Results

Figures 42a and 42b show the 95% confidence intervals for the means of the percent correct values adjusted for chance for the monitor display condition over the two testing days. The graphs show that there are no differences in rVA measured when performing the modified Landolt C test over the five contrast levels and between the two days of testing. All percent correct adjusted for chance values were 90% or above for all contrast and gap size levels. The means for .2 contrast and gap size of 1 arc min for each testing day were $M_{Day1} = 97.2$, $SD = 2.75$ and $M_{Day2} = 91.5$, $SD = 7.23$.

a.

Monitor Display Type, Testing Day = 1

Monitor Display Type, Testing Day = 2
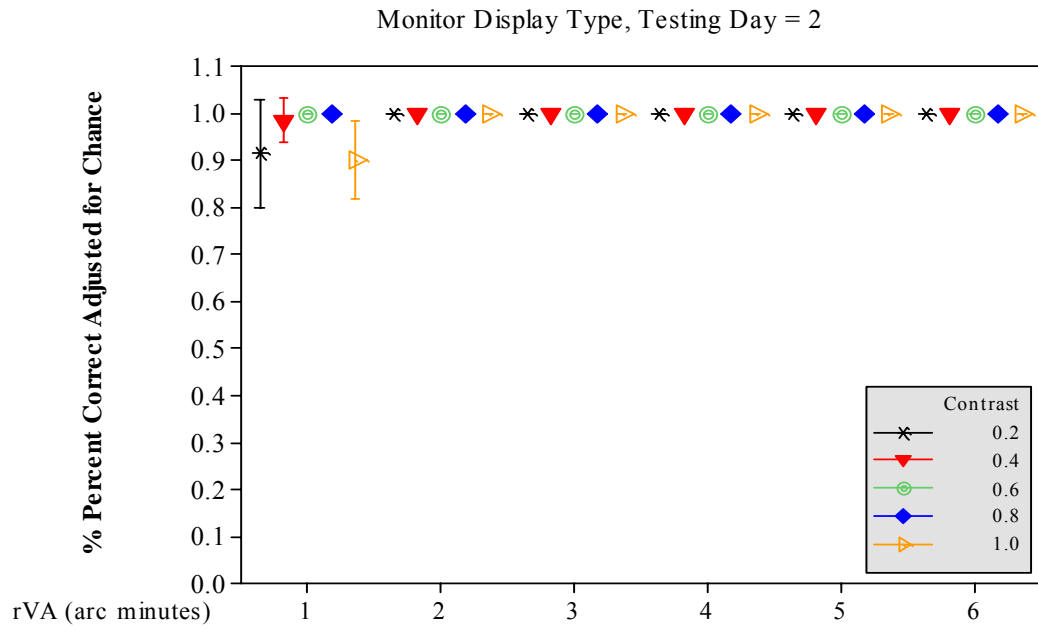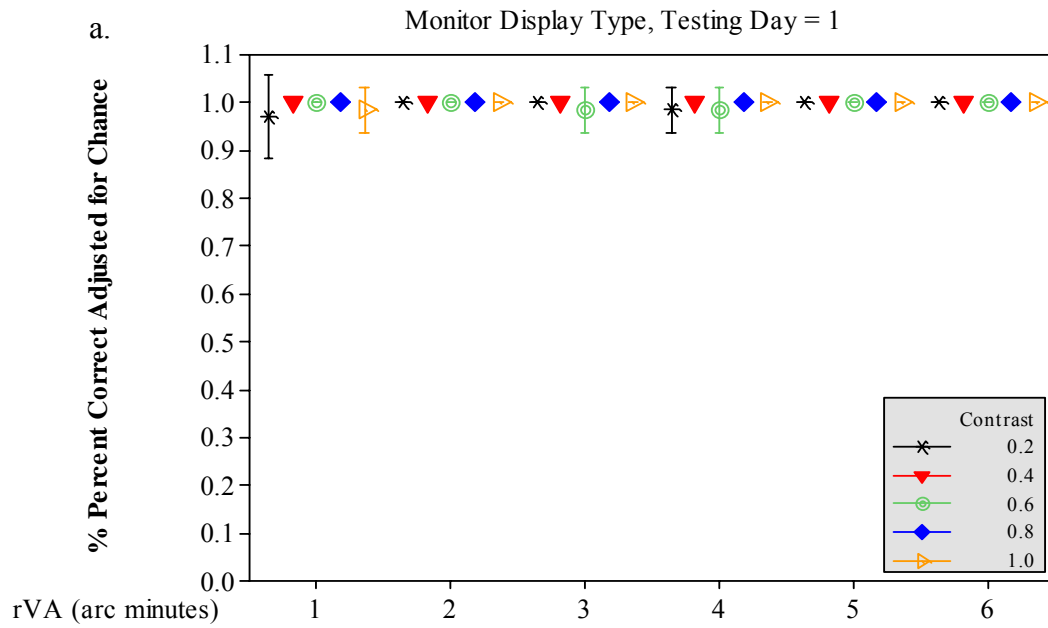
Figure 42 Real world viewing results over consecutive days

A repeated measures ANOVA (2 material types x 6 gap sizes x 5 contrast levels) was performed on the percent correct values adjusted for chance for the HMPD-1 display condition only. The results showed a two-way interaction between the type of material tested on and the gap size $F(5, 25) = 4.72$, $p < .01$, $\rho_I = .380$. Table 15 shows the means and the standard deviations of the % correct adjusted for chance outcome for the .2 contrast level and 4.1 arc minute gap size for both retroreflective materials. Participants viewing the modified Landolt C optotypes projected onto the beaded material performed worse than when viewing the stimuli projected onto the cubed material. Results from the Landolt C test for both material types are graphically displayed in Figure 43.

Table 15 % Correct adjusted for chance per contrast as a function of material type for 4.1 arc minute gap size.

Material	Contrast	%Correct Adjusted for Chance	
		Percent	
		Mean	SD
Beaded	.2	32.50	1.10
	.4	49.80	21.80
	.6	55.75	13.74
	.8	59.75	18.34
	1.0	76.50	21.40
Cubed	.2	58.25	4.96
	.4	100.00	0.00
	.6	93.25	5.20
	.8	94.25	2.25
	1.0	98.50	1.50

130
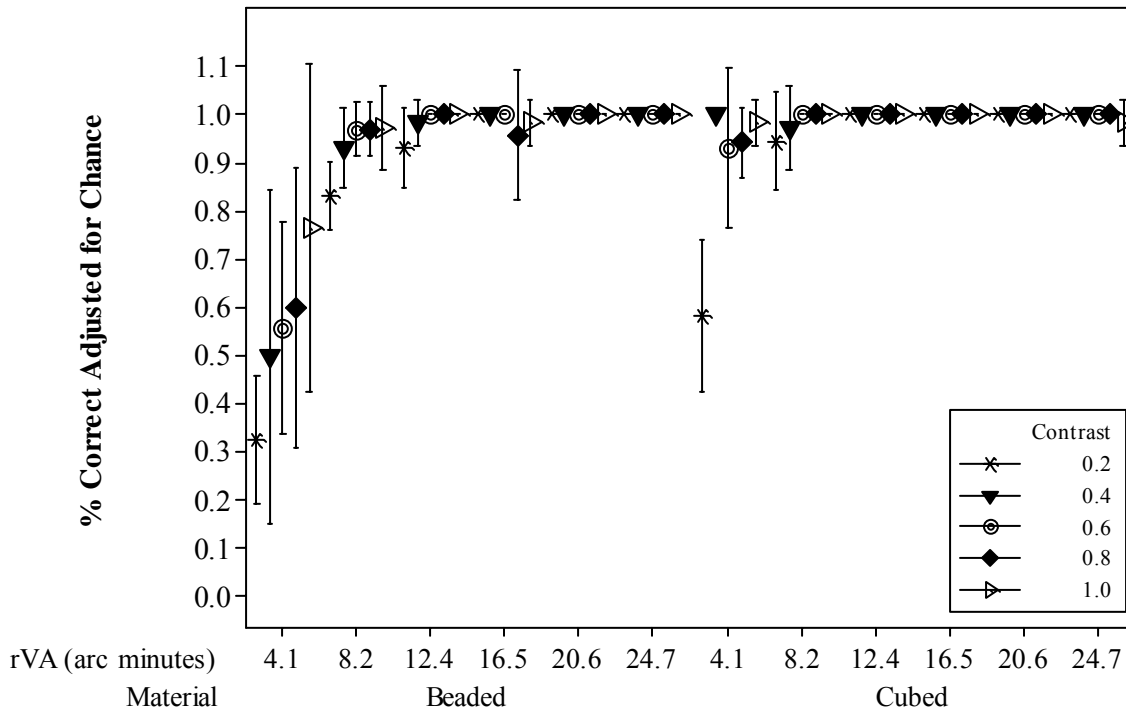
Figure 43 Percent correct adjusted for chance for different contrast optotypes and materials

The rVA thresholds predicted by the Probit analysis are shown in Table 16 and Figure 44. The predicted rVA while viewing the beaded and cubed material for a contrast of .2 was 8.9 arc minutes and 6.2 arc minutes, respectively. A repeated measures ANOVA (2 material types x 5 contrast levels), with material types and contrast levels as independent variables and the predicted rVA as the dependent measure also showed a main effect for both contrast and material $C_{Contrast}(4, 20) = 12.4$, p < .01, $\rho_I = .655$ and $F_{Material} (1, 5) = 11.9$, p < .02, $\rho_I = .645$, respectively. As expected, participants' rVA decreases with contrast level and is poorer when viewing the optotypes projected onto the beaded retroreflective material.

		Visual Acuity (arc minutes)	
		rVA	
Material	Contrast	Predicted	SD
	.2	8.9	1.17
Beaded	.4	6.8	2.12
	.6	6.1	2.25
	.8	5.5	2.12
	1.0	4.2	0.00
Cubed	.2	6.2	2.19
	.4	4.2	0.00
	.6	4.2	0.00
	.8	4.2	0.00
	1.0	4.2	0.00

Table 16 Predicted resolution visual acuity (arc minutes) from Probit analysis.

Estimated Resolution Visual Acuity Threshold
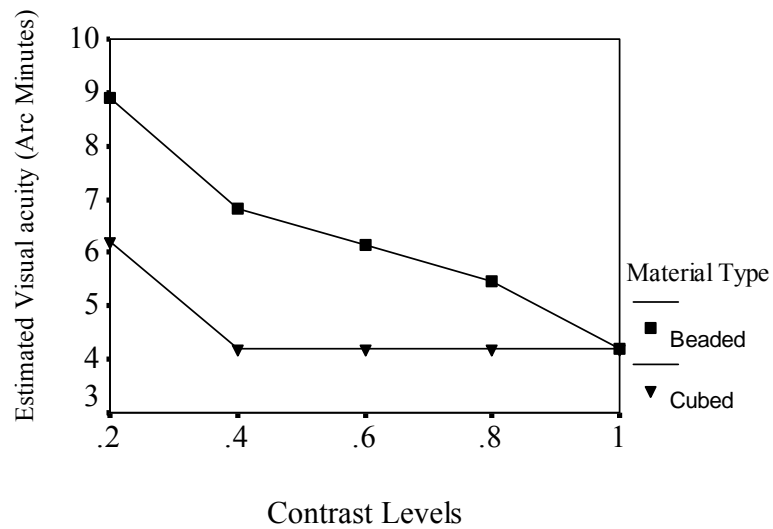
Figure 44 Estimated rVA for different contrasts over different materials

6.3.5 Conclusion Low Contrast Experiment

The modified Landolt C test reliably assessed rVA for different observers across different testing days. Further, the monitor display test environment provided a screening tool to assess user performance on the contrast stimuli prior to the participants entering the experiment. From an experimental design perspective, the contrast rVA assessment allowed us to identify participants who perform poorly when viewing low contrast stimuli in general. The data obtained to assess the differences between the retroreflective materials are not confounded by inherent visual limitations of the user. More importantly the assessment was able to identify those participants who may have altered vision due to corrective surgical procedures either from severe eye injury or vision correction (e.g., Laser-Assisted *In Situ* Keratomileusis, LASIK surgery).

Two of the participants who failed to meet threshold on the monitor based assessment had LASIK surgery approximately two years prior to participating in the experiment. LASIK surgery, reshaping of the corneal surface with a laser, has become increasing popular to correct for myopia or nearsightedness. However, one possible adverse effect is decreased contrast sensitivity (Sugar et al., 2002). The percent correct for each participant was 25% and 44% respectively for the .2 contrast level and 4.1 arc minute gap sizes. A third participant had suffered a previous eye injury that required surgical intervention that left scarring on the retina. The percent correct for this participant for the same contrast and gap size was also well below threshold at 38%. Each participant was either corrected to 20/20 or had normal vision and passed the rVA assessment with high contrast optotypes in Experiment 1. It is unclear why the other three participants performed below threshold on the monitor based version of the assessment. However, this example illustrates how user characteristics influence image quality outcomes as outline in Chapter 3, Figure 3.

133

If we had not screened the participants prior to performing the experiment, we could assess the VE system more negatively than true performance. Further, we could falsely attribute failure of the VE system to the HMD, and not the beaded material as the results clearly suggest. When determining appropriate technology for use with clinical populations, accurate system assessment is imperative to ensure the safety of the patients and preserve the validity of the results. This study aids in defining how the properties of the custom made cubed retroreflective material actually augments user performance. In addition, by evaluating the system as a whole, we determined how the prototype HMPD and supporting technology limit or augment each other.

In sum, results from experiment 2 clearly show that there is a difference in rVA threshold between the two retroreflective materials as indicated by the respective point-spread functions. The differences in rVA between the materials are only for the low contrast targets. These results confirm that multiple tests of system parameters are necessary for accurate assessment of the VE system.

CHAPTER 7 STEREOACUITY AND DARK FOCUS FOR DIFFERENT HMDS

7.1 Introduction

Each HMD has associated parameters such as FOV and display size. As discussed in Chapter 5, these parameters assist in determining the optimal range which maximizes HMD performance and subsequently user performance. As results from Chapter 6 show, VE system components other than the HMD can contribute to degraded user performance. Thus, it is important to understand how well the VE technology chosen for a particular application actually performs in comparison to its specifications. In many cases, the information provided about a particular technology may be incomplete or overstated. Each of these factors plays a part in the overall evaluation of the integrated VE system, which is why system calibration is an important step in the VE design cycle.

The work reported here is an extension of Rolland, Gibson, & Ariely (1995) and Rolland, Meyer, Arthur, & Rinalducci (2002). As described in Chapter 4, these studies were conducted using bench prototypes that allowed for maximal control over the implementation of their experiments. Here we use the working head-mounted prototype HMPD-1. There are several aims of this research that are different from the original studies.

First, in previous experiments the Howard-Dolman task was used to establish predictive measures of user performance within the VE. In the current research, we employ the Howard-Dolman task to not only establish baseline stereoacuity for the individual participants, but also to monitor changes in real world performance due to the VE exposure. Second, we allow for free motion of the head to the degree that the person can adjust if they fell uncomfortable. Thus, we

135

are introducing depth cues that may not be represented in the original work. Third, as noted by the authors, individual differences in dark focus and dark convergence may degrade user performance when viewing images that mismatch accommodation and convergence eye mechanisms. Although Rolland et al. (2002) controlled for the mismatch between accommodation and convergence by placing the monocular optical images at the same depth as the rendered object, they did not measure the dark focus or dark convergence of their participants. In these initial experiments, we include a measure of dark focus and in subsequent experiments we measure dark focus. Fifth, only accuracy and precision of perceived depth expressed in mm were evaluated in previous studies. Here we convert the CE and VE_r values to measures of stereoaccuracy and stereoacuity, which are angular metrics, as outline in Chapter 5.

In the sections to follow we report the results of two experiments. The first of which assesses the stereoacuity and associated errors as outlined in Chapter 4 for the HMPD-1 over three different viewing distances, 800, 1500, and 3000 mm. The goal of this study is to evaluate the metrics chosen with regard to HMD and usability assessment. To evaluate the extendibility of the HMD assessment, we evaluate two off-the-shelf HMDs: the Canon Coastar™ and the VR6. Experimental design issues not covered in previous chapters (i.e., dark focus) will be addressed in their respective sections.

7.2 HMPD-1: Error Metrics and Dark Focus over Three Viewing Distances

7.2.1 Hypotheses

Given the current VE system setup, we can make the following predictions:

Hypothesis 1: We expect the precision errors (VE_r) to fall within the depth of resolution ranges for the HMPD-1 at each respective viewing distance. Stereoacuity is limited by the resolution of the display and should be roughly 246 arc seconds or 4.1 arc minutes.

Hypothesis 2: Given that we are aligning the optical image with the rendered image on the depth plane, we expect no shifts of dark focus across trials. Additionally, we do not expect adaptation changes as reflected in performance on the Howard-Dolman peg test.

Hypothesis 3: Given the exposure time in the virtual environment is less than 15 minutes per testing condition, we hypothesize that participants will experience lower cybersickness scores on the SSQ.

7.2.2 Method for Evaluating the HMPD-1 over Three Viewing Distances

Participants

This reach was approved by the Institutional Review Board (IRB) of the University of

Central Florida. The IRB protocol and consent form are listed in Appendix A. From a pool of 15

participants, five participants (all male, mean age = 29.8, SD = 5.26) were randomly chosen to

perform the experiment using the HMPD-1. The other participants performed a modified version

of the experiment reported in the next section. Each participant in this experiment was either

corrected for or had 20/20 vision. Glasses or contacts were worn during each part of the

experiment.

VE System

We used the HMPD-1 with a 52 degree FOV. The microdisplays for the HMD were off-the-shelf

Liquid Crystal Displays with a VGA resolution of 640x480. Software ran on a computer system

with Microsoft XP Pro and a dual processor graphics card. We used the full ARC display. Based

on the results reported in Chapter 6, only the Scotchlite^{TM} 3M Film Silver (Cubed) panels were

used in the experiment.

Experimental Design

Participants were randomly placed in a viewing distance sequence (e.g., 800, 1500,

3000). Only one participant's viewing distance order overlapped with another participant. This

order was 1500, 800, 3000. Each viewing distance was presented on a different day to minimize

order effects. Each participant performed the Howard-Dolman peg test three times during the

experiment, before VE exposure, after the first VE exposure, and after the second VE exposure. Each trial consisted of 10 repetitions. The response variables of CE, |CE|, VE_r and η were calculated from each test trial for the real world viewing condition. A repeated measures ANOVA was performed using only the |CE| and VE_r values.

Dark focus and dark convergence were also measured before VE exposure, after the first VE exposure, and after the second VE exposure. The metric for dark focus is Diopters. This metric was added to evaluate participants for changes after each VE exposure.

In addition to calculating the response variables of CE, |CE|, VE_r, and η, the % correct for evaluating whether the octahedron was in front of or behind the cylinder at the start of each stimulus presentation was computed. The octahedron appeared equal times to the right or to the left of the fixed target and an equal number of times in front or behind the object. Thus, the front or back judgments constituted a 2AFC design. Participants performed the Virtual Howard-Dolman two times in the experiment. There were a total of 32 responses per trial.

After each exposure in the VE, the participants also filled out the simulator sickness questionnaire (SSQ)[4], also listed in the Appendix X. The total sickness score and subscore for oculomotor effects were compared for each condition. Another questionnaire was administered to monitor general health and eye fatigue. This questionnaire is also listed in the appendix, but the results are not reported since the SSQ provided more information about participant health status.

---

[4] Thanks to Julie Drexler of RSK, inc for her assistance in scoring the SSQ data

<u>Procedure</u>

Prior to entering the VE testing area, the participants completed a general questionnaire on daily tasks performed by the participant which could produce eye strain or alter performance on the assessments. Once completed, the SSQ pretest was administered (Kennedy, Lane, Berbaum & Lilienthal, 1993). Participants were then asked to perform the Titmus Stereo Test. A passing score on the Circle test is considered as obtaining 6 out of the 11 judgments correct or 80 arc seconds. We used this value as a screening criterion for the experiment. The interpupillary distance (IPD) of each subject was measured using a pupillometer. This value was entered into the Virtual Howard-Dolman application to set the correct viewing position for each of the test stimuli.

The participants were then seated in front of the Howard-Dolman apparatus at the viewing distance that would be tested in the VE. Using a random number generator, the experimenter predetermined the initial position of the moveable rod for each of the 10 judgments for a single trial. The input to the random number generator was the 20 possible whole number positions in cm along the Howard-Dolman scale. A separate set of starting positions was given for each Howard-Dolman trial. Once the 10 judgments for the Howard-Dolman task were completed, the participants' dark focus was measured.

Dark focus was measured using a hand held stigmatoscope[5] as shown in Figure 46. The participant held the stigmatoscope over the eye of their preference and moved the telescoping section inward and out until the internal light point being viewed came into focus. The barrel of the stigmatoscope was graduated in Diopters. Once the participant focused the light, the

[5] Stigmatoscope, Dark Vergence Test, and Titmus test provided by RSK, inc

experimenter was able to read the Diopter value for that trial. There were four trials performed, two times with the barrel extended upon starting and two times with the barrel fully collapsed upon start.

Figure 45 Example of dark focus testing using the stigmatoscope

The participants were then seated to the selected viewing distance from the retroreflective material and fitted with the HMPD-1. Prior to the start of each Virtual H-D trial, the participants' visual acuity was tested. A string of 11 modified Landolt C optotypes each with a randomly placed gap position were displayed. The participants read off the gap directions for each string. A total of 5 strings were presented each representing 20.6. 16.5, 12.4, 8.2, and 4.1 arc minutes respectively. Based on the results presented in Chapter 6, the high contrast stimuli presented upon the cubed material should yield scores 100 % at each rVA level. Any score below 80% or 9 out of 11 correct should be indicative of negative VE exposure effects.

The participants were then shown a practice stimulus and given the input device, which is pictured in Figure 46. We set the two dials on the Shuttle Express to either move the octahedron using larger steps or finer steps. Rotating either dial to the right moved the object further away

141

from the participant on the Z axis. Likewise, tuning either dial to the left moved the object closer

to the observer. The resolution of each of the dials was set based on the work Rolland et al.

(2002). These values are given in Table 17.
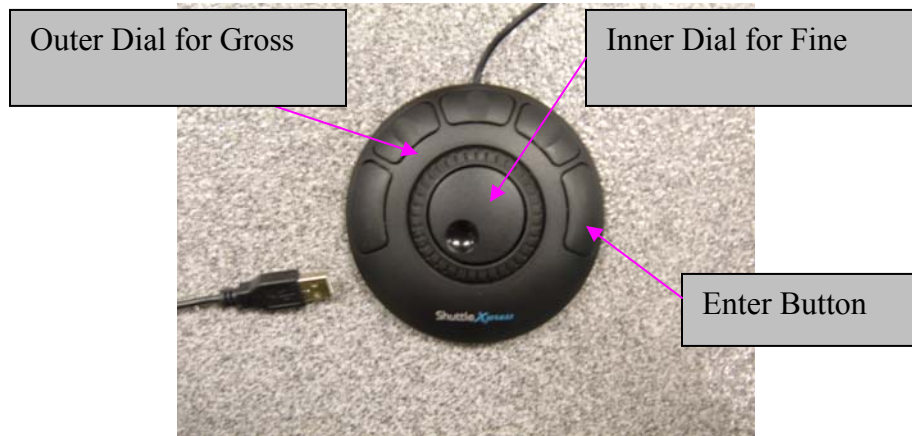
Figure 46 Shuttle dial used for moving the octahedron in virtual space

Table 17 Gross and fine adjustments for the input device

Viewing Distance (mm)	Gross Adjustment (mm)	Fine Adjustment (mm)
800	12	4
1500	25	6
3000	100	25

Once the participants were comfortable with moving the stimuli, they started the

experiment. The octahedron could appear anywhere within a ± 50 mm range, where the center of

this range was consider the nominal distance. Over the 32 repetitions, the octahedron appeared

equally in front or behind the fixed cylinder, as well as to the right and to the left of the cylinder.

Prior to moving the stimulus, the participants reported whether the octahedron appeared in front

of or behind the cylinder. The participants then moved the octahedron to the point of subjective

equality, where the octahedron and the cylinder appeared to be at the same distance. Once the

PSD was determined, they pressed the enter button shown in Figure 44. There was no time limit in responding. The program terminated upon completion of the assessment.

The HMPD-1 was removed from the participants and they filled out the SSQ post exposure questionnaire. They then were positioned in front of the Howard-Dolman apparatus and performed 10 PSE adjustments. Their dark focus was remeasured. This procedure was repeated for another sequence. Participants performed the experiment another two times at a different viewing distance.

7.2.3 Results of HMPD-1 over the Three Viewing Distances

In all experiments, the ANOVAs were performed using SPSS 11.5 and all graphs were constructed using Minitab 14. Table 18 gives the CE means and the standard deviations for each participant at each viewing distance and for both testing environments. Overall, participants show a slight positive bias on both the Howard-Dolman (H-D) task and the Virtual Howard-Dolman (V-HD) task. Variability in responding increased over distance for each testing environment; however, variability was the greatest in the V-HD condition.

Figure 47 illustrates individual differences in bias across the two testing environments and viewing distances. Some participants consistently underestimate the distance of the cylinder, as evidenced by their negative scores. Others perceive the cylinder to be further away and have a positive bias. Overall, the means for H-Dolman and V-H Dolman are similar to each other over all trials and reflect that the participants across all conditions are accurate in aligning the octahedron with respect to the nominal distance of the cylinder.

Table 18 Constant errors (mm) per participant (HMPD-1) as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman						Virtual-Dolman			
		M			SD			M		SD	
		Pre	Post1	Post2	Pre	Post1	Post2	Pre	Post	Pre	Post
800	1	2.5	1.7	0.7	1.8	3.2	2.8	-7.7	-4.6	18.1	12.1
	2	-0.1	1.9	0.8	1.8	3.7	1.1	0.0	0.0	3.0	2.8
	3	1.3	1.0	1.7	2.2	2.5	2.4	2.3	3.6	9.4	11.4
	4	0.3	1.6	0.3	1.3	1.4	1.6	4.6	3.4	10.9	9.7
	5	-0.4	-0.5	1.6	1.0	1.0	1.6	9.9	8.5	6.1	7.0
	$\overline{X}$	0.7	1.1	1.0				1.8	2.2		
	SD	1.2	1.0	0.6				6.5	4.9		
1500	1	1.2	-0.9	2.5	2.8	3.3	1.4	-4.8	-0.5	25.9	28.9
	2	0.5	7.0	3.2	3.5	10.5	2.9	-4.6	-5.5	13.5	18.5
	3	10.7	5.9	6.6	9.1	5.7	1.1	8.5	2.5	23.3	21.9
	4	-3.6	-2.1	0.6	2.6	3.2	1.8	12.5	5.9	17.8	15.6
	5	-0.1	1.0	1.4	2.1	1.7	3.2	8.8	2.1	29.9	29.9
	$\overline{X}$	1.7	2.2	2.9				4.1	0.9		
	SD	5.3	4.1	2.3				8.2	4.2		
3000	1	-0.1	-0.5	1.2	4.5	3.9	4.7	-11.2	-2.8	24.9	31.6
	2	2.8	0.3	0.8	6.6	1.2	1.6	-0.1	1.9	22.3	18.3
	3	10.7	9.4	9.1	4.6	5.9	6.6	3.7	2.8	29.4	26.4
	4	5.2	0.8	-1.3	5.3	6.1	6.4	-6.9	-11.8	33.9	28.9
	5	4.0	0.9	-8.7	3.7	5.6	4.6	19.8	18.9	32.0	35.0
	$\overline{X}$	4.5	2.2	0.2				1.0	1.8		
	SD	4.0	4.1	6.4				12.0	11.2		
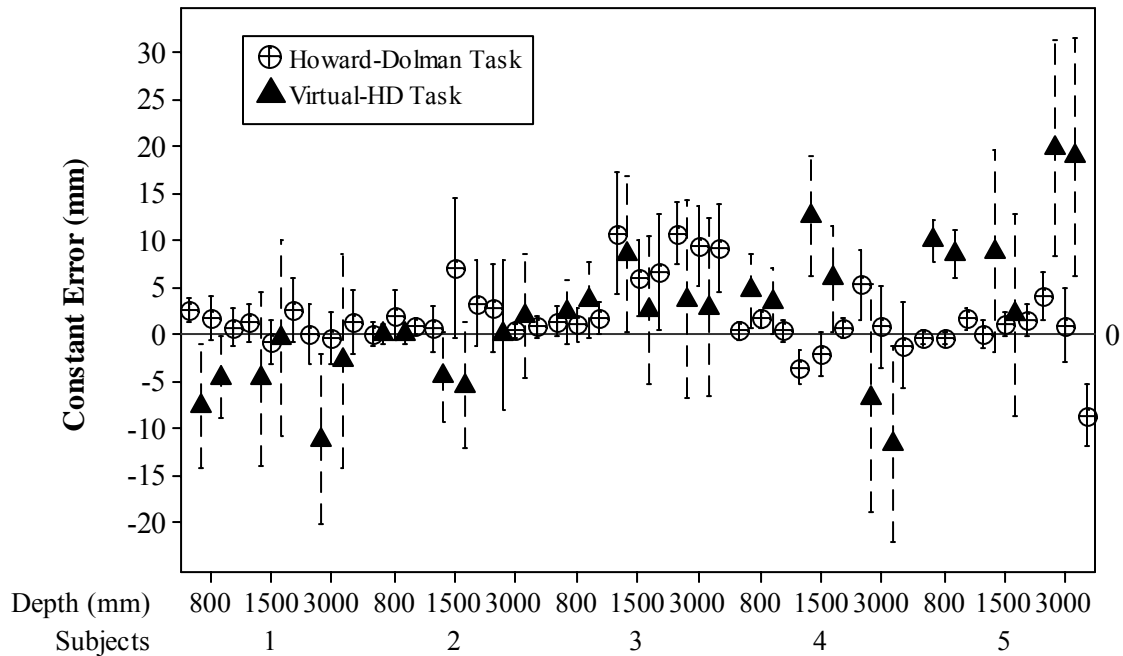
Figure 47 CE for each subject in the HMPD-1 condition at each test viewing distance (800, 1500, and 3000 mm) for the Howard-Dolman task and the Virtual-HD task. The results are ordered in the same manner they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). Individual differences in accuracy are illustrated along with their biases. The center points represent the mean of the signed deviations between the participant's responses and the nominal depth (800, 1500, or 3000 mm). The error bars represent the 95% Confidence Interval of the means.

Initial inspection of the V-HD data showed that the distribution was positively skewed and positively kurtotic. In comparison to a normal distribution for which the repeated measures ANOVA was designed, the data are shifted to the left and are more peaked than one would expect. Because the data do not appear to be from a normal distribution, other assumptions of the repeated measures ANOVA must be met before this statistic can be used to evaluate the data. A

145

test for additivity was performed to check that the participants' responses were consistent across all levels of viewing distances. Also a check for homogeneity of variance was performed to confirm that variance across each factor was similar. Of the two tests, the test to ensure additivity is more important since nonadditivity in a repeated measures design leads to lower power. Given the small sample size, the analysis is already at low power.

Given that both tests failed, the data were transformed using the following logarithmic transformation:

Let Y = Raw test score

Y' = Transformed test score

$$Y' = Log(Y + 1) \qquad\qquad (24)$$

The log transformed Y' values for both the |CE| and VEr were used in the repeated measures ANOVA. The log transformed scores for the |CE| are shown in Table 19. The geometric means are also shown. Inspection of the results show that compared to the pretest at 1500 mm, participants became more accurate on their posttests for the VE condition, pretest (M = 8.32 mm, 95% CI = 5.22 to 17.78) and posttest (M = 3.71 mm, 95% CI = 1.72 to 8.07). The results are graphically displayed in Figure 48.

Table 19 Log transformed absolute constant errors |CE| for each participant (HMPD-1) and group means as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman			Virtual-Dolman	
		Log			Log	
		Pre	Post1	Post2	Pre	Post
800	1	0.54	0.43	0.23	0.94	0.75
	2	0.04	0.46	0.26	0.01	0.01
	3	0.36	0.30	0.43	0.52	0.66
	4	0.11	0.41	0.11	0.75	0.65
	5	0.15	0.18	0.41	1.04	0.98
	$\overline{X}$	0.24	0.36	0.29	0.65	0.61
	SD	0.21	0.12	0.13	0.41	0.36
Geometric	$\overline{X}$	1.73	2.29	1.94	4.46	4.07
1500	1	0.34	0.28	0.54	0.76	0.18
	2	0.18	0.90	0.62	0.75	0.81
	3	1.07	0.84	0.88	0.98	0.55
	4	0.66	0.49	0.20	1.13	0.84
	5	0.04	0.30	0.38	0.99	0.48
	$\overline{X}$	0.46	0.56	0.53	0.92	0.57
	SD	0.41	0.29	0.26	0.16	0.27
Geometric	$\overline{X}$	2.88	3.63	3.39	8.32	3.71
3000	1	0.04	0.18	0.34	1.09	0.59
	2	0.58	0.11	0.26	0.05	0.46
	3	1.07	1.02	1.00	0.67	0.58
	4	0.79	0.26	0.36	0.90	1.11
	5	0.70	0.28	0.99	1.32	1.30
	$\overline{X}$	0.64	0.37	0.59	0.81	0.81
	SD	0.38	0.37	0.37	0.48	0.37
Geometric	$\overline{X}$	4.36	7.41	3.89	6.46	6.46
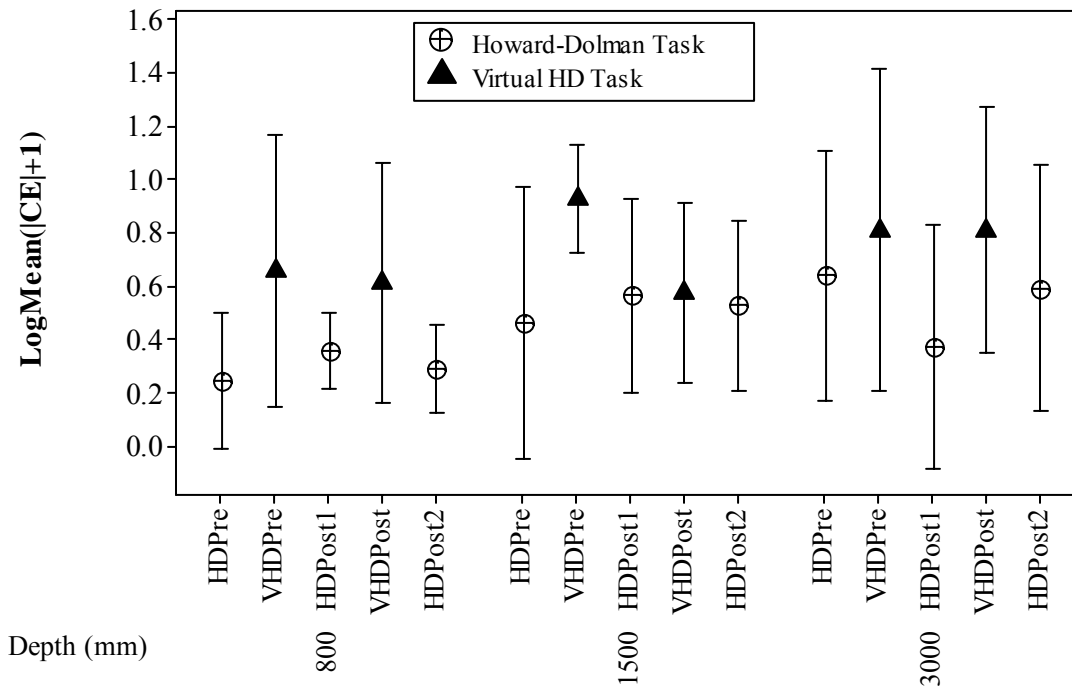
Figure 48 Log transformed means for the absolute constant error |CE| for testing condition, Howard-Dolman or Virtual-HD, at each viewing distance (800, 1500, and 3000 mm). Participants were wearing the HMPD. The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). The center points represent the log mean of the |CE| for the grouped means. The error bars represent the 95% Confidence Interval of the log means.

Visual inspection of the Figure 48 suggests that there is no effect of Trial or Viewing Distance on performance in the H-D condition. A repeated measures ANOVA (3 Viewing Distance x 3 Trials) confirmed that there was no significant effect of Viewing Distance, $F(2, 8) =$ 2.17, $p = .177$, $\rho_I = .189$, or Trial, $F(2, 8) = .102$, $p = .904$, $\rho_I = .22$ for the H-D task. The repeated measures ANOVA (3 Viewing Distance x 2 Trials) for the V-HD task showed a significant interaction between Viewing Distance and Trial, $F(2, 8) = 5.77$, $p = .028$, $\rho_I = .61$. The interaction plot for this result is shown in Figure 49.

148

The plot suggests that the difference in means between the pre and post test for the viewing distance of 1500 mm is significant. A paired samples t-test was performed on the pre and post tests at each depth. Only the pre and posttests for the 1500 mm viewing distance was significant, $t(4) = 3.09$, $p = .037$.
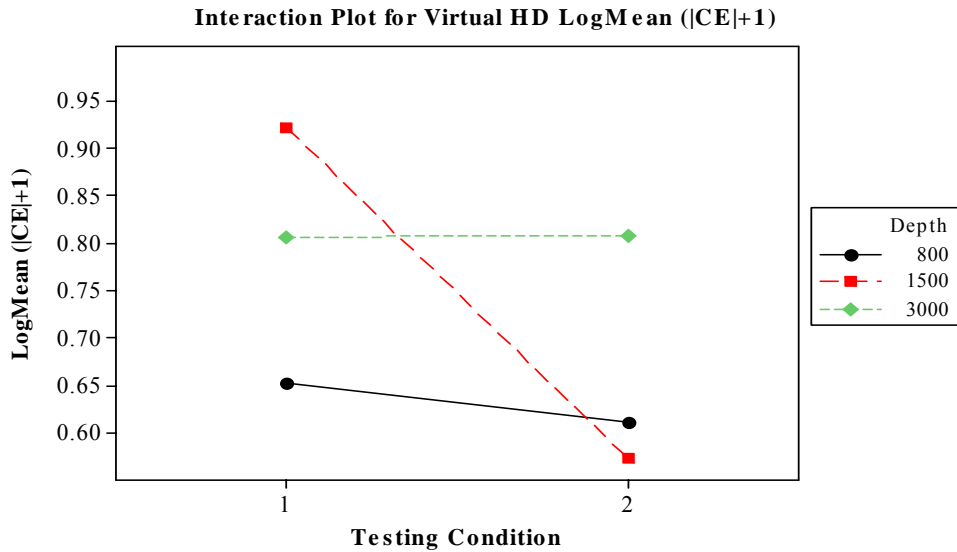
Figure 49 Log transformed means for the Virtual HD for the Pre and Post-testing conditions at each viewing distance (800, 1500, and 3000 mm) are represented illustrating the interaction between viewing distance and trial. Participants in the HMPD-1 condition were significantly more accurate in their posttest performance at the 1500 mm viewing distance. There were no Pre and Post differences in accuracy at the 800 or the 3000 mm viewing distances.

The results were pooled over Trial, but not Viewing Distance, to determine if the participants were more accurate in performing the H-D task than they were performing the V-HD task. A repeated measures ANOVA (3 Viewing Distances x 2 Environments) showed that there was no main effect of Environment, $F(1, 4) = 3.51$, $p = .134$, $\rho_I = .56$. There was no significant interaction. These results are presented graphically in Figure 50.
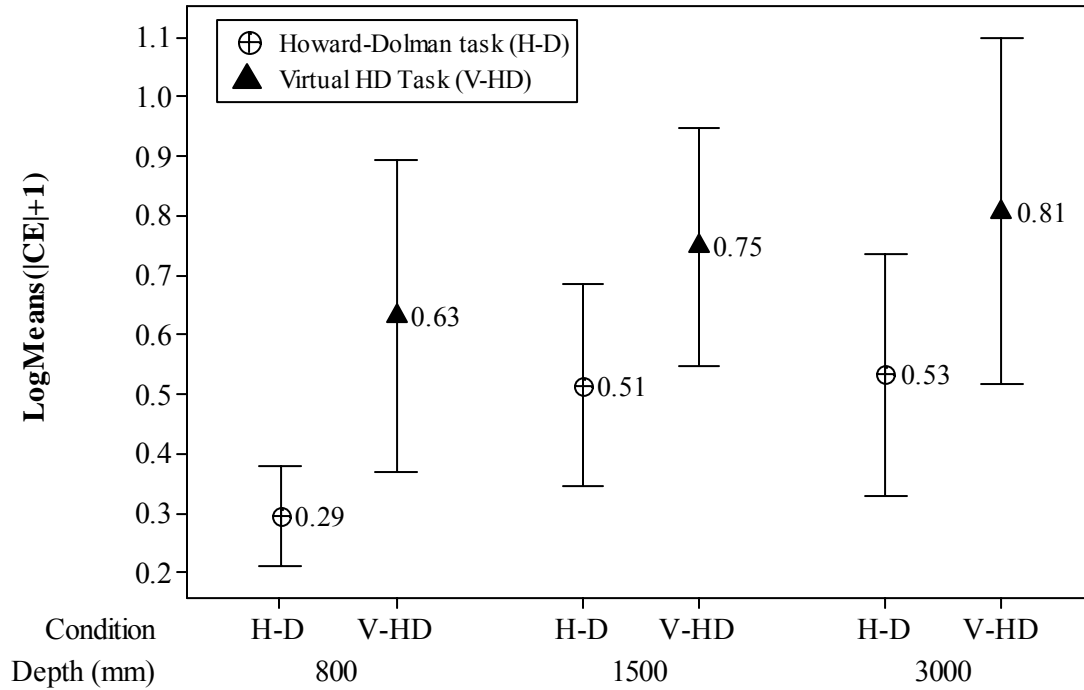
Figure 50 Combined log means of the absolute constant error |CE| for the different viewing distances (800, 1500, 3000 mm) representing the Howard-Dolman task and the Virtual-HD task. The center points signify the log mean of |CE| pooled over all subjects. The error bars represent the 95% Confidence Interval of the log means.

Table 20 shows the untransformed VE_r for each of the participants over viewing distances and testing environments. The results are presented graphically in Figure 51. VE_r is a measure of precision, thus the data show that participants are less consistent when performing the V-HD than when they perform the H-D task. VE_r also corresponds to the resolvable depth computed for pixel resolution as shown in Table 21. As Table 21 shows all values of the VE_r are accurate to within 1/4 to 1/10 of a pixel resolution. This result demonstrates the interdependence of the computer graphics and the HMD with respect to user performance. The anti-aliasing techniques applied to the rendering process of the 3D objects may assist participants in performing above expectation given the HMD parameters. To perform the repeated measures ANOVA analysis the data were log transformed using Equation 24. The log transformed data is shown in Figure 22. The same analysis was carried out as with the |CE|.

Table 20 Untransformed variable error VE_r (mm) values per participant and group means as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman Task Magnitude			Virtual-Dolman Task Magnitude	
		Pre	Post 1	Post 2	Pre	Post
	1	1.75	3.07	2.61	17.80	11.86
800	2	1.70	3.53	1.08	2.98	2.73
	3	2.05	2.41	2.28	9.23	11.18
	4	1.27	1.36	1.49	10.73	9.57
	5	0.92	0.92	1.56	5.99	6.94
	$\overline{X}$	1.54	2.26	1.80	9.35	8.46
	SD	0.45	1.11	0.63	5.59	3.72
1500	1	2.64	3.14	4.57	25.45	28.49
	2	3.35	9.97	6.08	13.25	18.18
	3	8.66	5.45	8.22	22.91	21.52
	4	2.46	3.08	1.43	17.56	15.39
	5	1.97	1.61	2.24	29.43	29.46
	$\overline{X}$	3.82	4.65	4.51	21.72	22.61
	SD	2.75	3.27	2.78	6.40	6.22
3000	1	4.25	3.72	4.45	24.50	31.14
	2	6.27	1.10	1.54	21.97	17.99
	3	4.38	5.59	6.30	28.92	26.01
	4	5.02	5.79	6.07	33.32	28.46
	5	3.46	5.28	4.41	31.53	34.49
	$\overline{X}$	4.68	4.30	4.55	28.05	27.62
	SD	1.05	1.96	1.90	4.75	6.24
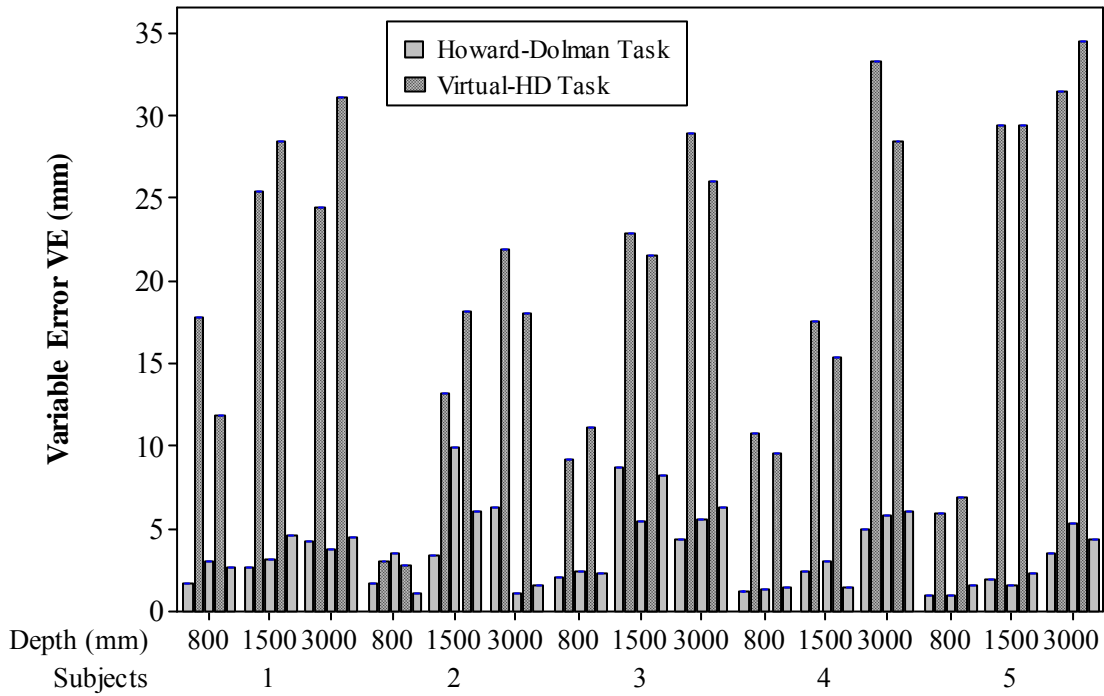
Figure 51 Variable error (VE$_r$) values for each participant performing the Howard-Dolman task or the Virtual-HD task at the different viewing distances (800, 1500, 3000 mm). The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). The height of the bar represents the magnitude of the VE$_r$. The log transform of these values were used in the inferential statistics analysis. The actual values are shown in Table 19.

Table 21 Predicted resolvable depth in mm for the HMPD-1 as a function of pixel size and viewing distance.

Pixel Size	Depth (mm)		Resolvable Depth (mm)
		800	18.36
1		1500	65.93
		3000	267.34
		800	9.18
1/2		1500	32.95
		3000	133.47
		800	4.59
1/4		1500	16.47
		3000	66.71
		800	1.84
1/10		1500	6.59
		3000	26.68

A repeated measures ANOVA (3 Viewing Depths x 3 Trials) performed on the H-D data showed a significant main effect of Viewing Distance, $F(2, 8) = 5.72$, $p = .029$, $\rho_I = .61$. Post Hoc analysis using the Bonferroni post hoc criterion for significance confirmed that the participants are significantly more precise at the 800mm viewing distance (M= 2.77 mm, 95% CI = 2.11 to 3.61) than at 1500 mm (M = 4.74 mm, 95% CI = 2.68 to 7.97) or 3000 mm (5.22 mm, 95%CI = 3.75 to 7.24). This result is shown graphically in Figure 52. The results are identical for the V-HD task.

A repeated measures ANOVA (3 Viewing Distances x 2 Trials) was conducted for the V-HD data. The results showed a main effect of depth, $F(2, 8) = 24.56$, $p = .001$, $\rho_I = .89$. Post Hoc analysis using the Bonferroni post hoc criterion for significance confirmed that the participants are significantly more precise at the 800mm viewing distance (M= 8.91 mm, 95% CI = 4.57 to 17.37) than at 1500 mm (M =22.38 mm, 95% CI = 15.85 to 30.90) or 3000 mm (28.18 mm,

95%CI = 22.38 to 35.48). This result is shown graphically in Figure 53. The results are identical

for the V-HD task. There was no effect of Trial.

Table 22 Log transformed variable errors (VE$_r$) for each participant and group means as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman Log			Virtual-Dolman Log	
		Pre	Post1	Post2	Pre	Post
800	1	0.44	0.61	0.56	1.27	1.11
	2	0.43	0.66	0.32	0.60	0.57
	3	0.48	0.53	0.52	1.01	1.09
	4	0.36	0.37	0.40	1.07	1.02
	5	0.28	0.28	0.41	0.84	0.90
	$\overline{X}$	0.40	0.49	0.44	0.96	0.94
	SD	0.08	0.16	0.10	0.25	0.22
Geometric	$\overline{X}$	2.51	3.09	3.63	9.12	8.71
1500	1	0.56	0.62	0.75	1.42	1.47
	2	0.64	1.04	0.85	1.15	1.28
	3	0.99	0.81	0.96	1.38	1.35
	4	0.54	0.61	0.39	1.27	1.21
	5	0.47	0.42	0.51	1.48	1.48
	$\overline{X}$	0.64	0.70	0.69	1.34	1.36
	SD	0.20	0.24	0.24	0.13	0.12
Geometric	$\overline{X}$	4.36	5.01	4.90	21.88	22.91
3000	1	0.72	0.67	0.74	1.41	1.51
	2	0.86	0.32	0.40	1.36	1.28
	3	0.73	0.82	0.86	1.48	1.43
	4	0.78	0.83	0.85	1.54	1.47
	5	0.65	0.80	0.73	1.51	1.55
	$\overline{X}$	0.75	0.69	0.72	1.46	1.45
	SD	0.08	0.21	0.19	0.07	0.10
Geometric	$\overline{X}$	5.62	4.90	5.25	28.84	28.18

Howard-Dolman Task
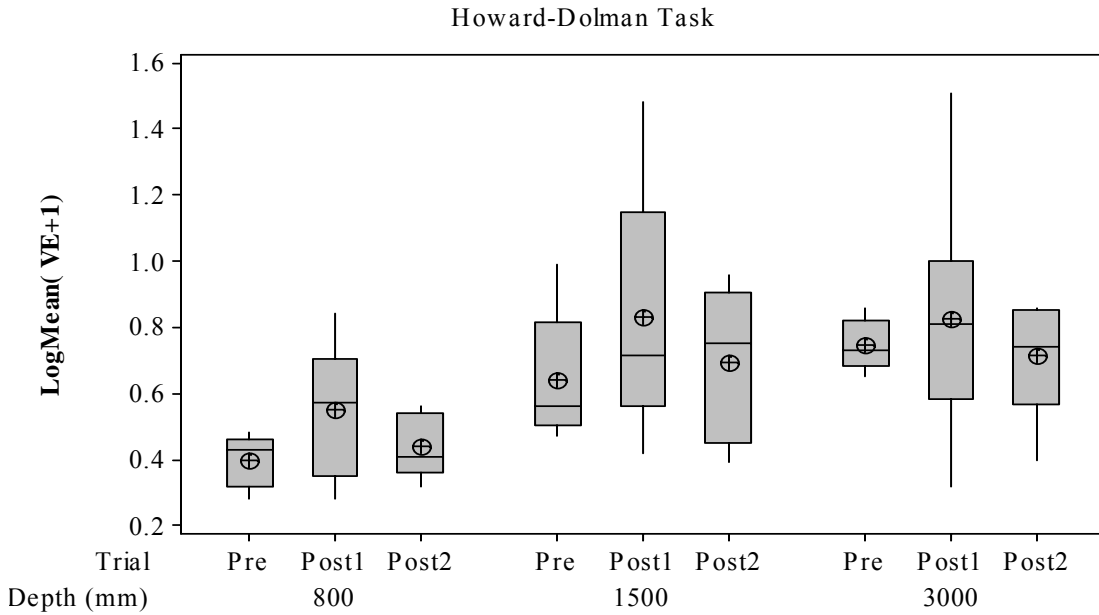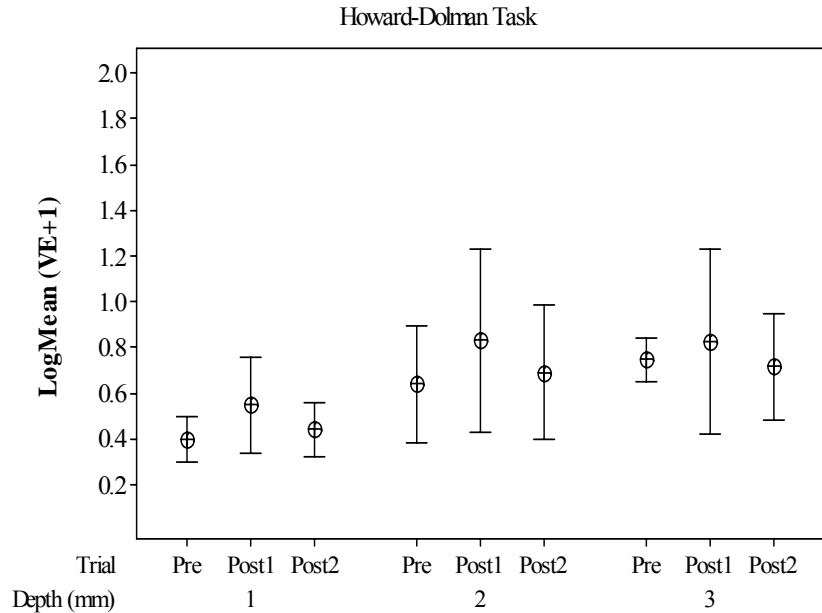
Howard-Dolman Task

Figure 52 Log variable error (VE$_r$) means (bull's eye), medians (line), and interquartile ranges for the participants' grouped responses on the Howard-Dolman task performed at the different depths (800, 1500, 3000 mm) and on the separate trials (Pre, Posttest 1, and Posttest 2). Precision is significantly lower for the 800 mm viewing distance than at 1500 or 3000 mm. There are no significant differences in precision between 1500 mm and 3000 mm viewing depths.

Virtual-HD Task
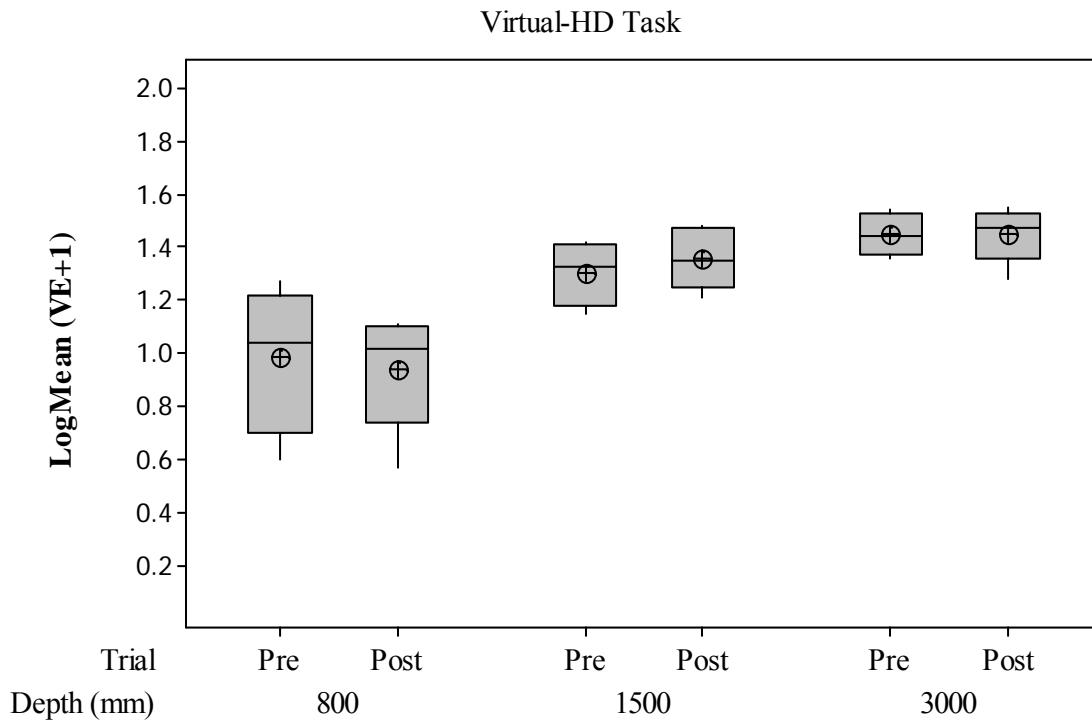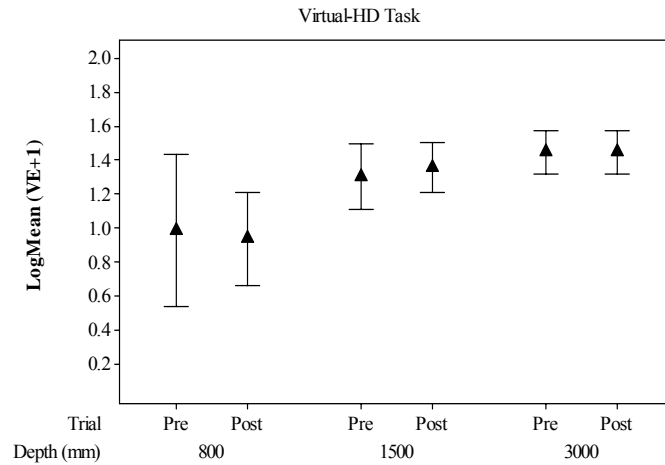
Virtual-HD Task

Figure 53 Log variable error (VE$_r$) means (bull's eye), medians (line), and interquartile ranges for the participants' grouped responses on the Virtual-HD task performed at the different viewing distances (800, 1500, 3000 mm) and on the separate Pre and Posttest trials. The mean averages illustrate the increase in VE$_r$ error between the 800 mm and 1500 or 3000 mm. There are no significant differences between the Pre and Post tests at any depth.

Because there was only a significant effect of Viewing Distance, the data was pooled over Trials, but not viewing distance. A repeated measures ANOVA (3 Viewing Distances x 2 Environments) shows a main effect for Viewing Distance, $F(2, 8) = 22.36$, p. $= .001$, $\rho_I = .88$, and for Environment, $F(1, 4) = 94.09$, $p = .001$, $\rho_I = .97$. Participants were more precise in judging the distance of the octahedron at the 800 mm Viewing Distance ($M = 4.96$ mm, 95% CI $= 3.37$ to 7.31) than at 1500 mm ($M = 10.23$, 95% CI $= 7.78$ to 13.70), or 3000 mm ($M = 12.12$, 95% CI $= 9.31$ to 15.88). Overall, participants' responses were more consistent on the H-D task ($M = 4.09$ mm, 95% CI $= 12.56$ to 25.29) than on the V-HD task (17.78 mm, 95% CI $= 12.56$ to 25.29. These results are shown graphically in Figure 54.
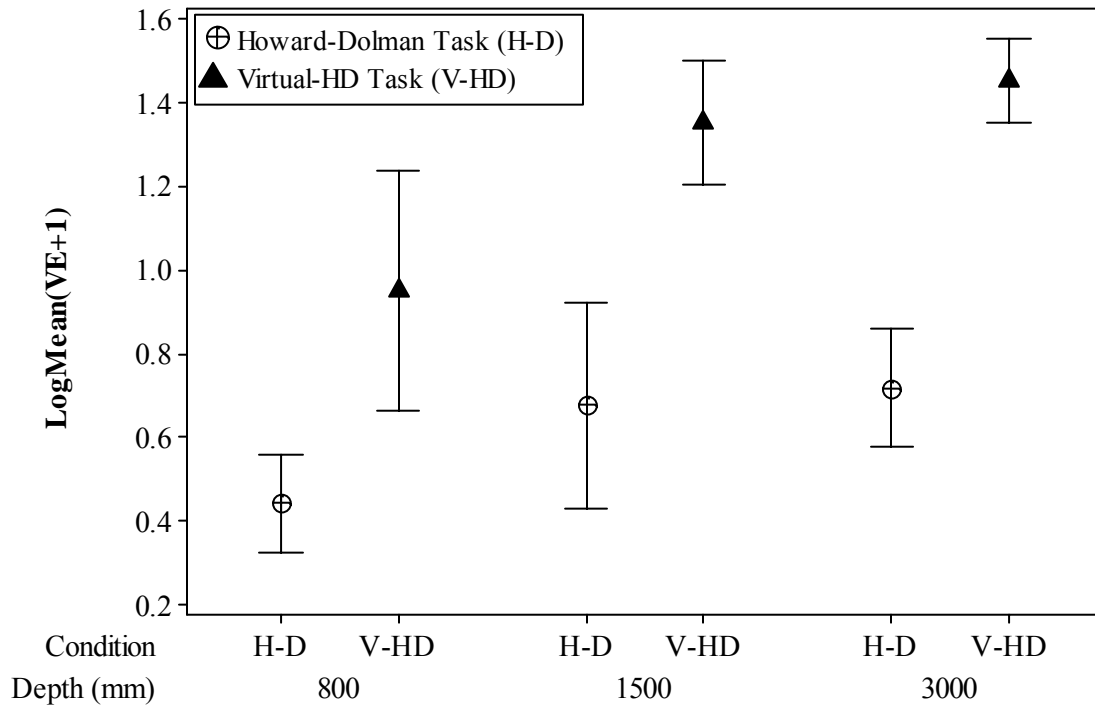
Figure 54 Combined log means of the variable error VE_r for the different viewing distances (800, 1500, 3000 mm) representing the Howard-Dolman task (H-D) and the Virtual-HD task (V-HD). The center points signify the log mean of the variable error VE_r pooled over all subjects. The error bars represent the 95% Confidence Interval of the log means. The graph shows that there is more VE_r over the different viewing conditions when participants perform the Virtual H-D task than when they perform the H-D task.

Table 23 shows the stereoaccuracy in arc seconds for each participant for each viewing distance and testing environment. The reference line in the graph is taken as the fixation point or the nominal distance from the observer. If the objects are adjusted to the same distance, stereoaccuracy should be zero. For the H-D task, stereoaccuracy shows that participants are able to judge the target distance well with respect to the fixed cylinder location at all depths. They are the best, however, at judging the distance of the octahedron closer to the normal testing distance for the H-D task at 3000 mm. As with the H-D task, participants are not as accurate at judging distances at the viewing distance of 800 mm as they are at 3000 mm. Figure 55 graphically displays the stereoaccuracy for each of the observers, while figure 56 shows the grouped means for each viewing distance and testing environment.

Table 23 Stereoaccuracy (η_{ce}) in arc seconds per Participant (HMPD-1) as a Function of Depth for the Howard-Dolman Task and Virtual-Dolman Task

Depth (mm)	Subjects	Howard-Dolman M			Howard-Dolman SD			Virtual-Dolman M		Virtual-Dolman SD	
		Pre	Post1	Post2	Pre	Post1	Post2	Pre	Post	Pre	Post
800	1	32.2	21.9	9.0	23.7	41.7	35.4	-161.3	-10.6	379.0	606.4
	2	-1.4	26.7	11.3	25.2	52.4	16.0	-0.7	-0.4	69.3	63.5
	3	16.2	12.5	21.2	27.0	31.7	30.1	46.2	73.3	190.5	230.7
	4	3.9	20.6	3.9	17.2	18.4	20.2	95.5	71.8	228.3	203.8
	5	-5.2	-6.5	20.9	12.6	12.7	21.5	210.6	180.0	129.5	149.9
	$\overline{X}$	9.1	15.0	13.3				38.1	62.8		
	SD	15.2	13.1	7.6				136.3	76.4		
1500	1	5.4	-4.1	11.3	12.6	15.0	21.8	-28.7	-6.3	154.1	172.5
	2	2.5	34.7	15.8	17.5	52.0	31.7	-29.8	-35.8	87.6	120.2
	3	47.0	25.9	29.0	40.1	25.2	38.1	48.9	14.5	134.4	126.2
	4	-16.3	-9.5	2.7	11.7	14.7	6.8	74.6	35.0	106.3	93.2
	5	-0.5	4.6	6.4	9.6	7.8	10.9	53.0	12.4	180.9	181.1
	$\overline{X}$	7.6	10.3	13.1				23.6	4.0		
	SD	23.6	19.2	10.2				49.2	26.6		
3000	1	-0.1	-0.6	1.6	5.8	5.1	6.1	-16.7	-4.2	37.1	47.1
	2	4.0	0.4	1.1	9.3	1.6	2.3	-0.2	3.1	36.3	29.7
	3	13.4	11.8	11.4	5.8	7.4	8.3	5.3	4.1	42.4	38.2
	4	6.7	1.0	-1.7	6.8	7.9	8.3	-10.3	-17.5	50.4	43.1
	5	5.3	1.2	-11.4	4.8	7.3	6.1	29.9	28.5	48.5	53.0
	$\overline{X}$	5.8	2.8	0.2				1.6	2.8		
	SD	4.9	5.1	8.2				18.0	5.1		
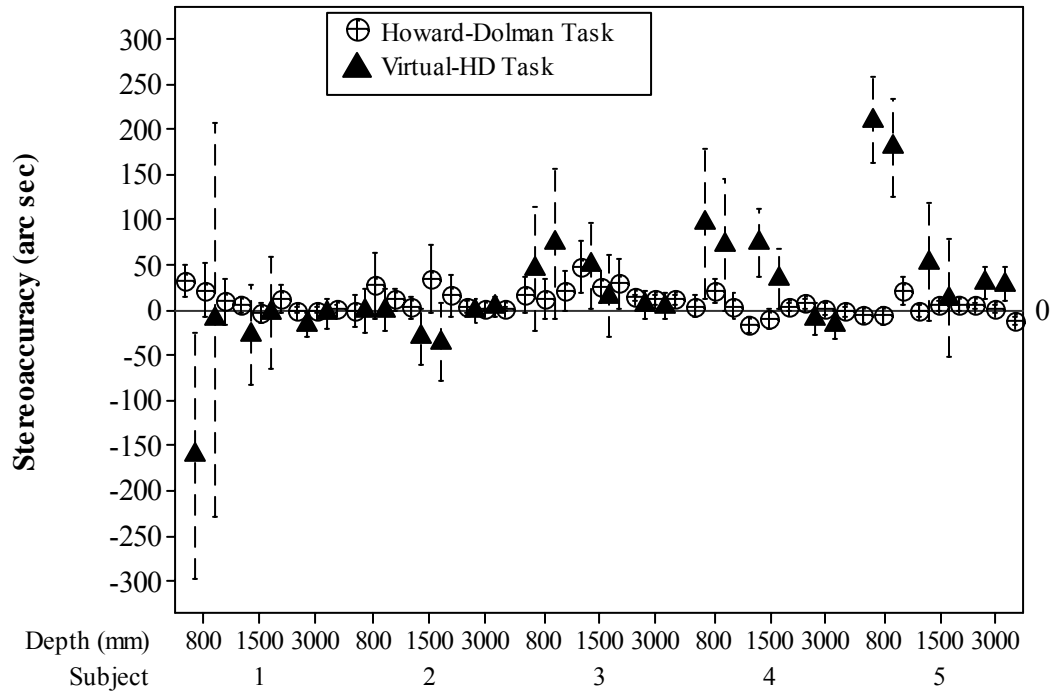
Figure 55 Stereoaccuracy η_{ce} for each subject at each viewing distance (800, 1500, and 3000 mm) for the Howard-Dolman task and the Virtual-HD task. Participants performed the task wearing the HMPD-1. The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). Individual differences in accuracy are illustrated along with their biases. The center points represent the mean of stereoaccuracy as calculated from the CE, viewing depth (800, 1500, or 3000 mm), and interpupillary eye distance (IPD) of the observer. The error bars represent the 95% Confidence Interval of the means.

162
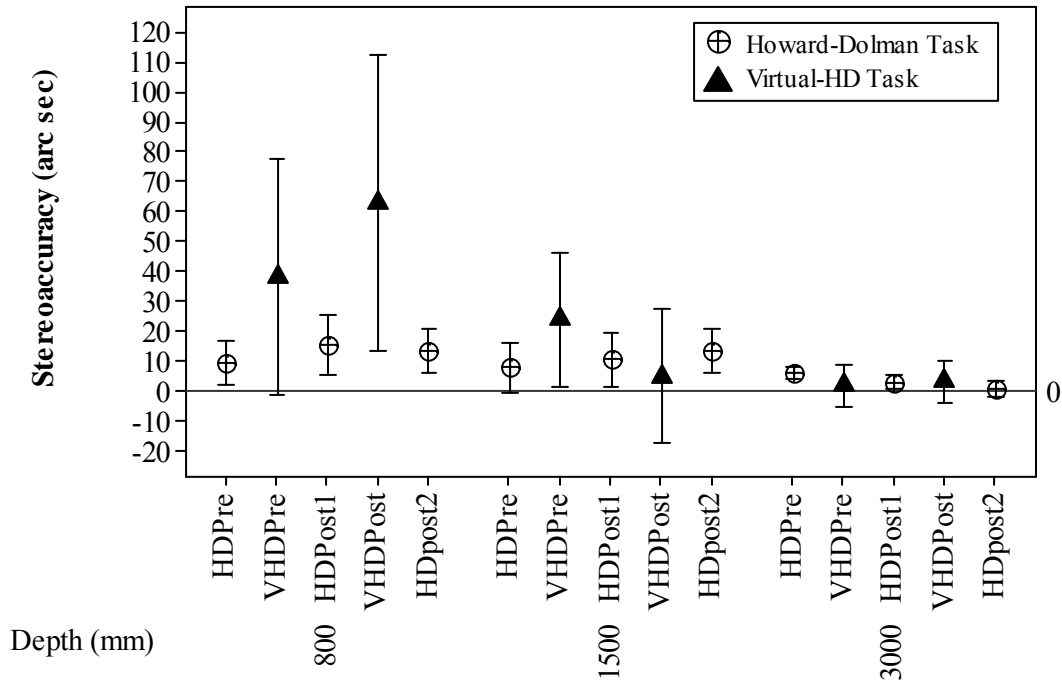
Figure 56 Stereoaccuracy η_{ce} for each viewing distance (800, 1500, and 3000 mm) for the Howard-Dolman task and the Virtual-HD task. Participants performed the task wearing the HMPD-1. The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). Pooled differences in stereoaccuracy are illustrated along with their biases. The center points represent the mean of stereoaccuracy as calculated from grouping data over participants for each viewing depth (800, 1500, or 3000 mm). The error bars represent the 95% Confidence Interval of the means. A zero value for stereoaccuracy suggests that the participants were able to perfectly align the test objects.

163

The stereoacuity thresholds for each observer are given in Table 24. As shown, stereoacuity varies greatly over trials for both testing environments per individual. Figure 57 shows the overall stereoacuity for each viewing depth and testing environment. Although the data is noisy, the average stereoacuities for all participants are below the predicted threshold of the HMPD-1, which is 246 arc seconds. The reference line at 20 arc seconds is the standard real world viewing stereoacuity value. Participants' performance on the H-D task is close to this value at each viewing distance. Figure 57 clearly shows that stereoacuity changes with distance in the VE condition. This shift in stereoacuity is not as great for the H-D task, but is significant, $F(2, 28) = 19.65$, $p < .001$, $\rho_I = .88$. This result would not be expected if the H-D task was devoid of all monocular cues (Wong, Woods, & Peli, 2002). However, given that the stereoacuity value measured is within appropriate limits, the H-D task seems appropriate for its use with HMD testing.

Table 24 Stereoacuity (η_{ve}) in arc seconds per Participant (HMPD-1) as a Function of Depth for the Howard-Dolman Task and Virtual-Dolman Task.

Depth (mm)	Subjects	Howard-Dolman Task M			Virtual-Dolman Task M	
800	1	22.51	39.53	19.98	372.99	248.48
	2	23.93	49.75	34.36	68.24	62.46
	3	25.63	30.08	16.10	187.46	227.09
	4	16.35	17.48	14.32	224.70	200.54
	5	11.99	12.06	43.44	127.45	147.56
	$\overline{X}$	20.08	29.78	25.64	196.17	177.23
	SD	5.72	15.48	12.69	115.37	74.43
1500	1	11.96	14.25	20.69	151.64	169.77
	2	16.60	49.35	30.09	86.26	118.31
	3	38.04	23.93	36.13	132.32	124.26
	4	11.14	13.96	6.47	104.64	91.71
	5	9.08	7.42	10.33	178.06	178.28
	$\overline{X}$	17.36	21.78	20.74	130.58	136.46
	SD	11.88	16.50	12.61	36.53	36.53
3000	1	5.50	4.81	5.75	36.50	46.38
	2	8.61	1.51	2.11	35.75	29.27
	3	5.75	7.34	8.27	41.76	37.56
	4	6.39	7.38	7.72	49.64	42.40
	5	4.48	6.83	5.70	47.69	52.16
	$\overline{X}$	6.15	5.57	5.91	42.27	41.56
	SD	1.54	2.50	2.42	6.32	8.71
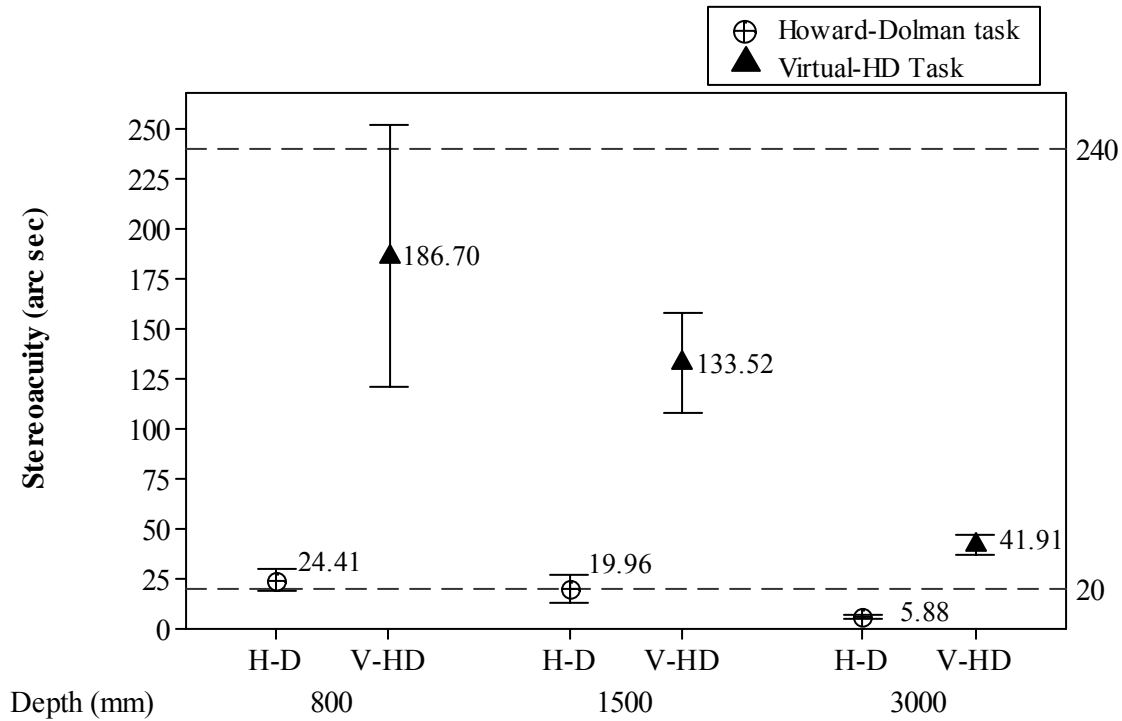
Figure 57 Overall stereoacuity η_{ve} means and 95% Confidence Interval for the participants'
(HMPD-1) grouped responses on the Howard-Dolman and Virtual-HD task performed at the
different depths (800, 1500, 3000 mm). The mean averages illustrate the change in stereoacuity
over the different viewing depths. The reference line at 20 arc seconds represents average
responding in the population at large. The reference line at 120 arc seconds represents the
predicted stereoacuity based on the HMPD parameters.

Table 25 gives the dark focus results for each Viewing Distance and Trial. The dark focus values are consistent across Trial and Viewing Distance. The only notable differences are for the first posttest at 1500 mm. All participants experienced a downward shift in dark focus as noted in Figure 58. A paired sampled t-test performed between the Pretest and Posttest for the 1500 mm Viewing Distance, was significant, $t(4) = 4.88$, $p = .008$. The participants who experienced greater shifts (e.g., participant 2) did not recover to baseline during the course of the experiment.

Table 26 shows the oculomotor subset scores for the SSQ given in both post sessions. Typically scores above 20 indicate that the participant is experiencing some effect of the VE (Kennedy, Kennedy, & Drexler, 2006). Participants 2 and 4 are suggestively high on this SSQ metric. However, these participants actually show a significant improvement in performance not seen at any other viewing distance. Since this is a repeated measures design where all the participants experience all conditions and thus act as their own control, the result may be more than spurious. However, a larger sample size is needed to quantify these effects further.

Table 25 Mean dark focus in Diopters per participant (HMPD-1) as a function of depth and testing condition

Depth (mm)		Dark Focus M		
		Pre	Post1	Post2
	1	1.7	1.6	1.9
800	2	3.0	2.9	2.8
	3	1.2	1.1	1.4
	4	0.9	1.3	1.0
	5	0.1	0.6	0.4
	$\overline{X}$	1.4	1.5	1.5
	SD	1.1	0.9	0.9
1500	1	1.2	1.1	1.5
	2	1.4	0.9	0.6
	3	1.4	1.0	1.1
	4	1.4	0.8	1.2
	5	1.7	1.2	1.2
	$\overline{X}$	1.4	1.0	1.1
	SD	0.2	0.1	0.3
3000	1	1.6	1.7	1.9
	2	1.1	1.0	0.9
	3	1.1	1.4	1.7
	4	1.3	0.6	2.1
	5	1.4	1.8	1.6
	$\overline{X}$	1.3	1.3	1.6
	SD	0.2	0.5	0.4
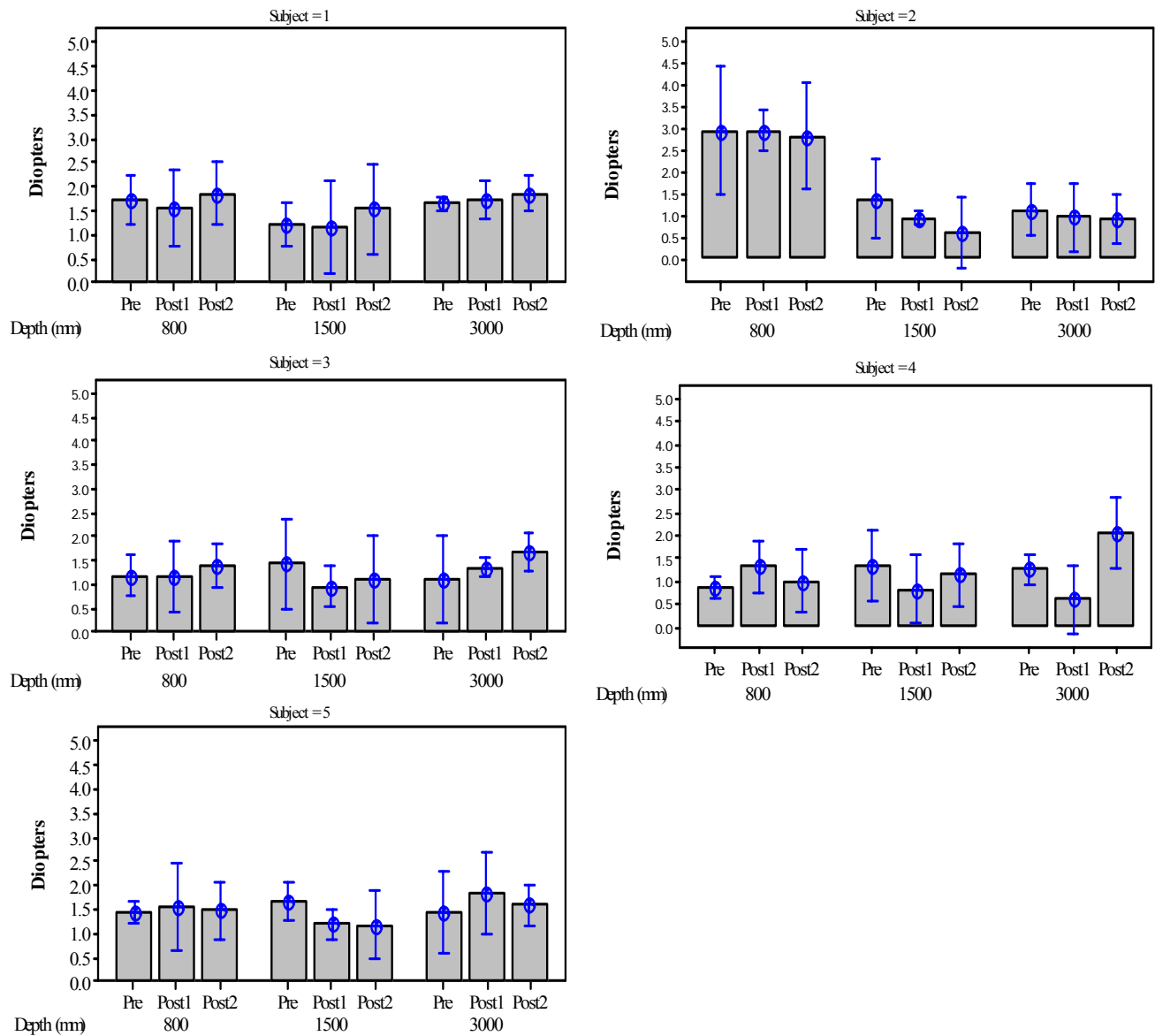
Figure 58 Participants' dark focus means for Pretest and Posttest trials after performing the Virtual-HD task at each viewing depth (800, 1500, or 3000 mm). The 95% Confidence Interval for each dark focus mean is given. The mean averages illustrate that participants seem to change in dark focus from the first and second trials for the 1500 mm viewing depth. For some participants their dark focus does not return to baseline before the third trial.

Table 26 Cybersickness scores for the oculomotor subset per participant (HMPD-1) for viewing distance and testing condition

Depth (mm)		SSQ	
		OD	
		Post1	Post2
	1	0	0
800	2	38	23
	3	0	0
	4	23	30
	5	8	0
1500	1	23	0
	2	61	83
	3	0	0
	4	30	53
	5	8	8
3000	1	8	0
	2	23	28
	3	0	0
	4	15	56
	5	8	0

7.2.4 HMPD-1 Error Metrics and Viewing Distance Conclusion

The results of this experiment are important for two reasons: 1) they demonstrate the difference between assessments for HMDs versus those for user assessment and 2) they open the door to future research when considering usability within the VE design cycle. The metrics for this experiment were chosen to identify individual differences, which they did. However, once identified there was not enough power when performing the analysis to meaningfully interpret these differences. Thus to extend to a real usability paradigm more participants are needed. This point is clearly illustrated with the participants' performance at the 1500 mm viewing distance.

Although all participants showed a downward shift in dark focus, which may be indicative of cybersickness (Fowlkes, Kennedy, Hettinger, & Harm, 1993), we are unable to further evaluate the possibility that the participants were overcompensating in some manner. For example, participant 2, whose sickness scores were the highest, increased his CE from .5 to 7 mm from the H-D pretest to the H-D posttest1. He also flipped sign on the CE metric between testing environments, (i.e., + for H-D and – for V-HD). Participant 4, who also experienced higher sickness scores, also flipped sign on CE in the opposite direction. The overall improvement shown by that testing condition could signal learning effects; however, the current experimental design was not aimed at these finding.

In contrast, the V-HD task was quite capable at evaluating the HMPD-1 against its specifications. Figure 58 shows that for the V-HD overall stereoacuity values are below the 240 arc second limit set by the HMPD-1, ($M_{800} = 186.70$ arc sec, $SD_{800} = 92.10$ arc sec; $M_{1500} = 133.52$ arc sec, $SD = 34.6$; $M_{3000} = 41.91$ arc sec, $SD_{800} = 7.18$). Further, participants performed within 1/4 to 1/10 of the pixel resolution of the display; thus demonstrating the supportive

interplay of the computer graphics and the HMD. However, these results show that observers perceive the V-HD differently than they do the real world task.

Although the HMPD-1 is set to project images at the same location as the object is rendered, participants performed worse at the near distance. This adjustment to the HMPD-1 was done to eliminate the mismatch between accommodation and convergence and to assist the user in performing the task. One explanation for the poor stereoacuity performances at the closer viewing distance could be that although the HMPD was set to function at 800 mm, it is designed for projecting collimated images. We see a vast improvement in stereoacuity as the projection moves toward the infinity setting, which could be set at 6 meters. Further, Wong, Woods, & Peli (2002) showed that stereoacuity is independent of viewing distance when monocular cues are eliminated. It may be that there are depth cue conflicts (e.g., lighting of the virtual objects) because of the HMPD-1 setting.

In Chapter 4, we showed that the H-D task could be a reliable assessment for stereoacuity. Although there was a significant difference in stereoacuity over the viewing distances, this result could suggest a violation of the assumptions of the stereo threshold by the equation. Computations of the exact equation for calculating the difference of the vergence angles were compared to the estimated values; the results were similar. This result suggests that there may be other distance cues represented in both the H-D and V-HD task that may facilitate stereoacuity in some instances and degrade it in others.

Wong, Woods, & Peli (2002) cited research showing that a change in vergence, when accommodation is left unchanged, reduces stereoacuity. These dark vergence changes may persist and carryover into other aspects of the experiment. Since we did not test dark vergence in these studies, we added such a metric to the experiment performed in Chapter 8.

7.3 Virtual Howard-Dolman Assessment of Different HMDs at 1500 mm

The goal of this study was to apply the specifications from an off-the-shelf HMD to the V-HD software to assess the stereoacuity level of the observer while wearing that display. As outlined in Chapter 5 the off-the-shelf HMDs evaluated for this experiment were the Canon Coastar™ (Canon) and the Virtual Research 6 (VR6). The available metrics from the user's manual were applied to the V-HD application to set the parameters of each respective display. The participants were only tested at the 1500 mm viewing distance.

7.3.1 Hypotheses

Given the current VE system setup and the HMD parameters given in Table 7, we can make the following predictions:

Hypotheses for Canon HMD:

- Based on the fixed IPD and a larger pixel size, participants performing the V-HD task wearing the Canon should perform more variably when locating the nominal distance (i.e., CE should be more variable).
- The optical image plane (2000 mm) and the rendered depth of the stimuli (1500 mm) are offset, but are more matched than in the VR6 which may improve precision (VE_r), but not accuracy (CE).
- The estimated stereoacuity based on information from their documentation is 5.8 arc minutes, which equates to 384 arc seconds. From the results of the previous study we expect the graphics to support an overall stereoacuity better than 348 arc seconds.
- Because of the larger pixel size, we expect lower % correct responding on front and back determinations than for HMPD-1.
- Because the optical image location is set beyond the 3D image plane, the mismatch may cause some shift in dark focus.

Hypotheses for the VR6:

- Based on the mismatch between the optical image plane and the rendered depth of the stimuli, (i.e., the optical image plane is set to 911 mm in the VR6), one would expect both CE and VE_r measures to show high variability. However, rendering the virtual objects with respect to the user's own IPD may improve CE variability.
- The estimated stereoacuity for this HMD is 4.8 arc minutes, which equates to 288 arc seconds. From the results of the previous study we expect the graphics to support an overall stereoacuity better than 288 arc seconds.
- Because of the larger pixel size, we expect lower % correct responding on front and back determinations than for HMPD-1.
- Because of the mismatch between the optical image location and the 3D image plane, there may be a shift in dark focus.

Hypothesis for the HMPD-1:

- Because the optical image is set at the same distance as the image plane, we expect low variability in localizing the fixed position of the cylinder. Both measures of CE and VE_r should show lower error.
- Because the smaller pixel size, we expect a higher % correct in responding on front and back determinations than for the other HMDs.

Because the overall time per virtual environment was less than 15 minutes per group, we expected no large affect of cybersickness on the results per HMD condition.

7.3.2 Method of Evaluating Different HMD Types

<u>Participants</u>

This research was approved by the Institutional Review Board (IRB) of the University of Central Florida. The IRB protocol and consent form are listed in Appendix A. Fifteen participants (11 men and 4 women, mean age = 27.2 years) were randomly assigned to an HMD

condition. Each participant was either corrected for or had 20/20 vision. Glasses or contacts were worn during each part of the experiment.

VE System

HMPD-1

The participants performed the Virtual-HD task wearing a first generation see-through projection head mounted display (HMPD-1) with a 52 degree FOV and a predicted 4.1 arc minute resolution. The microdisplays for the HMD were off-the-shelf Liquid Crystal Displays with a VGA resolution of 640x480. Software ran on a computer system with Microsoft XP Pro and a dual processor graphics card. Participants performed the task in the full ARC display. Based on the results reported in Chapter 6, only the Scotchlite[TM] 3M Film Silver (Cubed) panels were used in the experiment.

VR6

The VR6 has a 48 x 36 degree FOV. The microdisplays are Liquid Crystal Displays with a VGA resolution of 640 x 480. The V-HD ran on a computer system with Red Hat Linux 9 as an operating system and a dual Intel Xeon 2.0 GHz processor with a GeForce4 Ti 4600 graphics card. The VR6 has a variable IPD, but the image plane is set at .91 meters. The predicted resolution is 4.81 arc minutes. Participants performed the V-HD task in the ARC although the VR6 is an immersive HMD.

Canon

The Canon Coastar™ hybrid video-see through HMD which corrects for parallax errors, the image displacement between the camera and the eye, by tracking the FOV of the user was run on the same computer as the HMPD-1. The Canon has a 640 x 480 FOV with a set IPD distance of 63 mm and a fixed image plane of 2000 mm. It has a predicted resolution of 5.8 arc minutes.

Experimental Design

Participants were randomly placed in an HMD condition. The HMPD-1 was set to project the displayed image to the retroreflective material 1500 mm from the observer. The Canon was used in immersive mode and the object was rendered 1500 mm in the image space of the HMD. The image for the VR6 was similarly presented to the user. Each participant performed the Howard-Dolman peg test three trials during the experiment, before VE exposure, after the first VE exposure, and after the second VE exposure. Each trial consisted of 10 repetitions. The response variables of CE, |CE|, VE_r and η were calculated from each test trial for the real world viewing condition. Descriptive statistics were computed for each HMD. Since the analysis is not to cross compare the HMDs, no ANOVAs were performed on these metrics. However, a mixed repeated measures ANOVA was performed for dark focus.

Dark focus was measured before VE exposure, after the first VE exposure, and after the second VE exposure. The metric for dark focus is Diopters. This metric was added to evaluate change in visual system responding after each VE exposure.

In addition to calculating the response variables of CE, |CE|, VE_r and η, % correct for evaluating whether the moveable virtual object appeared in front or behind the fixed target upon

176

the start of each repetition was computed. The octahedron appeared equal times to the right or to the left of the fixed target and an equal number of times in front or behind the object. Thus, the front or back judgments constituted a 2AFC design. Participants performed the Virtual Howard-Dolman two times in the experiment. There were a total of 32 responses per trial.

After each exposure in the VE, the participants also filled out the simulator sickness questionnaire (SSQ), also listed in the Appendix. The Total sickness score and subscore for oculomotor effects were compared for each condition. Another questionnaire was administered to monitor general health and eye fatigue. This questionnaire is also listed in the appendix, but the results are not reported since the SSQ provided more information about participant health status.

Procedure

The procedure for this experiment was exactly the same as in the first experiment. The only difference was the adjustment of the displays for the users. The VR6 required more assistance to comfortably fit the users head.

7.3.3 Results of the Virtual-HD Assessment for the Different HMDs

Table 27 shows the CE for each of the participants and their respective HMD cohort. The data show that all participants perform similarly on the H-D task. As expected, the results from the V-HD task are more variable across each HMD condition. The Canon group is consistently biased in front of the point of fixation or nominal distance (Pre_{Canon} = -4.6 mm, SD_{Canon} = 18.4 and $Post_{Canon}$= -6.5 mm, SD_{Canon}= 11.7). The other HMD groups show less variability and more

of a positive bias toward the nominal distance. The patterns of bias are illustrated in Figure 59.

Those participants who wore the Canon HMD show more variability while performing the V-HD

task, while those participants wearing the HMPD-1 showed more variability in their H-D scores.

The VR6 group was more uniform in their responding over both testing conditions.

Table 27 Constant errors (mm) per participant as a function of HMD for the Howard-Dolman
task and Virtual-Dolman task at a viewing distance of 1500 mm

HMD	Subjects	Howard-Dolman						Virtual-Dolman			
		M			SD			M		SD	
		Pre	Post1	Post2	Pre	Post1	Post2	Pre	Post	Pre	Post
HMPD 1	1	1.2	-0.9	2.5	2.8	3.3	1.4	-4.8	-0.5	25.9	28.9
	2	0.5	7.0	3.2	3.5	10.5	2.9	-4.6	-5.5	13.5	18.5
	3	10.7	5.9	6.6	9.1	5.7	1.1	8.5	2.5	23.3	21.9
	4	-3.6	-2.1	0.6	2.6	3.2	1.8	12.5	5.9	17.8	15.6
	5	-0.1	1.0	1.4	2.1	1.7	3.2	8.8	2.1	29.9	29.9
	$\overline{X}$	1.7	2.2	2.9				4.1	0.9		
	SD	5.3	4.1	2.3				8.2	4.2		
Canon	1	0.6	-2.5	-2.3	1.6	3.0	2.7	-5.4	-11.0	23.0	24.1
	2	3.5	3.1	2.1	2.4	2.9	4.0	4.1	-4.3	25.2	24.6
	3	1.9	2.9	3.4	4.4	1.9	2.5	-35.4	-24.2	18.3	23.8
	4	3.5	4.4	2.0	2.5	2.5	1.9	12.4	5.3	22.1	20.4
	5	1.9	2.1	-0.7	3.8	2.3	3.6	1.3	1.6	20.3	21.8
	$\overline{X}$	2.3	2.0	0.9				-4.6	-6.5		
	SD	1.2	2.6	2.3				18.4	11.7		
VR6	1	3.0	0.8	3.2	2.9	3.5	2.2	-9.2	-0.5	25.5	28.0
	2	2.1	0.0	2.2	2.7	2.7	3.4	0.0	7.7	22.8	19.3
	3	3.8	3.7	1.9	1.6	2.7	2.6	5.5	13.9	22.4	11.4
	4	2.5	4.6	1.9	2.5	2.8	2.6	-2.6	-10.3	18.9	23.2
	5	0.4	-1.6	1.1	2.5	2.3	1.7	3.8	3.8	26.8	26.8
	$\overline{X}$	2.4	1.5	2.1				-0.5	2.9		
	SD	1.3	2.6	0.8				5.8	9.1		
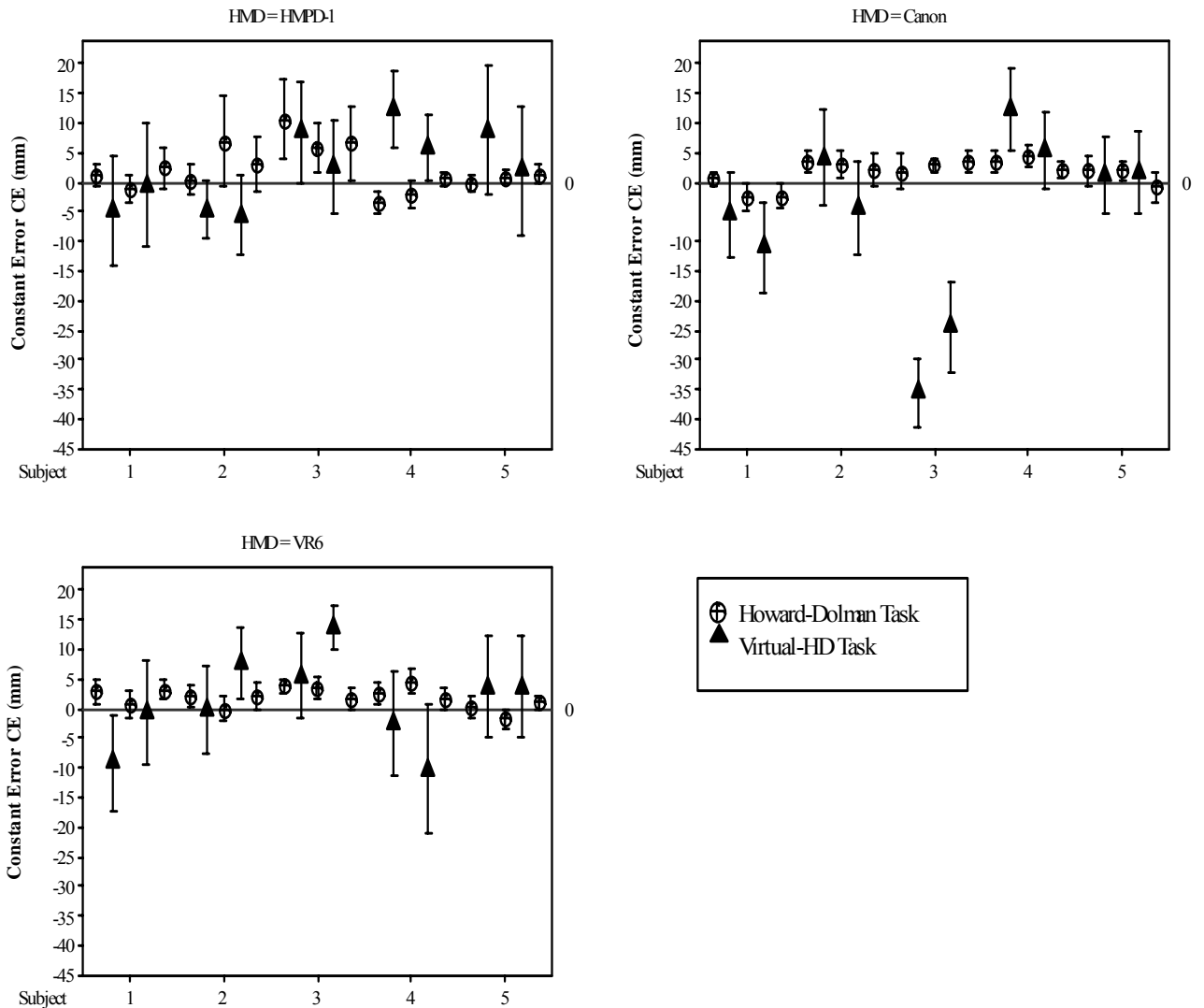
Figure 59 Constant error (CE) for each participant performing at the test viewing distance of 1500 mm for the Howard-Dolman task and the Virtual-HD task. The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). Individual differences in accuracy are illustrated along with their biases for each HMD condition. The center points represent the mean of the signed deviations between the participant's responses and the nominal depth (800, 1500, or 3000 mm). The error bars represent the 95% Confidence Interval of the means.

Table 28 shows the VE_r error means for each participant in their respective HMD cohort. As reflected in the VE_r scores, the HMPD-1 group shows higher precision errors on the H-D task than participants in either of the other HMD groups. The overall H-D variable error means for each HMD cohort are: M_{HMPD-1} = 4.32 mm, SD = 2.75, M_{Canon} = 2.66 mm, SD = 0.78, and M_{VR6} = 2.43 mm, SD = 0.48. The overall precision for each group on the V-HD task for both sessions is equivalent, although the HMPD-1 group shows higher variability over both trials (M_{Pre} = 21.72 mm, SD = 6.40 and M_{Post} =22.61 mm, SD = 6.22. The subpixel based resolution for each of the displays is between ½ and ¼ of the size of a pixel as measured for each display. The pattern of VE_r is shown more clearly in Figure 60.

As the 95% Confidence Intervals (CI) show, the HMPD-1 group consistently performed with higher variability across all trials of the H-D task. This variability is also reflected in their V-HD scores, although the means for the pretest and posttest do not vary. For the Canon group, they show uniform precision over trials on the H-D task. The 95% CIs are more compact suggesting higher precision. This pattern of precision is reflected as well in their V-HD scores. Those persons in the VR6 group performed consistently on the H-D task, yet they were much less precise at aligning the octahedron and the cylinder in the virtual environment. The 95% CI for this group shows a trend of increasing error from the V-HD pretest to the posttest.

Table 28 Variable error VE_r (mm) values per participant and group means as a function of HMD for the Howard-Dolman task and Virtual-Dolman task, viewing distance 1500 mm

HMD	Subjects	Howard-Dolman Task Magnitude			Virtual-Dolman Task Magnitude	
		Pre	Post 1	Post 2	Pre	Post
HMPD-1	1	2.64	3.14	4.57	25.45	28.49
	2	3.35	9.97	6.08	13.25	18.18
	3	8.66	5.45	8.22	22.91	21.52
	4	2.46	3.08	1.43	17.56	15.39
	5	1.97	1.61	2.24	29.43	29.46
	$\overline{X}$	3.82	4.65	4.51	21.72	22.61
	SD	2.75	3.27	2.78	6.40	6.22
Canon	1	1.50	2.87	2.57	22.71	23.84
	2	2.25	2.77	3.75	24.86	24.31
	3	4.16	1.76	2.42	18.07	23.47
	4	2.38	2.37	1.79	21.83	20.12
	5	3.64	2.20	3.41	20.07	21.56
	$\overline{X}$	2.78	2.40	2.79	21.51	22.66
	SD	1.09	0.45	0.79	2.58	1.76
VR6	1	2.75	3.30	2.08	25.13	31.14
	2	2.52	2.57	3.19	22.48	17.99
	3	1.54	2.53	2.43	22.11	26.01
	4	2.33	2.65	2.43	19.92	28.46
	5	2.33	2.15	1.58	29.63	34.49
	$\overline{X}$	2.29	2.64	2.34	23.85	21.41
	SD	0.46	0.41	0.59	3.72	6.63
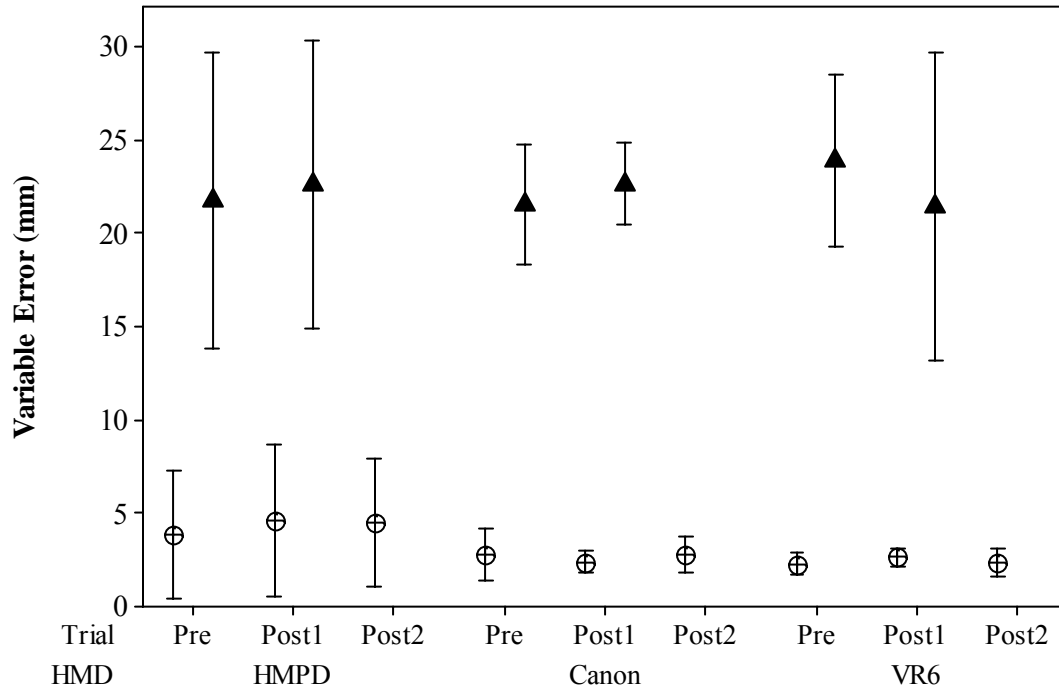
Figure 60 Variable error (VE_r) means for participants performing the Howard-Dolman task or the Virtual-HD task with different HMDs at a viewing distance of 1500 mm. The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). The 95% Confidence Interval is also represented. The participants performed similarly on both tasks regardless of the HMD.

Stereoaccuracy allows better understanding of the vergence angles of the observer attained when assessing the distance, and subsequently aligning the moveable object. The sign of the stereoacuity value gives an indication of the convergence angle related to the fixed object. As Equation 20 illustrates, when the convergence angle relating to the fixated object is bigger than that of the moveable object, stereoaccuracy will take on a negative value. Thus, the Canon group overall judged the cylinder to be closer than its fixed position, $M_{CanonPre}$ = -28.9 mm, SD = 112.0 and $M_{CanonPost}$ = -33.0 mm, SD = 74.4. Similarly those participants in the HMPD-1 group perceived the cylinder as being further away from the nominal position ($M_{HMPD-1Pre}$ = 23.6 mm, SD = 49.2), and then nearer to the nominal depth on the subsequent trial ($M_{HMPD-1Post}$ = 4.0, SD = 26.6). The VR6 group at first aligned the objects quite well (M_{VR6Pre} = 2.2, SD = 35.0), and then on the second trial judged the cylinder to be closer than the nominal depth ($M_{VR6Post}$ = -33.0, SD = 74.4). The pattern of results is shown in Figure 61.

Table 29 Stereoaccuracy (η_{ce}) in arc seconds per participant as a function of HMD for the Howard-Dolman task and Virtual-Dolman task

| | | Howard-Dolman | | | | | | Virtual-Dolman | | | |
| | | M | | | SD | | | M | | SD | |
HMD	Subjects	Pre	Post1	Post2	Pre	Post1	Post2	Pre	Post	Pre	Post
HMPD	1	5.4	-4.1	11.3	12.6	15.0	21.8	-28.7	-6.3	154.1	172.5
	2	2.5	34.7	15.8	17.5	52.0	31.7	-29.8	-35.8	87.6	120.2
	3	47.0	25.9	29.0	40.1	25.2	38.1	48.9	14.5	134.4	126.2
	4	-16.3	-9.5	2.7	11.7	14.7	6.8	74.6	35.0	106.3	93.2
	5	-0.5	4.6	6.4	9.6	7.8	10.9	53.0	12.4	180.9	181.1
	$\overline{X}$	7.6	10.3	13.1				23.6	4.0		
	SD	23.6	19.2	10.2				49.2	26.6		
Canon	1	2.7	-11.3	-10.4	7.1	13.7	12.3	-32.4	-65.5	137.0	143.9
	2	15.4	13.6	9.2	10.4	12.8	17.4	23.6	8.0	145.4	121.1
	3	8.8	13.5	16.0	20.5	8.7	11.9	-217.3	-148.8	112.4	146.0
	4	16.0	19.8	9.0	11.4	11.3	8.6	73.8	31.7	131.7	121.4
	5	8.8	9.5	-3.2	17.4	10.5	16.3	8.0	9.8	121.1	130.1
	$\overline{X}$	10.3	9.0	4.1				-28.9	-33.0		
	SD	5.5	12.0	10.7				112.0	74.4		
VR6	1	14.2	3.7	1.6	13.7	16.5	10.4	-57.1	-65.5	158.7	174.8
	2	9.3	0.0	1.1	11.8	12.1	15.0	-0.2	8.0	133.5	113.1
	3	15.9	15.5	11.4	6.8	11.2	10.7	30.3	-148.8	123.2	62.6
	4	11.5	21.2	-1.7	11.3	12.9	11.8	16.4	31.7	125.7	120.1
	5	1.8	-7.3	-11.4	11.1	10.3	7.5	21.7	9.8	155.0	155.0
	$\overline{X}$	10.5	6.6	0.2				2.2	-33.0		
	SD	5.5	11.6	8.2				35.0	74.4		
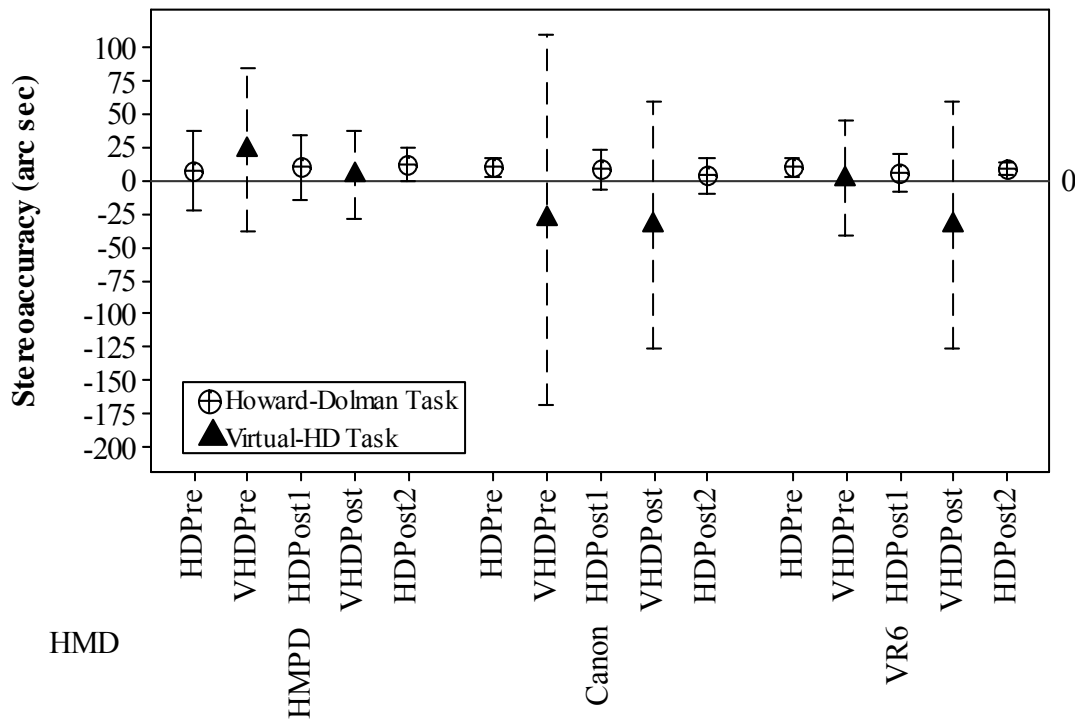
Figure 61 Stereoaccuracy η_{ce} for each HMD condition at the viewing distance of 1500 mm for the Howard-Dolman task and the Virtual-HD task. The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). Combined accuracy for each HMD condition illustrated along with their biases. The center points represent the mean of stereoaccuracy as calculated from the CE, viewing depth 1500 mm, and interpupillary eye distance (IPD) of the observer. The error bars represent the 95% Confidence Interval of the means.

As expected the participants performed at the stereoacuity limit for real world viewing, 20 arc seconds. Participants in all groups were not different from each other in stereoacuity as measured by the H-D task. Importantly, there are no changes from baseline over repeated trials within any HMD group.

The predicted resolution in arc seconds for each display is: 246, HMPD-1; 288, VR6; and 348, Canon. Table 30 lists the stereoacuity values and overall stereoacuity means and standard deviations for each participant within an HMD type. All HMDs performed better than the limit of the microdisplay used in the HMD. This result is shown graphically in Figure 62.

Table 30 Stereoacuity (η_{ve}) in arc seconds per participant as a function of HMD for the Howard-Dolman task and Virtual-Dolman task

HMD	Subjects	Howard-Dolman Task $\underline{M}$			Virtual-Dolman Task $\underline{M}$	
HMPD-1	1	11.96	14.25	20.69	151.64	169.77
	2	16.60	49.35	30.09	86.26	118.31
	3	38.04	23.93	36.13	132.32	124.26
	4	11.14	13.96	6.47	104.64	91.71
	5	9.08	7.42	10.33	178.06	178.28
	$\overline{X}$	17.36	21.78	20.74	130.58	136.46
	SD	11.88	16.50	12.61	36.53	36.53
Canon	1	6.78	13.02	11.65	135.32	142.08
	2	9.87	12.18	16.49	143.57	140.38
	3	19.42	8.21	11.30	110.99	144.16
	4	10.78	10.76	8.11	130.08	119.87
	5	16.48	9.99	15.44	119.60	128.46
	$\overline{X}$	12.67	10.83	12.60	127.91	134.99
	SD	5.15	1.88	3.39	12.85	10.42
VR6	1	13.03	15.63	9.86	156.67	172.63
	2	11.23	11.46	14.22	131.87	111.68
	3	6.43	10.59	10.15	121.60	61.78
	4	10.74	12.21	11.17	120.55	136.53
	5	10.24	9.46	6.93	171.14	153.07
	$\overline{X}$	10.33	11.87	10.47	140.37	127.14
	SD	2.42	2.34	2.63	22.53	42.84
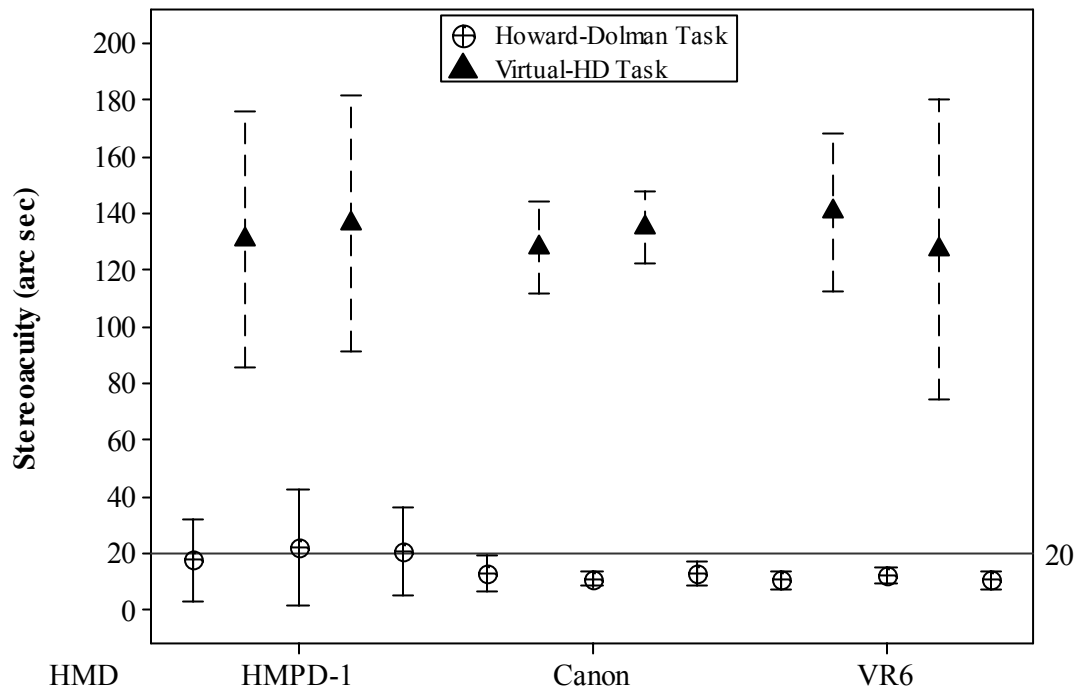
Figure 62 Stereoacuity η_{ve} means and 95% Confidence Interval for the participants' HMD grouped responses on the Howard-Dolman and Virtual-HD tasks performed at the different depths (800, 1500, 3000 mm) and on the separate Pre and Posttest trials. The mean averages illustrate that stereoacuity is similar across HMDs. The reference line at 20 arc seconds represents average responding in the population at large. The predicted stereoacuity for the HMPD is 240 arc seconds.

Table 31 displays the percent correct in identifying whether the initial starting position for the octahedron was in the front or the back of the fixed cylinder. As discussed in Chapter 4, stereopsis is based on relative disparities. These disparities allow the user to discriminate objects which are closer to the observer from those that are further away from a fixed point. If the participants were truly using disparity as a cue for distance perception, then they should be able to discriminate the front and back initial positions of the octahedron relative to the fixed cylinder at above chance levels. However, this discrimination is also limited by the resolvable depth range of each respective HMD. Because this discrimination assessment is a 2AFC design, the threshold is set at 75 percent.

Table 31 Percent correct for front and back responses as a function of HMD for the Virtual-Dolman task

HMD		Virtual-Dolman Task	
		$\underline{M}$	
		Trial 1	Trial 2
HMPD-1	$\overline{X}$	0.75	0.78
	SD	0.18	0.11
Canon	$\overline{X}$	0.62	0.65
	SD	0.20	0.12
VR6	$\overline{X}$	0.58	0.50
	SD	0.15	0.26

The data from Table 31 are graphically displayed in Figure 63 Participants are at threshold when discerning front and back locations of the octahedron while wearing the HMPD-1. In each of the other two HMD groups, participants consistently responded below the threshold.
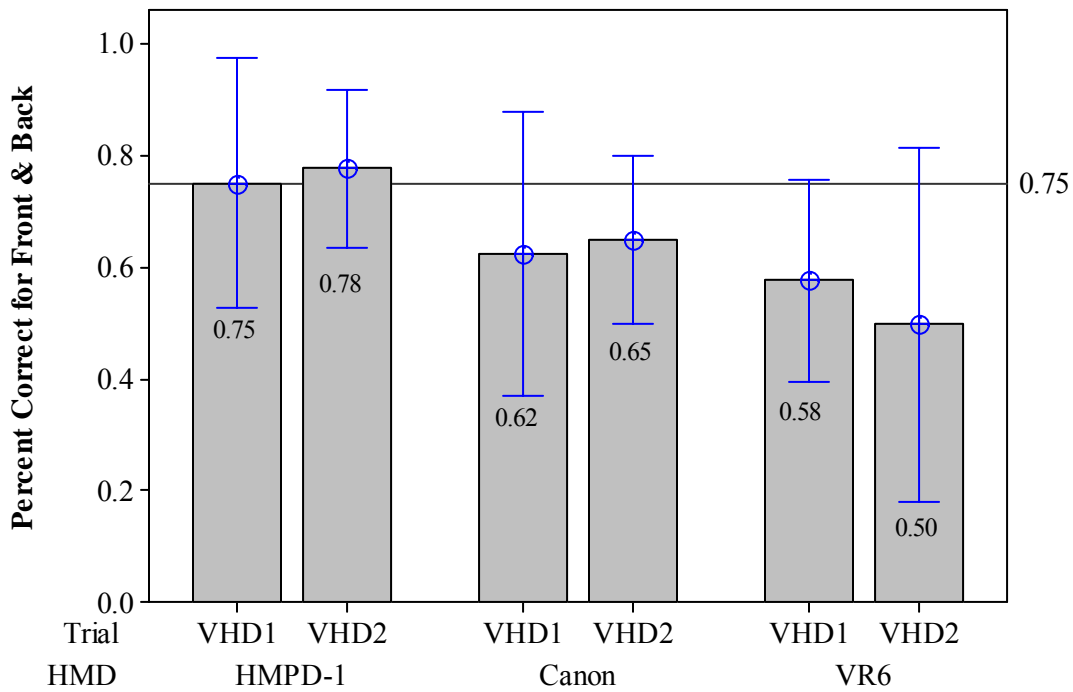
Figure 63 Percent correct for front and back judgments on both trials of the V-HD task grouped over HMD type. The mean and 95% CI are given. The reference line at 75% represents the detection threshold for a 2AFC design.

Table 32 shows the dark focus measures for each individual with their respective HMD cohort. Within each HMD group, the dark focus values are consistent. A One-Way Analysis of Variance showed that there was a difference in dark focus over the HMD types on each trial, Dark Focus$_{pre}$, $F(2, 12) = 13.39$, $p = .001$, $\eta_e = .83$, Dark Focus$_{Post1}$= 10.87, $p = .001$, $\eta_e = .80$, Dark Focus$_{Post2}$ = 20.84, $p = .001$, $\eta_e^6 = .88$. A Tukey HSD post hoc analysis confirmed that the dark focus of participants in the HMPD-1 group were significantly different than those of the other two HMD conditions. The dark focus measured for participants in the Canon and VR6 groups were not different from each other nor were they different over trial. As reported in experiment 1, the participants in the HMD-1 experience a significant downward shift in dark focus from their baseline score on Trial 2. These results are graphically displayed in Figure 64.

---

[6] η_e is used to distinguish the effect size Eta from the Stereoacuity value.

Table 32 Mean dark focus in Diopters per participant as a function of HMD and testing condition.

HMD		Dark Focus Diopters M		
		Pre	Post1	Post2
HMPD-1	1	1.2	1.1	1.5
	2	1.4	0.9	0.6
	3	1.4	1.0	1.1
	4	1.4	0.8	1.2
	5	1.7	1.2	1.2
	$\overline{\overline{X}}$	1.4	1.0	1.1
	SD	0.2	0.1	0.3
Canon	1	2.2	2.0	2.2
	2	2.5	3.2	2.9
	3	2.4	2.3	2.5
	4	2.2	1.7	2.4
	5	2.3	2.3	2.5
	$\overline{X}$	2.3	2.3	2.5
	SD	0.1	0.6	0.3
VR6	1	2.7	2.2	1.7
	2	2.2	3.0	2.8
	3	1.5	1.4	1.9
	4	2.4	2.3	2.2
	5	2.2	2.2	2.1
	$\overline{X}$	2.2	2.2	2.1
	SD	0.5	0.6	0.4
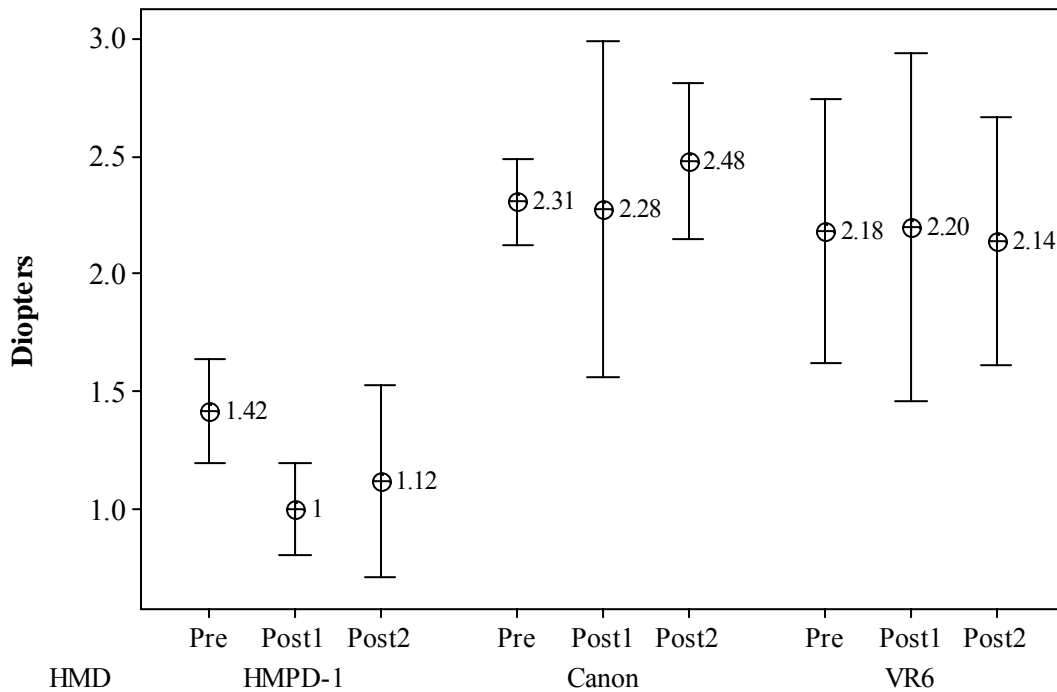
Figure 64 Overall dark focus means for Pretest and Posttest trials. The 95% Confidence Interval for each dark focus mean is given. The average dark focus in Diopters measured after participants performed the Virtual-HD task while wearing the Canon or the VR6 is significantly different from the dark focus of those who performed the task wearing the HMPD-1. They are no different from each other and show no shift from baseline.

SSQ scores are presented in Table 33. Participants from each HMD type presented with Oculomotor subscore on the SSQ. A score of 20 or greater signaled that the VE exposure was having an effect on the user. The general trend in the data showed that participants from each HMD condition do experience elevated scores on the SSQ Oculomotor subtest after the first virtual environment exposure. These symptoms either became worse or lessened through the course of the experiment. More data is needed to make further conclusions.

Table 33 Cybersickness scores for the oculomotor subset per participant for HMD type and testing condition

		SSQ	
HMD		OD	
		Post1	Post2
HMPD-1	1	23	0
	2	61	23
	3	0	0
	4	30	30
	5	8	0
Canon	1	30	45
	2	23	23
	3	98	106
	4	0	0
	5	0	0
VR6	1	30	8
	2	68	83
	3	38	53
	4	15	15
	5	0	0

7.3.4 Conclusion of HMD Comparison

The results of this experiment showed that the V-HD task in conjunction with the H-D task provide a method for assessing HMD of differing types (e.g., immersive versus see-though) and parameters (e.g., fixed focal plane versus variable) with respect to their documented specifications. For example, based upon the fixed IPD and larger pixel size of the Canon, we expected that the CE measured would have higher variability caused by inaccurate positioning of the octahedron with respect to the cylinder's fixed distance over successive repetitions. This pattern of results was shown in the analysis.

The resolvable depth limits for each HMD predicted the direction of percent correct responding for front and back judgments on whether the octahedron was positioned in front of or behind the cylinder at the start of each repetition. More specifically, the subpixel resolution step was finer for the HMPD-1; consequently, participants were able to judge front and back starting locations of the octahedron with respect to the cylinder's fixed distance at threshold or 75%. We expect that with a higher resolution display, such as the HMPD-2, the percent correct outcomes should be well above the threshold. However, participants tested while wearing either of the other types of HMDs performed below threshold for this metric possibly suggesting that other cues to depth (e.g., stimulus size) were being used to perform the task. The resolvable depth also predicted ranges based on pixel and sub pixel resolution that an HMD could achieve given optimal viewing conditions. For each HMD, participants achieved subpixel perfromance between ½ and ¼ the size of a pixel as measured by the microdisplay properties of each HMD.

The stereoacuity metric showed that all HMDs performed better than the resolution limit imposed by their microdisplay. The average stereoacuity across all HMD types was similar despite the increased resolution of the HMPD-1 and the alignment of the optical image plane

with the 3D image plane. The Canon also performed better than expected given the information in the owner's manual. These results further support benchmarking for HMDs with a user-in-the-loop methodology. However, because the HMPD-1 was built in-house, the potential sources of performance degradations are more easily pinpointed. For example, the results of experiment 1 of this chapter clearly show that the HMD fails to perform optimally at distances away from its designed focal plane of optical infinity. We may be able to attribute the improvement in the Canon to the optical design, which aligns the optical axis with the display axis to reduce parallax or the mismatch between the cameras FOV and the FOV of the eye. This alignment may allow for a more accurate depth image to be presented to the user (Biocca & Rolland, 1998; Keller & Colluci, 1998).

Localization errors are generally thought to occur due to a bias toward dark focus when the stimuli in the environment are inadequate (Owen, 1984). However, there was no change in dark focus for participants who performed the V-HD task wearing either the Canon or the VR6. On average the participants spent less than 15 minutes per V-HD trial. This duration may not be long enough or the stimuli are too simplistic to cause a shift in dark focus. The suboptimal focal plane settings may have introduced more blurring of the image across the 800 and 1500 mm depths for the user wearing the HMPD-1. However, we should have observed greater dark focus shifts at 800 mm in comparison to the shift found at 1500 mm, which does not occur.

Although all HMD types show some form of cybersickness across participants, the small sample size of the experiment precludes any further analysis. However, these findings do indicate next steps for future work, such as probing the mismatch between the focal plane and the image plane with real and virtual objects. With a larger sample size, the effects of cybersickness on performance of the V-HD task and the shift in dark focus as a measure of

196

cybersickness or eye strain may prove important for extending the current capabilities of the test to encompass usability.

CHAPTER 8 HMPD-2: STEREOACUITY, DARK FOCUS, AND DARK CONVERGENCE FOR HMPD-2

8.1 Introduction

The methodology developed in Chapter 7 was extended to include a measure of dark convergence, distance where a person converges when there are no stimuli in the viewing area. Owens & Leibowitz (1976) showed that persons with a near dark convergence underestimate depth more than those persons with a far dark focus. Further, they showed that under scotopic lighting conditions, vergence and accommodation are decoupled and shifted toward different resting points. Thus, dark vergence, and not dark accommodation should change under the degraded viewing conditions of the virtual environment. Dark vergence would be a better metric to evaluate the HMPD-2, where the optics are designed an optimized to change focal planes.

We are also replicating the work of Chapter 7 to show that the V-HD task is reliable enough to estimate the stereoacuity of a higher resolution microdisplay (800 x 600) with a resolution of 2.6 arc minutes or 156 arc seconds. The resolvable depth for the HMPD-2 is also finer than that of the HMPD-1. Finally, while the microdisplays assessed in Chapter 7 were all LCDs, this microdisplay is an OLED with higher brightness and better contrast.

8.2 Hypotheses

Given the current VE system setup, we can make the following predictions:

Hypothesis 1: We expect the precision errors (VE_r) to fall within the resolvable depth ranges for the HMPD-2 at each respective viewing distance. Stereoacuity is limited by the resolution of the display and should be roughly 156 arc seconds or 2.6 arc minutes.

Hypothesis 2: Given that the HMDP-2 is optimized for viewing at each depth and we are aligning the optical image with the rendered image on the depth plane, we expect no shifts of dark focus or dark convergence across trials. Additionally, we do not expect adaptation changes as reflected in performance on the Howard-Dolman depth test.

Hypothesis 3: Based on the resolvable depth, the percent correct on identifying the front or back initial positions of the octahedron with respect to the cylinder on each repetition should be well above threshold.

Hypothesis 4: Given the exposure time in the virtual environment is less than 15 minutes per testing condition, we hypothesize that participants will experience low cybersickness scores on the SSQ.

8.3 Method for Evaluating the HMPD-2 over Three Viewing Distances

Participants

This research was approved by the Institutional Review Board (IRB) of the University of Central Florida. The IRB protocol and consent forms are listed in Appendix A. Five participants (all male, mean age = 30.6, SD = 5.36) performed the experiment using the HMPD-2. Each participant in this experiment was either corrected for or had 20/20 vision. Glasses or contacts were worn during each part of the experiment.

VE System

We used the HMPD-2, which currently has a 40 degree FOV. The microdisplays for the HMD were off-the-shelf OLEDs manufactured by eMagin with a resolution of 800 x 600. Software ran on a computer system with Microsoft XP Pro and a dual processor graphics card. We used the quarter ARC display. Based on the results reported in Chapter 6, only the Scotchlite[TM] 3M Film Silver (Cubed) panels were used in the experiment.

Experimental Design

The experimental design was similar to that used with the HMPD-1. Participants were randomly placed in a viewing distance sequence (e.g., 800, 1500, and 3000). Only one participant's viewing distance order overlapped with another participant. This order was 1500, 800, 3000. Each viewing distance was presented on a different day to minimize order effects.

Each participant performed the Howard-Dolman peg test three times during the experiment, before VE exposure, after the first VE exposure, and after the second VE exposure. Each trial consisted of 10 repetitions. The response variables of CE, |CE|, VE_r and η were calculated from each test trial for the real world and virtual viewing condition. A repeated measures ANOVA was performed using only the |CE| and VE_r values. A repeated measures ANOVA was performed using only the |CE| and VE_r values.

Dark focus and dark convergence were also measured before VE exposure, after the first VE exposure, and after the second VE exposure. The metric for dark focus is Diopters and that of convergence is the meter-angle. Each was evaluated for changes after each VE exposure.

In addition to calculating the response variables of CE, |CE|, VE_r, and η, the % correct for evaluating whether the octahedron was in front of or behind the cylinder at the start of each stimulus presentation was computed. The octahedron appeared equal times to the right or to the left of the fixed target and an equal number of times in front or behind the object. Thus, the front or back judgments constituted a 2AFC design. Participants performed the Virtual Howard-Dolman two times in the experiment. There were a total of 32 responses per trial.

After each exposure in the VE, the participants also filled out the SSQ, also listed in Appendix B. The total sickness score and subscore for oculomotor effects were compared for each condition. Another questionnaire was administered to monitor general health and eye fatigue. This questionnaire is also listed in Appendix B, but the results are not reported since the SSQ provided more information about participant health status.

<u>Procedure</u>

The procedure for this experiment is outlined in Chapter 7.2.2. There are three changes to the procedure. The code was re-evaluated for accommodating the new resolution and resolvable depth for the HMPD-2. Thus, we changed the gross and fine adjustments to reflect the finer resolution of the HMPD-2. Table 34 gives these new values.

Table 34 Gross and fine adjustments for the input device

Viewing Distance (mm)	Gross Adjustment (mm)	Fine Adjustment (mm)
800	15	1.5
1500	50	5
300	200	20

We also changed the stimulus presentation depth interval to reflect the distance change for each viewing distances. This was done to ensure that the participant could perceive the movement of the octahedron when turning the fine and gross adjustors. For the 800 mm viewing distance, the octahedron could appear anywhere within a ± 50 mm range, where the center of this range was considered the nominal distance. The range was increased to a ± 75 mm range for the 1500 mm viewing distance and to a ± 300 mm range for the viewing distance of 3000 mm. These values were chosen based on the linear magnification, or the ratio of the object size to the image size.

The dark vergence was measured using a horizontal grating and a red flashing light affixed to a support at the observer's eye height. Figure 65 gives an example of the setup. With the horizontal grating covering his preferred eye, the participant saw the flashing as a vertical line. The red light appeared as a red dot in the uncovered eye. Participants adjusted their distance

until the dot and the horizontal line were aligned one on top of each other. The distance from the wall to the participant's eye was measured in millimeters. This distance was converted to meter-angles by computing the reciprocal of the measured distance (i.e., 1/distance in meters).
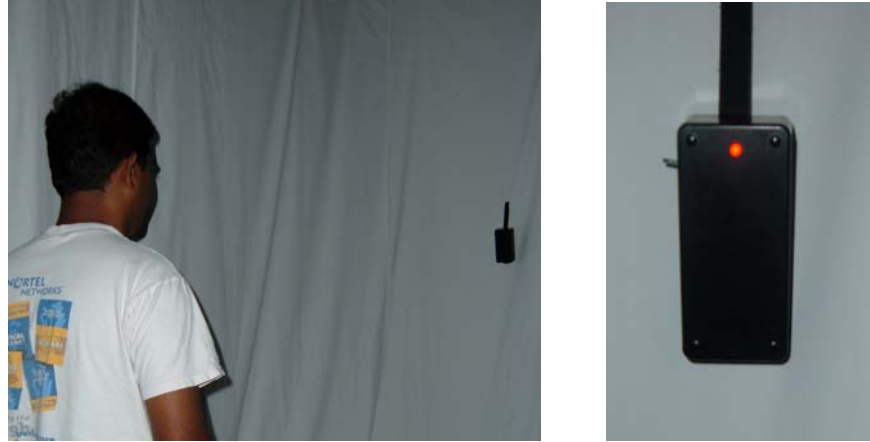
Figure 65 Participant performing the dark convergence test

8.4 Results of HMPD-2 over the Three Viewing Distances

Table 35 gives the CE means and the standard deviations for each participant at each viewing distance and for both testing environments. For the H-D task, the overall bias was minimal at the 800 and 1500 mm viewing depths, $M_{H-D800} = 0.24$ mm, SD = 3.55 and $M_{H-D1500} = -0.89$, SD = 4.00. The overall bias for the 3000 mm viewing distance was negative and more variable, $M_{H-D1500} = -7.36$ mm, SD = 6.88. For the V-HD task, the variability in responding increased over distance for each pretest and posttest outcomes.

Table 35 Constant errors (mm) per participant (HMPD-2) as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman M Pre	Post1	Post2	Howard-Dolman SD Pre	Post1	Post2	Virtual-Dolman M Pre	Post	Virtual-Dolman SD Pre	Post
		Pre	**Post1**	**Post2**	**Pre**	**Post1**	**Post2**	**Pre**	**Post**	**Pre**	**Post**
800	1	2.6	3.6	1.2	1.6	1.8	1.7	0.1	0.0	5.2	5.5
	2	-6.9	-6.3	-4.9	4.0	2.4	2.7	3.6	3.5	3.7	3.1
	3	1.4	2.1	0.8	1.4	1.5	1.3	-3.0	-3.0	3.5	1.3
	4	-0.5	0.4	0.5	-0.5	1.6	1.3	3.1	3.9	4.4	4.6
	5	1.4	3.1	5.1	1.4	2.5	3.7	6.1	7.7	3.3	6.1
	$\overline{X}$	-0.4	0.6	0.5				2.0	2.4		
	SD	3.8	4.0	3.6				3.5	4.1		
1500	1	-0.1	2.3	1.0	3.1	2.2	1.4	-8.9	-7.5	27.3	29.0
	2	-3.3	-12.9	-4.7	3.3	10.5	2.9	12.1	4.2	20.3	17.2
	3	2.3	3.3	1.7	1.3	0.9	1.1	-3.4	-1.6	14.2	12.5
	4	-2.5	0.0	-0.7	1.8	2.4	1.8	1.2	-1.0	12.7	11.3
	5	0.3	- 1.6	1.5	2.3	4.2	3.2	15.6	1.0	18.8	16.7
	$\overline{X}$	-0.7	-1.8	-0.2				3.3	-1.0		
	SD	2.3	6.5	2.7				10.3	4.3		
3000	1	-8.7	-4.6	-1.1	8.5	7.1	5.6	-34.7	-17.5	69.3	67.9
	2	-19.0	-17.2	-14.3	8.1	6.4	2.9	23.1	24.8	44.8	40.5
	3	2.9	0.8	-0.5	3.2	3.4	2.9	-17.8	5.7	39.3	37.9
	4	-9.4	-16.7	-7.8	8.5	6.6	4.4	10.0	32.2	18.4	39.6
	5	-5.3	-6.4	-3.1	8.9	4.9	2.0	-10.3	-11.6	54.4	39.8
	$\overline{X}$	-7.9	-8.8	-5.4				-5.9	6.7		
	SD	7.9	7.9	5.8				22.8	21.8		

The Log transformed |CE| vales are given in Table 36. A repeated measures ANOVA (3 Viewing Depths x 3 Trials) performed on the log transformed H-D |CE| showed no significant main effect of Viewing Distance or Trial. This result is shown graphically in Figure 66. The repeated measures ANOVA (3 Viewing Distances x 2 Trials) for the V-HD task showed a main effect of depth, $F(2, 8) = 7.62$, $p = .014$, $\rho_I = .69$. Post Hoc analysis using the Bonferroni post hoc criterion for significance confirmed that the participants are least accurate performing at the 3000 mm viewing distance, ($M_{3000Pre} = 18.19$ mm, 95% CI = 9.54 to 33.11 and $M_{3000Post} = 16.59$ mm, 95% CI = 7.58 to 36.31). There was no effect of Trial.

Table 36 Log transformed absolute constant errors |CE| for each participant and group means as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman Log			Virtual-Dolman Log	
		Pre	Post1	Post2	Pre	Post
800	1	0.56	0.66	0.34	0.05	0.01
	2	0.90	0.86	0.77	0.66	0.65
	3	0.38	0.49	0.26	0.60	0.60
	4	0.18	0.15	0.18	0.61	0.69
	5	0.38	0.61	0.79	0.85	0.94
	$\overline{X}$	0.48	0.56	0.47	0.56	0.58
	SD	0.27	0.27	0.29	0.30	0.34
Geometric	$\overline{X}$	3.01	3.63	2.95	3.63	3.80
1500	1	0.04	0.52	0.30	0.99	0.93
	2	0.63	1.14	0.76	1.12	0.71
	3	0.52	0.63	0.43	0.64	0.42
	4	0.54	0.00	0.23	0.35	0.29
	5	0.11	0.41	0.40	1.22	0.29
	$\overline{X}$	0.37	0.54	0.42	0.86	0.53
	SD	0.27	0.41	0.20	0.36	0.28
Geometric	$\overline{X}$	2.34	3.46	2.63	7.24	3.38
3000	1	0.99	0.75	0.32	1.55	1.27
	2	1.30	1.26	1.18	1.38	1.41
	3	0.59	0.26	0.18	1.27	0.82
	4	1.02	1.25	0.94	1.04	1.52
	5	0.80	0.87	0.61	1.05	1.10
	$\overline{X}$	0.94	0.87	0.65	1.26	1.22
	SD	0.26	0.88	0.42	0.22	0.27
Geometric	$\overline{X}$	8.70	7.41	4.46	18.19	16.59
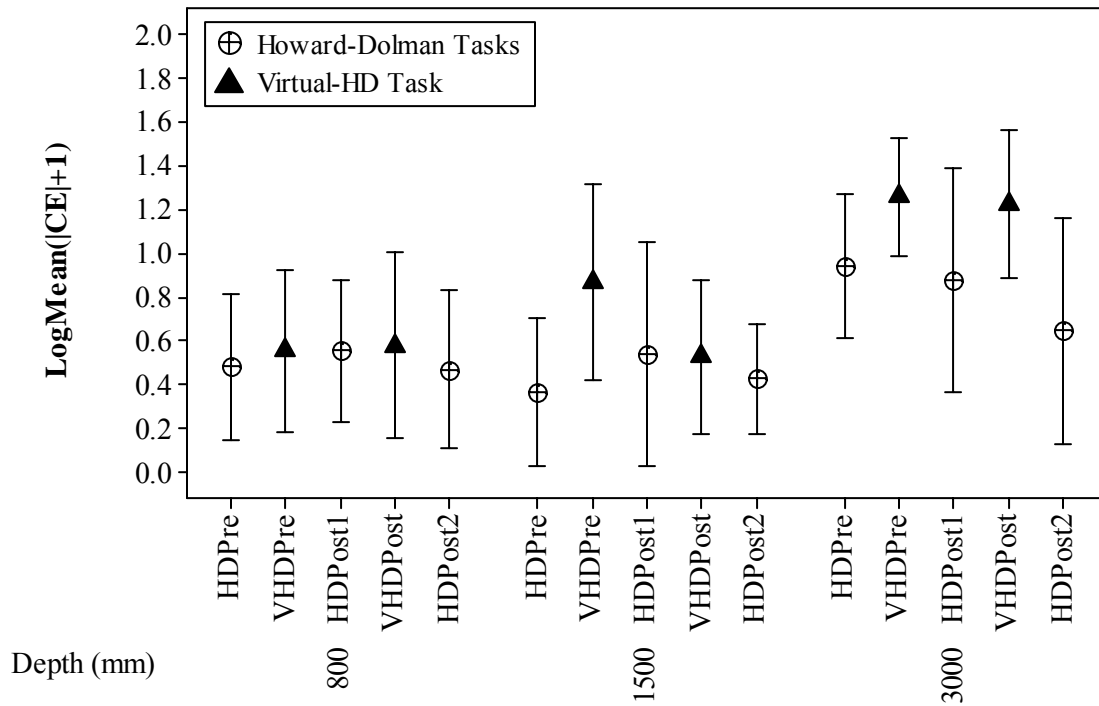
Figure 66 Log transformed means for the absolute constant error |CE| for testing condition, Howard-Dolman or Virtual-HD, at each viewing distance (800, 1500, and 3000 mm). The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). The center points represent the log mean of the absolute CE for the grouped means. The error bars represent the 95% Confidence Interval of the log means.

The |CE| results were pooled over Trial, but not Viewing Distance, to determine if the participants were more accurate in performing the H-D task than they were performing the V-HD task. A repeated measures ANOVA (3 Viewing Distances x 2 Environments) showed that there was a main effect of Viewing Distance, $F(2, 8) = 11.24$, p. $= .005$, $\rho_I = .77$ and Environment, $F(1, 4) = 16.26$, $p = .016$, $\rho_I = .88$. There was no significant interaction. These results are presented graphically in Figure 67. The overall geometric means for each condition are: $M_{800} = 3.68$ mm, 95%CI $= 1.83$ to 5.67, $M_{1500} = 4.96$ mm, 95% CI $= 2.99$ to 8.22 and $M_{3000} = 17.46$ mm, 95% CI $= 12.45$ to 24.35.
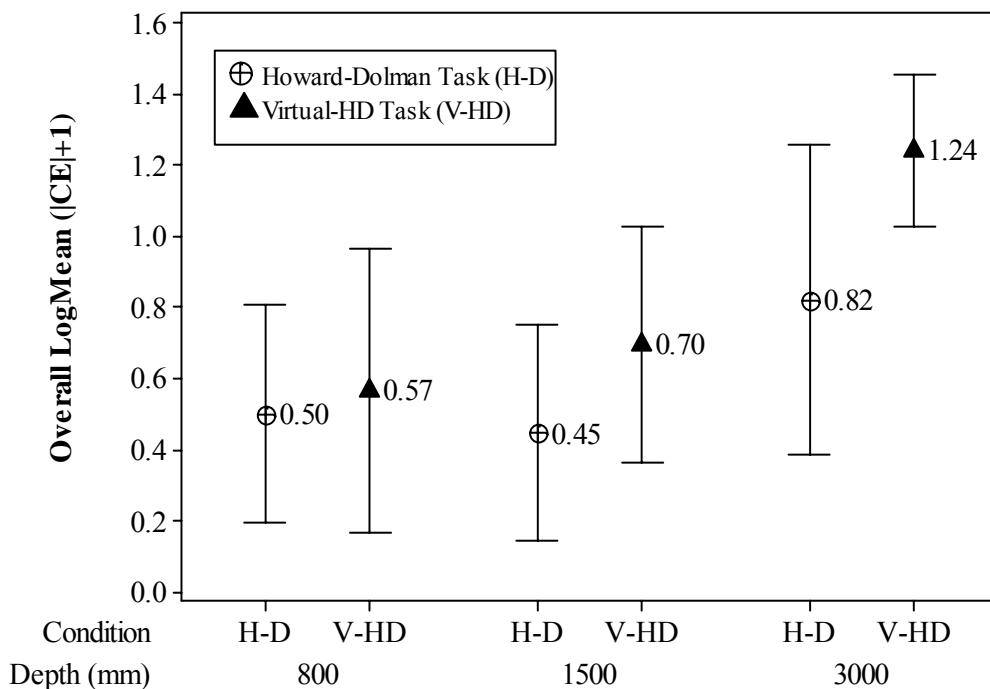
Figure 67 Combined log means of the absolute constant error |CE| for the different viewing distances (800, 1500, 3000 mm) representing the Howard-Dolman task and the Virtual-HD task. The center points signify the log mean of the absolute |CE| pooled over all subjects. The error bars represent the 95% Confidence Interval of the log means.

Table 37 shows the untransformed VE_r for each of the participants over viewing distances and testing environments. The results are presented graphically in Figure 68. For both the H-D task and the V-HD task, precision is decreased when viewing distance is increased. However, VE_r also corresponds to the resolvable depth computed for pixel resolution as shown in Table 38. Table 38 shows all values of the VE_r are precise to within 1/4 of a subpixel resolution.

Table 37 Untransformed variable error VE$_r$ (mm) values per participant and group means as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman Task			Virtual-Dolman Task	
		Magnitude			Magnitude	
		Pre	Post 1	Post 2	Pre	Post
800	1	1.56	1.69	1.60	5.12	5.41
	2	3.78	2.24	2.59	3.64	3.08
	3	1.36	1.45	1.25	3.44	1.25
	4	1.12	1.50	1.20	4.32	4.54
	5	0.92	2.39	3.53	3.21	6.05
	$\overline{X}$	1.75	1.85	2.03	3.95	4.06
	SD	1.16	0.43	1.01	0.78	1.93
1500	1	2.91	2.10	1.34	26.84	28.53
	2	3.13	9.97	2.72	19.96	16.90
	3	1.19	0.90	1.00	14.01	12.34
	4	1.69	2.32	1.68	12.47	11.07
	5	2.15	3.98	3.04	18.54	16.47
	$\overline{X}$	2.21	3.86	1.96	18.37	17.06
	SD	0.82	3.59	0.88	5.66	6.89
3000	1	8.10	6.74	5.28	68.20	82.93
	2	7.64	1.99	2.72	44.11	39.07
	3	3.08	3.19	2.73	38.71	37.34
	4	8.05	6.29	4.17	18.09	38.99
	5	8.41	4.69	1.87	53.53	39.21
	$\overline{X}$	7.06	4.58	3.35	44.53	47.51
	SD	2.24	2.02	1.36	18.54	19.82
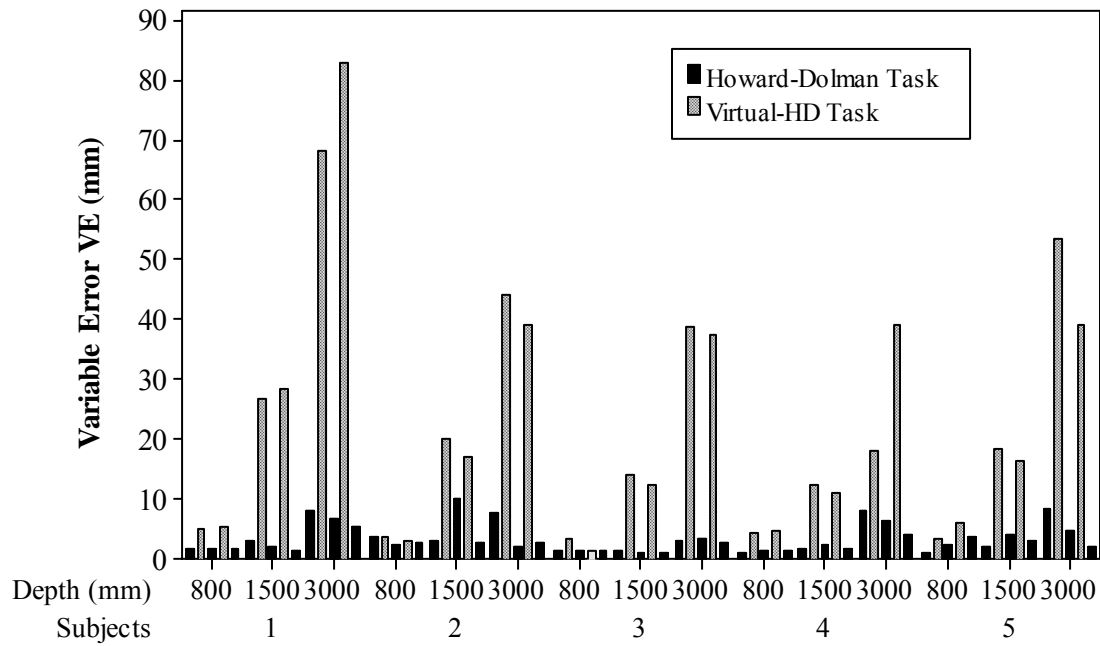
Figure 68 Variable error (VE_r) values for each participant performing the Howard-Dolman task or the Virtual-HD task at the different viewing distances (800, 1500, 3000 mm). The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). The height of the bar represents the magnitude of the VE_r. The log transform of these values were used in the inferential statistics analysis. The actual values are shown in Table 36.

Table 38 Predicted stereoacuity and resolvable depth in mm for the HMPD-2 as a function of pixel size and viewing depth.

Pixel Size	Depth (mm)	Resolvable Depth (mm)
1	800	14.98
	1500	53.30
	3000	214.79
1/2	800	7.49
	1500	26.64
	3000	107.29
1/4	800	3.74
	1500	13.32
	3000	53.63
1/10	800	1.50
	1500	5.33
	3000	21.45

The log transformed VE_r data are presented in Table 39 and are graphically displayed in Figure 69. Visual inspection shows that there is a main effect for distance viewed for the V-HD task. A repeated measures ANOVA confirmed this result, $F(2, 8)$ F = 166, p = .001, ρ_I = .98. Also, a paired samples t-test showed that the H-D pretest for 3000 mm is significantly different from the posttest2, $t(4) = 3.26$, p = .031. There were no other within groups' differences at either the 800 mm viewing distance or the 1500 mm distance. Further, there is no significant difference between the H-D 800 mm viewing distance and the H-D 1500 mm distance.

Table 39 Log transformed variable errors (VE_r) for each participant and group means as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman Log			Virtual-Dolman Log	
		Pre	Post1	Post2	Pre	Post
800	1	0.41	0.43	0.41	0.79	0.81
	2	0.68	0.51	0.55	0.67	0.61
	3	0.37	0.39	0.35	0.65	0.35
	4	0.33	0.40	0.34	0.73	0.74
	5	0.28	0.53	0.66	0.62	0.85
	$\overline{X}$	0.41	0.69	0.46	0.69	0.67
	SD	0.16	0.07	0.14	0.07	0.20
Geometric	$\overline{X}$	2.57	4.89	2.88	4.89	4.67
1500	1	0.59	0.49	0.37	1.44	1.47
	2	0.62	1.04	0.57	1.32	1.25
	3	0.34	0.28	0.30	1.18	1.13
	4	0.43	0.52	0.43	1.13	1.08
	5	0.50	0.70	0.61	1.29	1.24
	$\overline{X}$	0.50	0.61	0.46	1.27	1.23
	SD	0.11	0.28	0.13	0.12	0.15
Geometric	$\overline{X}$	3.16	4.07	2.88	18.62	29.51
3000	1	0.96	0.89	0.80	1.84	1.92
	2	0.94	0.48	0.57	1.65	1.60
	3	0.61	0.62	0.57	1.60	1.58
	4	0.96	0.86	0.71	1.28	1.60
	5	0.97	0.76	0.46	1.74	1.60
	$\overline{X}$	0.89	0.72	0.62	1.62	1.66
	SD	0.16	0.17	0.13	0.21	0.15
Geometric	$\overline{X}$	7.76	5.25	4.17	41.67	45.71

Howard-Dolman Task

Virtual H-D Task
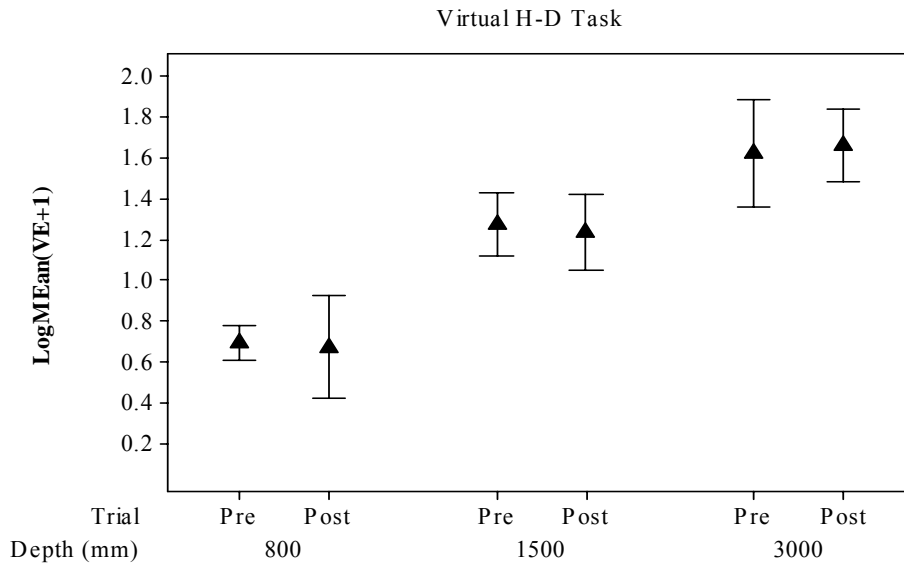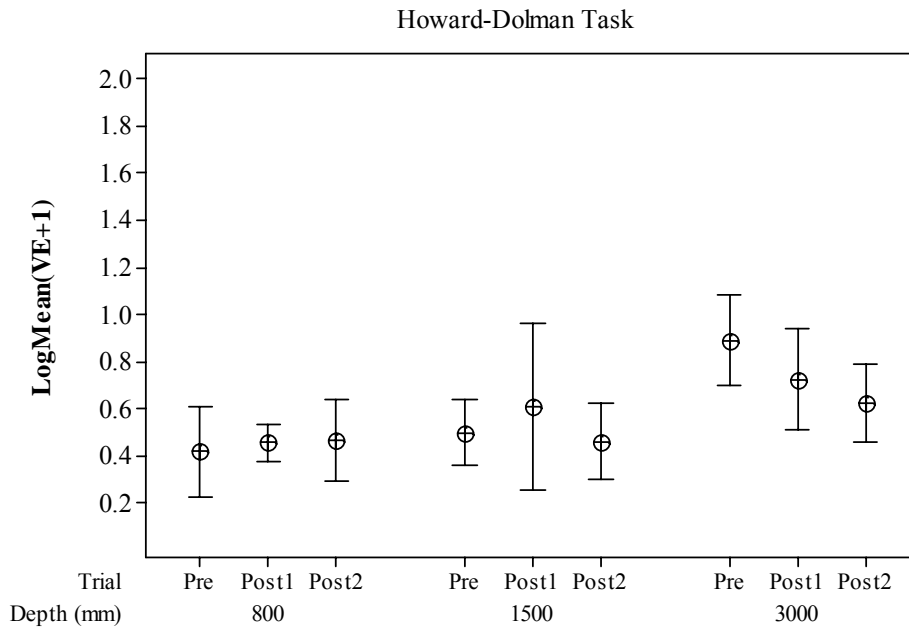
Figure 69 Log variable error (VE$_r$) means and 95% Confidence Intervals for grouped responses on the Howard-Dolman task performed at the different depths (800, 1500, 3000 mm) and on the separate trials (Pre, Posttest 1, and Posttest 2). There are no significant differences on responses across depths or trials on the Howard-Dolman task. However, there is a significant main effect of depth for the V-HD task.

The sign of the stereoacuity value gives an indication of the convergence angle related to the fixed object. As discussed previously, when the convergence angle relating to the fixated object is bigger than that of the moveable object, stereoaccuracy will take on a negative value. The negative bias across all viewing distances shown in Table 40 suggests that the fixed target is being perceived as closer to the observer. This effect is opposite when performing the V-HD task, namely the cylinder is perceived as further away than its fixed location. These results show more variability between trials than in the previous experiments. Also, the variability of the VE_r for the 800 mm viewing distance, regardless of testing environment, is quite high relative to the other viewing distances. Figure 70 depicting the grouped mean stereoaccuracy with their respective 95% CI graphically illustrates this pattern.

Table 40 Stereoaccuracy (η_{ce}) in arc seconds per participant as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman M			Howard-Dolman SD			Virtual-Dolman M		Virtual-Dolman SD	
		Pre	Post1	Post2	Pre	Post1	Post2	Pre	Post	Pre	Post
800	1	32.5	45.0	15.0	20.6	22.2	21.1	2.7	0.5	105.6	111.7
	2	-91.7	-83.7	-65.1	52.9	31.3	36.2	78.2	75.6	79.9	66.4
	3	18.0	27.1	10.3	18.4	19.6	17.0	-62.5	-62.8	73.2	26.5
	4	-5.9	4.8	5.9	14.0	18.8	15.1	59.5	74.9	84.9	89.2
	5	18.4	38.1	62.7	13.3	30.9	45.8	122.2	153.7	65.3	122.8
	$\overline{X}$	-5.7	6.2	5.8				40.0	48.4		
	SD	50.0	52.5	45.7				71.6	82.4		
1500	1	-0.4	10.1	4.4	8.4	6.0	3.9	-51.2	-43.6	54.6	58.0
	2	-15.9	-62.1	-22.6	9.8	31.3	8.6	76.4	26.3	44.5	37.6
	3	10.4	15.0	7.7	3.5	2.7	3.0	-20.0	-9.8	29.4	25.9
	4	-11.0	0.0	-3.1	4.8	6.7	4.8	7.1	-5.5	25.4	22.5
	5	1.3	-7.0	6.6	6.2	11.4	8.7	90.1	5.6	37.7	33.5
	$\overline{X}$	-3.1	-8.8	-1.4				20.5	-5.4		
	SD	10.4	31.0	12.6				61.1	25.5		
3000	1	-11.2	-5.9	-1.4	6.8	5.7	4.5	-51.7	-26.0	35.8	35.8
	2	-26.1	-23.6	-19.6	6.9	5.5	2.4	36.5	39.3	24.6	22.2
	3	3.8	1.1	-0.7	2.6	2.7	2.3	-27.0	8.6	20.6	19.9
	4	-12.0	-21.3	-9.9	10.7	7.6	4.7	14.6	47.3	9.3	20.1
	5	-6.9	-8.3	-4.0	10.9	5.7	2.1	-15.3	-17.3	28.1	20.6
	$\overline{X}$	-10.5	-11.6	-7.1				-8.6	10.4		
	SD	10.8	10.5	7.9				34.7	32.8		
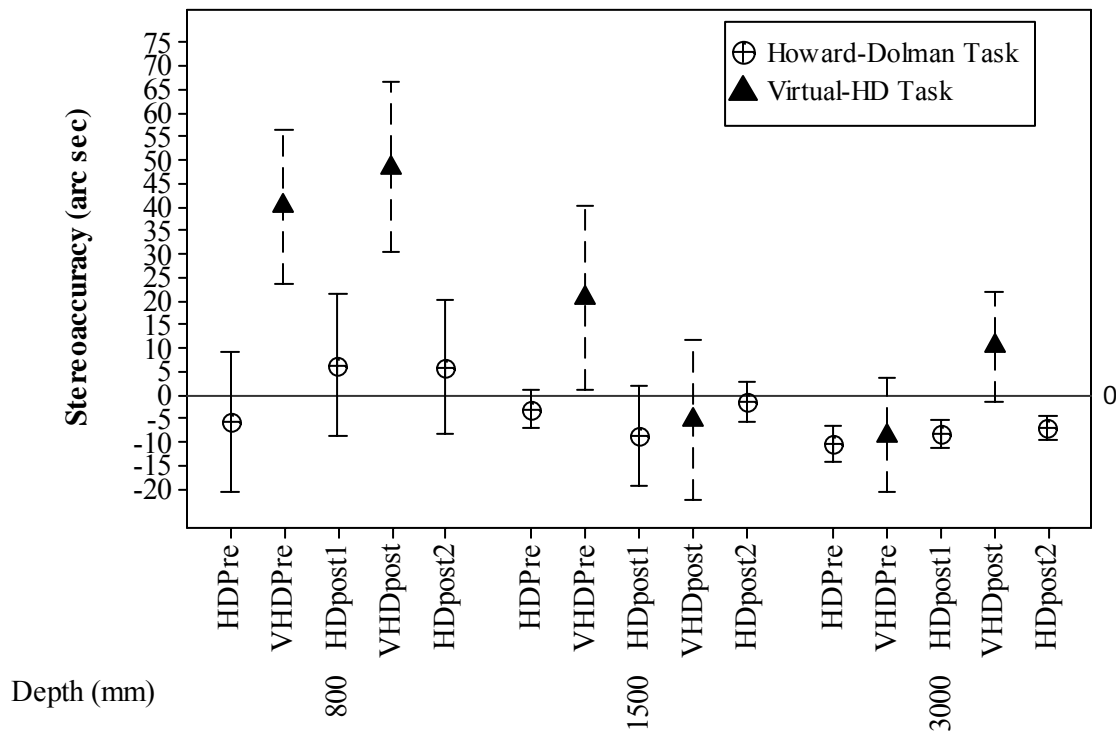
Figure 70 Stereoaccuracy η_{ce} for each viewing distance (800, 1500, and 3000 mm) for the Howard-Dolman task and the Virtual-HD task. The results are ordered in the same manner that they were performed in the experiment (i.e., H-D Pretest, V-HD Pretest, H-D Posttest 1, V-HD Posttest, and H-D Posttest 2). Pooled differences in stereoaccuracy are illustrated along with their biases. The center points represent the mean of stereoaccuracy as calculated from grouping data over participants for each viewing depth (800, 1500, or 3000 mm). The error bars represent the 95% Confidence Interval of the means. A zero value for stereoaccuracy suggests that the participants were able to perfectly align the test objects.

218

Table 41 shows the stereoacuity values attained for each viewing distance per participant. The results show that participants achieve 20 arc seconds or less on the H-D task across each viewing distance. Also, their stereoacuity levels as measured by the V-HD task are better than the 156 arc seconds predicted by the resolution of the microdisplay. The results are graphically depicted in Figure 73.

When performing the H-D task, the trend for stereoacuity is to improve with viewing distance (M_{H-D800} = 23.86, SD = 11.30, $M_{H-D1500}$ = 12.22, SD = 10.68, $M_{H-D3000}$ = 6.15, SD = 3.04). This result is similar to that found with the HMPD-1 in Chapter 7. However, there is no decreasing trend for stereoacuity with increasing viewing distance when performing the V-HD task ($M_{V-HD800}$ = 81.46, SD = 27.01, $M_{V-HD1500}$ = 104.80, SD = 34.70, $M_{V-HD3000}$ = 69.6, SD = 23.05).   As shown in Figure 71, the means are comparable and present with similar variability as shown by the 95% CIs.

Table 41 Stereoacuity (η_{ve}) in arc seconds per participant as a function of depth for the Howard-Dolman task and Virtual-Dolman task

Depth (mm)	Subjects	Howard-Dolman Task			Virtual-Dolman Task	
		$\underline{M}$			$\underline{M}$	
800	1	19.51	21.05	19.98	103.95	109.94
	2	50.21	29.73	34.36	78.64	67.49
	3	17.48	18.63	16.10	72.07	26.08
	4	13.30	17.80	14.32	83.60	87.75
	5	12.60	29.32	43.44	64.23	120.84
	$\overline{X}$	22.62	23.31	25.64	80.50	82.42
	SD	15.69	5.80	12.69	14.99	37.61
1500	1	12.80	9.22	5.89	155.00	164.78
	2	15.07	47.99	13.10	126.28	106.89
	3	5.38	4.08	4.55	83.50	73.55
	4	7.42	10.21	7.36	72.02	63.96
	5	9.43	17.48	13.36	107.09	95.14
	$\overline{X}$	10.02	17.80	8.85	100.86	100.86
	SD	3.93	17.54	4.12	39.57	39.57
3000	1	10.47	8.72	6.83	101.59	99.59
	2	10.49	2.73	3.74	70.87	87.37
	3	4.04	4.19	3.58	58.55	56.48
	4	10.25	8.01	5.30	26.54	57.18
	5	10.88	6.07	2.42	79.75	58.41
	$\overline{X}$	9.23	5.94	4.37	67.46	71.81
	SD	2.91	2.52	1.71	27.76	20.26
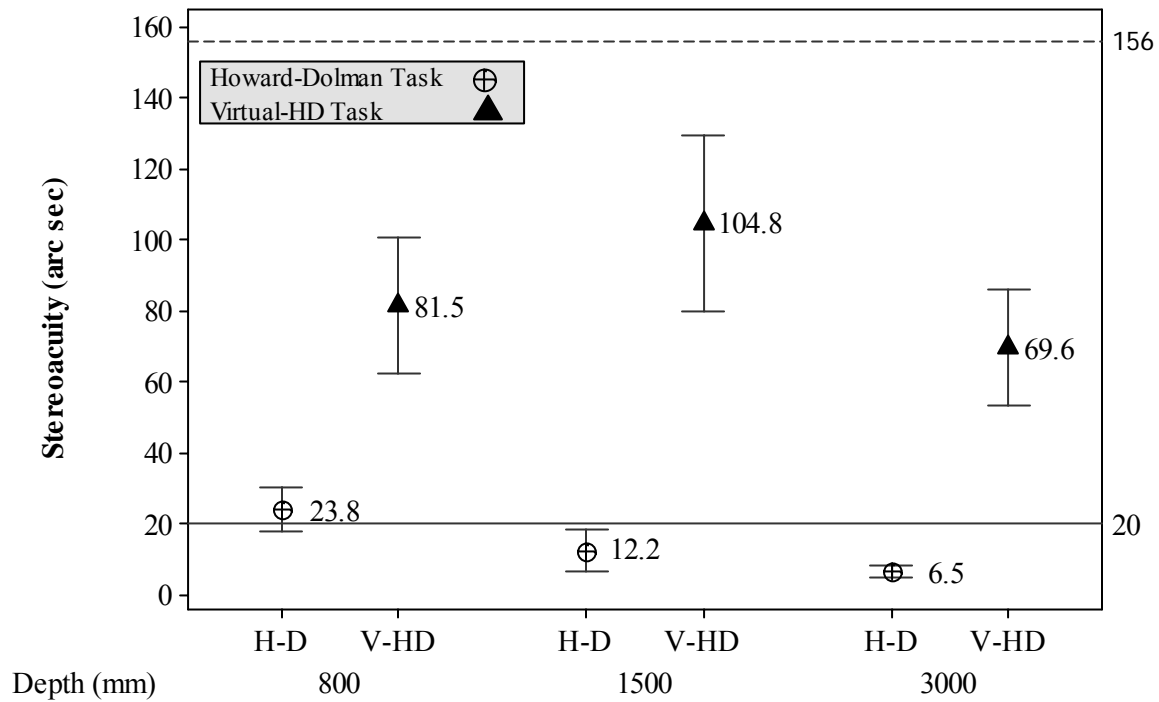
Figure 71 Stereoacuity η_{ve} means and 95% Confidence Interval for the participants' grouped responses on the Howard-Dolman and Virtual-HD task performed at the different depths (800, 1500, 3000 mm) and on the separate Pre and Posttest trials. The mean averages illustrate the change in stereoacuity over the different viewing depths. The reference line at 20 arc seconds represents average responding in the population at large. The reference line at 156 arc seconds represents the predicted stereoacuity based on the HMPD-2 parameters.

As expected, the means for the dark focus were relatively stable over each testing trial as shown in Table 42. Further, participants did not vary considerably from one another on the dark focus metric. Figure 72 shows the mean meter-angle measures for dark convergence with respect to the participant and viewing distance. While participants' dark convergences are consistent with respect to their own repeated measures, there is wide variability between participants on the measure of dark convergence. There is evidence of participants separating into near and far dark convergence groups, which is reported in the literature (Owens, 1984). However, the variability precludes this response variable from adding meaning to the analysis. If the clustering is a true effect, the results do demonstrate that individual differences are important to quantify and control for when performing these types of tasks within a virtual environment.

Table 42 Mean dark focus in Diopters and dark convergence in Mean-Angles per participant as a function of depth and testing condition

Depth (mm)		Dark Focus M			Dark Convergence M		
		Pre	Post1	Post2	Pre	Post1	Post2
	1	1.9	1.9	2.4	0.6	0.4	0.4
800	2	2.9	3.2	3.2	0.7	0.7	0.7
	3	1.4	1.6	1.6	2.7	2.4	2.5
	4	2.4	1.9	1.8	0.9	0.8	0.7
	5	1.9	1.7	2.6	2.7	2.0	2.2
	$\overline{X}$	2.1	2.0	2.3	1.5	1.3	1.3
	SD	0.5	0.7	0.6	1.1	0.9	1.0
1500	1	2.0	2.0	1.9	0.4	0.4	0.4
	2	2.3	2.6	2.3	0.8	0.7	0.8
	3	1.5	1.3	1.3	3.2	2.7	3.3
	4	2.1	2.3	1.6	1.2	1.3	1.1
	5	1.9	1.9	1.8	2.2	1.8	2.0
	$\overline{X}$	1.9	2.0	1.8	1.6	1.4	1.5
	SD	0.3	0.5	0.4	1.1	0.9	1.2
3000	1	1.8	2.8	2.6	0.4	0.4	0.4
	2	2.8	2.4	3.1	0.7	0.7	0.7
	3	1.4	1.1	1.2	2.2	2.8	2.8
	4	2.0	1.9	1.9	1.1	1.3	1.1
	5	1.5	2.0	1.4	2.3	2.2	1.7
	$\overline{X}$	1.9	2.0	2.0	1.3	1.5	1.4
	SD	0.6	0.6	0.8	0.9	1.0	1.0
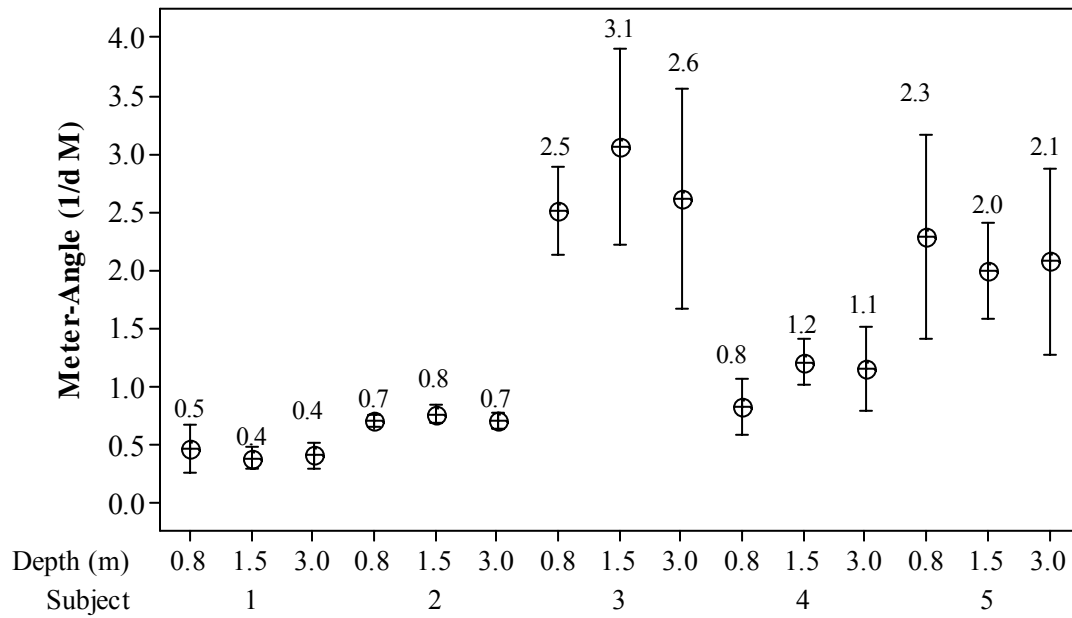
Figure 72 Dark vergence means per participant grouped over Pretest and Posttest at each viewing depth (800, 1500, 3000 mm). The 95% Confidence Interval for each vergence mean is given. The mean averages illustrate that participants do not change in dark vergence over the different viewing depths.

Table 43 displays the front and back discriminability of the participants across viewing

distances and trials. Given that the threshold for the 2AFC design is 75 %, the participants

performed well above this limit. The means as shown in Table 43 did not go bellow 90 %, even

for the viewing distance of 3000 mm. The data from Table 42 are graphically displayed in Figure

73.

Table 43 Percent correct for front and back responses as a function of HMD for the Virtual-
Dolman task

Depth mm		Virtual-Dolman Task $\underline{M}$	
		Trial 1	Trial 2
800	$\overline{X}$	0.95	0.93
	SD	0.01	0.04
1500	$\overline{X}$	0.96	0.94
	SD	0.03	0.03
3000	$\overline{X}$	0.96	0.90
	SD	0.02	0.12
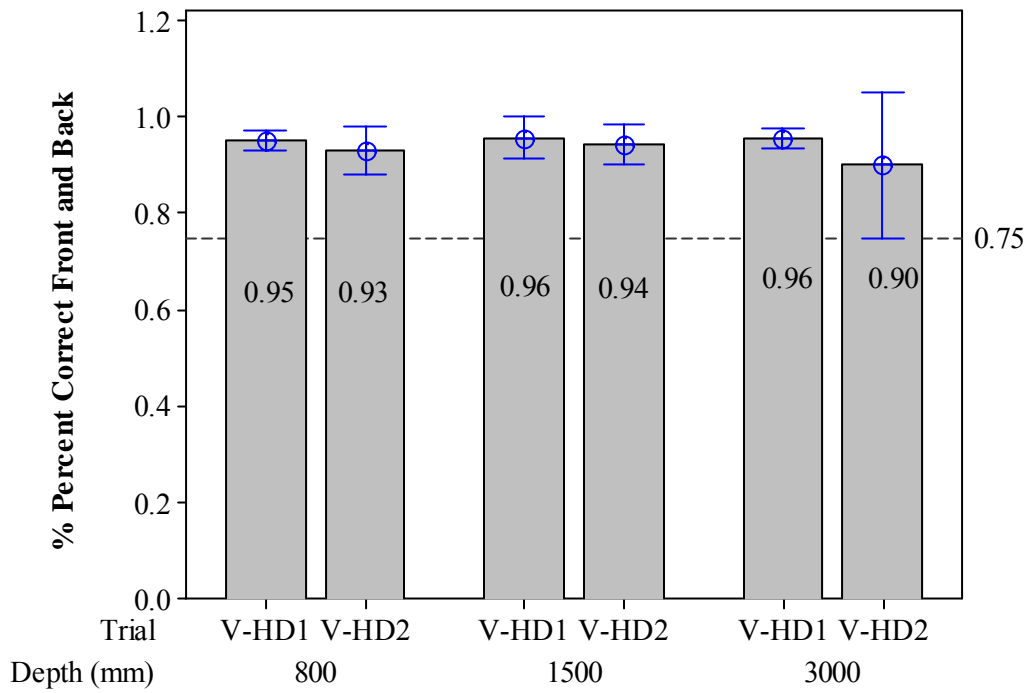
Figure 73 Percent correct for front and back judgments on both trials of the V-HD task grouped over HMD type. The mean and 95% CI are given. The reference line at 75% represents the detection threshold for a 2AFC design.

As expected, Table 44 shows that very few participants in this study experienced cybersickness. In fact, the same subject experienced different levels of cybersickness across the viewing distances. Inspection of VE$_r$ scores in Table 44 shows that participant 2 shifted from a CE of -3.3, SD = 3.3 to 12.9, SD = 10.05. This shift corresponds to the highest reported values of cybersickness for the experiment. With more data a more causative link between performance of the user and level of cybersickness could be achieved.

Table 44 Cybersickness scores for the oculomotor subset per participant for HMD type and testing condition

Depth		SSQ	
		OD	
		Post1	Post2
800	1	0	0
	2	30	15
	3	8	0
	4	8	0
	5	0	0
1500	1	0	0
	2	45	53
	3	0	8
	4	0	0
	5	0	0
3000	1	0	0
	2	15	0
	3	8	8
	4	0	0
	5	0	0

8.5 HMPD-2 Error Metrics and Viewing Distance Conclusion

The results of the V-HD task show that participants achieved stereoacuity levels at each viewing distance which were less than the predicted stereoacuity set by the microdisplay, $M_{V\text{-}HD800} = 81.46$, $SD = 27.01$, $M_{V\text{-}HD1500} = 104.8$, $SD = 34.7$, and $M_{V\text{-}HD3000} = 69.63$, $SD = 23.02$. The participants VE_r suggest that the resolvable depth for this HMD is ¼ subpixel resolution. How this result extends to usability or task based performance is unclear; however the results do suggest that we are correct in assessing the HMD, the graphics, and the user as an interdependent system. Thus, these metrics have allowed for multiple conservative views of the data to show where the technology limits occur versus where individual differences of the user play a role in overall outcomes.

For example, out of 5 participants, participant 2 experienced cybersickness on a scale suggestive of visual disturbances. Concomitant with the increased levels of cybersickness are shifts in performance on the H-D task. Although we do not have a causative link between cybersickness and possible adaptive responses on the H-D task, we are now building the capability of appropriately attributing error in the VE system to a more specific cause (e.g., user versus technology). Given that cybersickness may follow a pareto distribution whereby 20% of those exposed to virtual environments exhibit 80% of the cybersickness symptoms, the ability for the V-HD task to identify those individuals most susceptible to cybersickness before data analysis is critical to the overall assessment of the virtual environment (Kennedy, Kennedy, & Drexler, 2006).

Further, participants for this experiment reported less cybersickness over all the studies reported. This suggests that although there may be mechanical issues associated with this prototype display, the design of this display is moving in a direction capable of supporting a

broader range of tasks without compromising the user. As evidenced by the percent correct responding on front and back for the octahedron stimuli, participants in this study were clearly able to discern the front and back position of the object as presented on a curved surface of retroreflective material. All participants were able to discern depth into the plane of the retroreflective material and out of the plane of the material. For example, at the 3000 mm viewing distance, the virtual octahedron could be placed 30 cm into real space with out disruption of the space continuum. Alternatively, the space 30 cm into the plane of the material was discerned equally.

We plan to explore this result further with real and virtual objects to determine if the users perceive the material as a continuous space with real objects placed into view. Continuing this approach of methodologically evaluating the complete virtual environment system will allow us to scope out the best use of such systems over a wider range of tasks.

Although dark convergence did not provide a means of determining the source of localization errors in this study, we do have evidence of clustering of participants into far and near dark convergence. This dichotomy may assist in naturally classifying participants to provide better controls for experiments or to provide more meaning to the data analysis. As argued throughout this work, extending current virtual environment design to the domain of cognitive rehabilitation is more than choosing the correct HMD. As researchers, we have to identify all potentially confounding variables to assure that the rehabilitation protocols we follow are delivering the appropriate remediation for that given population.

CHAPTER 9 HEM TESTBED: FEASIBILITY STUDY FOR COGNITIVE REHABILITATION

9.1 Introduction

Szekeres, Ylvisaker, & Holland (1985) encouraged rehabilitation specialists "to generate novel intervention procedures while measuring their value against patient progress, not academic acceptability" (p 220). This statement is telling of the disconnect that sometimes exists between research for academia (e.g., push to publishing) versus the practical application of the basic research in a clinical setting. For example, how practical is a VE rehabilitation solution for private practice? The answer to that question depends on many factors, such as type of practice (e.g., physical versus physical rehabilitation) and needed equipment. However, how does a practitioner know what type of VE system they need? The answer lies with the universities, especially those that support technology transfer.

While technology can lead to profit for universities (Pounds, 2004), the transfer of such technology can lead to community business partnerships that supply direction and focus to developing the product further (Nature Editorial, 2006). Given that TBI is a serious community problem, it will need a supportive community effort to match the high-tech potential of university laboratories with the practical goal of independent rehabilitation specialists. Thus, VE technology developed for cognitive rehabilitation should not be inaccessible to the private practitioner. Although, most business models have as their main goal increased revenue, the guiding principles of such a community collaboration should be accessibility and financial independence such that the monies received continue to support existing technology development while encouraging innovation.

As first steps to building this community partnership, the Human Experience Modeler (HEM), developed at the University of Central Florida, was designed as a testbed to evaluate technologies of virtual, augmented and mixed reality that may enhance cognitive rehabilitation effectiveness (Fidopiastis et al., 2006). As Rizzo et al. (2005) acknowledged, a trial and error approach to VE design for rehabilitation is commonly employed due to the lack of appropriate methodologies for evaluating user interactions. The authors also contend that the expense and ever changing technology poses a challenge for proper benchmarking of VE systems for use with clinical populations.

The HEM provides a testbed with which to establish benchmarks for different system technologies and potential types of user interaction. The goal of such technology is forward looking from specific applications of cognitive remediation to more advanced transfer of training protocols. In both cases, benchmarking of the VE system supports the experimental design nature of cognitive rehabilitation approaches. As argued previously, without benchmarking, separating system error from the user is impossible and may lead to inappropriate conclusion about the technology or the rehabilitation method under study.

In this chapter we review results of a pilot study completed with the current version of the HEM. The goal of this initial study was to test the integrated components of the HEM and assess its feasibility for presenting contextualized environments that are personally meaningful to the patient (e.g., a room in the patient's home). The user interactions in such environments are more complex and require more realism than other rehabilitation tasks (e.g., pain distraction). The contextualization and the level of experimental control afforded by utilizing a VE framework allows researchers a unique opportunity to further explore transfer of training issues with brain injured populations.

This work is an extension of Zhang et al. (2003)'s meal preparation study. The authors

built a fully immersive generic kitchen as shown in Figure 74a. The participant was given a

scripted task and procedurally followed the steps for preparing a can of soup.

Figure 74 Fully immersive generic kitchen versus the HEM mixed reality kitchen

a   Model of Zhang et al.'s virtual kitchen   b Chroma-key kitchen for HEM

Within the HEM version of the task, the participant's home kitchen was modeled using both

virtual and real objects as shown in Figure 74b. The resulting mixed reality environment was

viewed through the Canon Coastar previously discussed. Further differences between the studies

are outlined in the next section.

9.2 HEM Breakfast Meal Preparation Methods

<u>Participant</u>

The participant for this study was a right-handed, 48 year old, Caucasian male who suffered an aneurysm in 2004 and presented with left frontal lobe damage. Prior to participating in the experiment, the participant was administered the following test battery from a speech/language pathologist and a psychologist: *Boston Diagnostic Aphasia Examination 3rd Edition (BDAE-III), Cognitive Linguistic Quick Test(CLQT),* oral peripheral examination, hearing screening, *Mini-Mental Status Exam (MMSE), Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), Wechsler Adult Intelligence Scale 3rd Edition (WAIS-III),* and the *Comprehensive Trail-Making Test (CTMT).* The participant's testing showed attention, memory, and executive functioning impairments consistent with frontal lobe damage. However, his ability to manipulate and process visuospatial stimuli remained intact. The participant's testing showed attention, memory, and executive functioning impairments consistent with frontal lobe damage. However, his ability to manipulate and process visuospatial stimuli remained intact. The participant was independent in his personal activities of daily living (ADLs) in his home but was not engaged in his meal preparation.

<u>Instrumentation</u>

The HEM is a virtuality system based on mixed reality (MR), which allows modeling of a complete real world setting (Hughes, Stapleton, Hughes, & Smith, 2005). As a first step, the spatial, audio, and visual environment of a location is captured utilizing 3D laser, image, and

acoustical recordings to accurately reproduce a space and its multi-sensory signature. Once captured, this real environment can be rendered within a mixed reality (a mixture of real and virtual objects and environmental conditions). Specifically, images are processed by the Mixed Reality Software Suite (MRSS), software tools developed for creating dynamic and interactive mixed reality experiences (O'Connor & Hughes, 2005).

The MRSS system is made up of four subsystems: three rendering engines simulate the multimodal simulation (Visual, Audio and Special Effects) while a fourth engine drives the integration and creates an interactive, non-linear scenario (story) of the chosen human experience. The output of the MRSS system is a mixed synthetic and real setting where an integrated system of sensors in the environment captures the user's performance for replay. Peripheral and environmental perception is rendered with a combination of 3D sensory displays such as the Canon Coastar, audio earbuds (earphones for Mixed Reality), surround sound and spatially registered audio, special effects (e.g., breezes and opening/closing doors) haptic vests and olfactory stimulation.

HEM Implementation

For this study, a depth camera (3DV Systems DMC100) and a 3D laser scanner (Riegl LMS420i) were used to capture the participant's home kitchen. The scanners output was edited in 3D Studio Max. The graphical representation of the kitchen, including the textures and appliance overlays, were imported into the MRSS. These images were seen through the Canon HMD discussed earlier.

In addition to the 3D graphics, parts of the real kitchen were reconstructed out of plywood to match the same dimensions and location of real stationary components of the kitchen (e.g., cupboards). Figure 75a shows a picture of the participant's own kitchen. The reconstructed

234

mock kitchen painted green for chroma-keying is represented in Figure 75b. There were some differences in the dimensions, brand of appliances, and number of fixtures between the mixed reality and real kitchen. Otherwise the spatial layout of target locations and target items was identical to the participant's own kitchen as shown in Figure 75c.

Nonessential areas such as the stove, microwave, and dishwashing machine were all virtually represented in the MR kitchen. Essential areas, such as the plate and cup cabinets along with the silverware drawer were fully functional as were the pantry door and refrigerator. The refrigerator contained milk and other non-target items, for example orange juice. As well, the pantry contained non-task relevant boxed items such as cake mix.

Figure 75 (a) Participant's actual kitchen. (b) Chroma-keyed mock-up. (c) Schematic of locations of target items and typical starting position of the participant.

MR Training Procedure

This study was approved by the IRB for the University of Central Florida. The protocols are listed in Appendix A. The participant was videotaped performing meal preparation in his own home at his regular breakfast hour. This single measure baseline followed a modified "Goal-Plan-Do-Review" or executive function map (Ylvisaker, 1998) where by the participant verbalized and wrote down a list of steps and materials needed to accomplish his goal. He was then asked to perform the cereal making task with the assistance of the therapist. The participant

235

was free to perform each identified step in any order, which was different from the Zhang et al procedure.

During each MR training session, the participant was fitted with the Canon, which was coupled with an IS900 wireless mini tracker. The participant was presented with the same protocol as was performed at home. There were 5 training sessions conducted over consecutive days. The Tracker Reviewer, a Java based software capturing the tracked movements of the participant, was started once the participant signaled that he was ready to begin. The facilitator remained in the virtual kitchen with the participant for the entire meal preparation. In addition, the facilitator monitored the participant's view of the MR on a display placed out of the line of sight of the participant. Once the participant had completed the cereal preparation task in MR, the SSQ was administered.

Upon completing the MR training, the participant was videotaped in his own home performing the cereal making task. The participant generated the script sequence and proceeded to make breakfast. Only one post-training home session was recorded due to time constraints.

9.3 HEM Transfer of Training Results

Location errors, time to locate target items, total time to complete the task, order of item retrieval, and number of cues were recorded in the real and simulated kitchen. In addition, total efficiency of movement was captured from the head tracker via Tracking Reviewer. Analysis of the overall location errors (total errors = 42) showed that the participant had the most difficulty remembering where the cereal (17 total errors) and the bowls (13 total errors) were kept in both his real kitchen and in the MR Kitchen.

Table 45 displays the time in seconds that the participant needed to locate a target item. Time was measured starting from when the participant verbalized what item he was about to retrieve and ending at the time he found the item in its appropriate location within the kitchen. The numbers in parentheses represent the order that each item was retrieved. Item retrieval order during the actual breakfast making task never matched the preplanning script. However, the participant's performance retrieval order (bowl, spoon, milk, and cereal) was the same for 3 out of the 5 MR training sessions and the post-test home session.

Table 45 Time in seconds to locate target items; order retrieved; and time to complete all tasks.

Location	Target	Pre	MR 1	MR 2	MR 3	MR 4	MR 5	Post
Pantry	Cereal	58(1)	112 (4)	76(4)	119 (4)	80(2)	18(3)	8 (4)
Cabinet	Bowl	13(2)	38 (1)	26(1)	43 (1)	28(3)	27(1)	18 (1)
Refrigerator	Milk	14(3)	44 (3)	24(3)	12 (3)	7(1)	14(4)	14 (3)
Drawer	Spoon	3 (4)	15 (2)	15(2)	3 (2)	6(5)	5 (2)	3 (2)
Counter	Make Cereal	73(5)	120 (5)	84 5)	41 (5)	40(4)	37(5)	51 (5)
	Total Time (s)	**240**	**379**	**315**	**341**	**236**	**177**	**158**
	Total Time (m:s)	**4:00**	**6:19**	**5:15**	**5:41**	**3:56**	**2:57**	**2:38**

Table 45 also shows that the participant took 4 minutes to prepare cereal in his own home during the pre-test, while he took over 6 minutes to make breakfast during the initial MR training session. By the end of the last MR training session, the breakfast task was performed in half this time. The participant's fastest cereal preparation time (2 minutes and 38 seconds) was achieved in his home during the post-test session. This result is shown graphically in Figure 76.

237
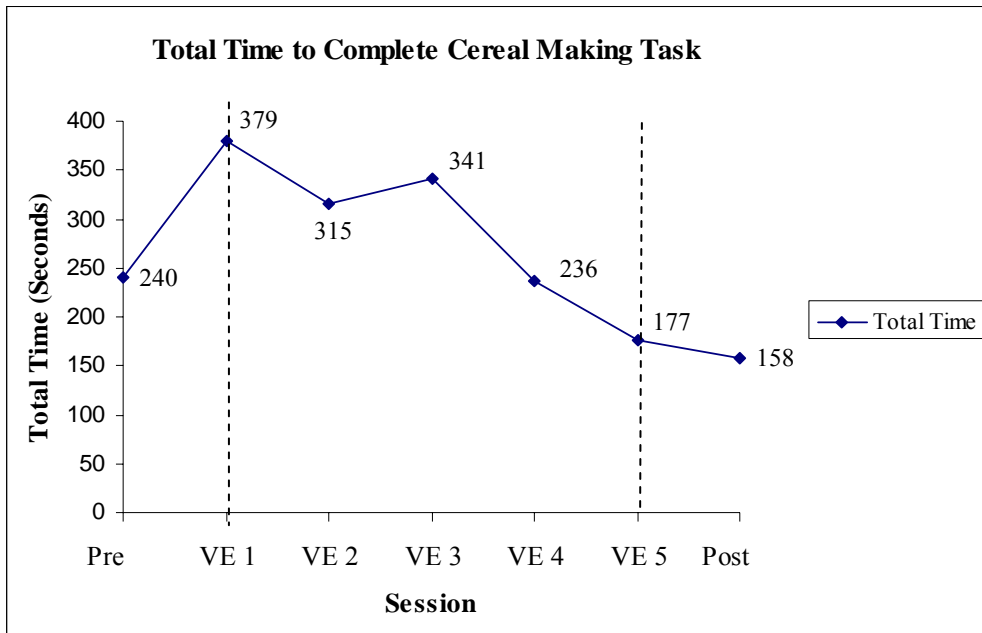
Figure 76 Time to complete the breakfast making task for all conditions

Figure 77 displays the captured tracking data of the participant as he prepared cereal in the MR kitchen during the first day of training and the last day of training. The overall track data visually highlights differences in movement behavior from an efficiency standpoint between the initial MR training and the final MR training session. The track in Figure 77a is very similar to the one presented in Figure 78a, a line drawing representation of the participant's track during the pre-testing home session, suggesting that the participant carried over errors in his searching behavior to the MR kitchen. Interestingly, in the MR kitchen, the nontarget areas were virtually represented. While the participant can open the target cabinet to find the bowls inside, the nontarget virtual cabinets only display the texture seen in the real kitchen.

Figure 78b shows that the participant makes fewer location errors and more efficiently completes the task during his final training session in the MR kitchen. In his own kitchen during

the post-test, the participant demonstrates more continuous movements between steps with less

searching behavior as evidenced by the line drawing track shown in Figure 78b.
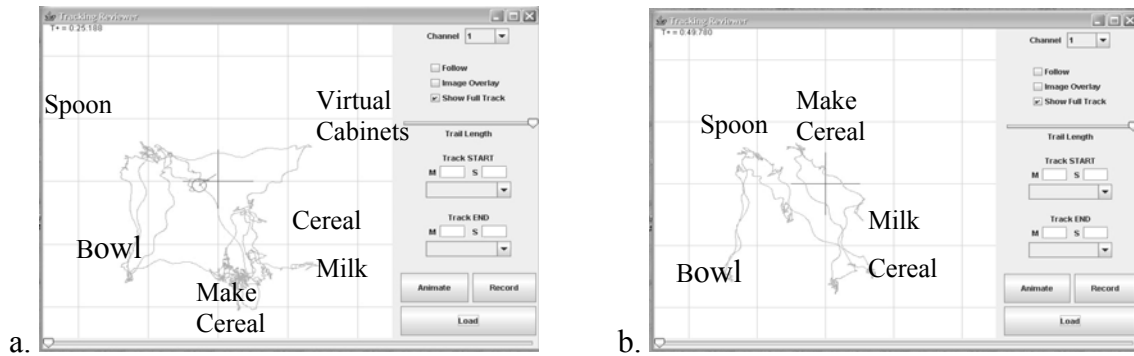
Figure 77 (a) Tracker Reviewer data from the participant's first MR training session; retrieval order was bowl, spoon, milk, cereal. (b) Tracker Reviewer data form participant's last MR trainings session; retrieval order was bowl, spoon, cereal, milk. Participant made cereal in different locations on each day.
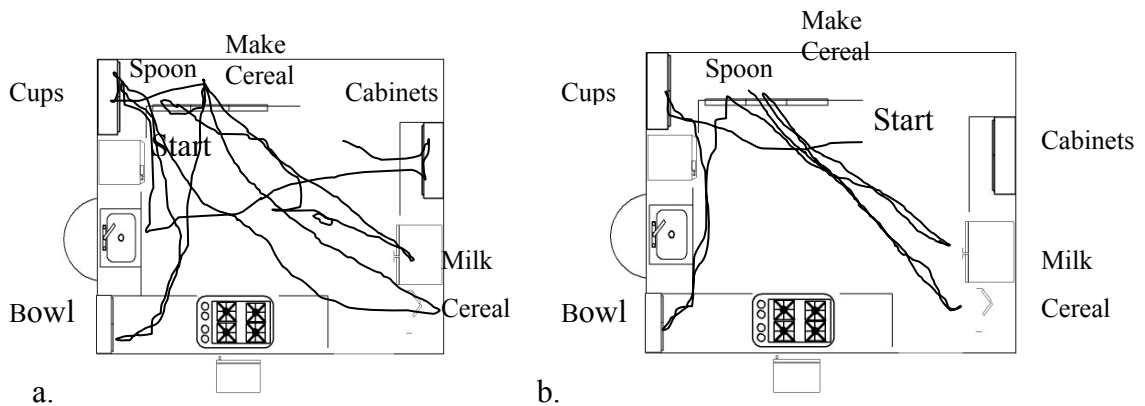
Figure 78 (a) Track at home during pre-training; order was cereal, bowl, milk, and spoon. (b) Track of post-training; order was bowl, spoon, milk, and cereal. Participant made cereal in same location each day.

9.4 HEM Conclusion

The participant's improvement over the 5 MR training sessions is evidence that the HEM mixed reality environment was an ecologically-valid environment within which a brain damaged adult could learn new information applicable to his home environment. Transfer of learning from the MR environment to his home environment was evidenced in decreased time spent on task, decreased number of location errors, and decreased wandering behavior. Research questions such as the degree and extent of replication needed for transfer of learning are currently being explored.

The Tracker Reviewer allowed us to capture the participant's internalized learning process that may have been undetected by traditional rehabilitative methods. More specifically, we were able to capture the implicit ordering effects of the participant who had no memory that he was performing the task similarly on each trial. The ability to scientifically visualize the process and then to track that process permitted analysis of emergent trends in the participant's performance. For example, the order of retrieval of target items possibly suggests that the training assisted the patient in developing an internal script which then carried over to the home environment. Although further research in this area is needed, the results are encouraging. Additionally, the HEM mixed reality environment afforded the interactive human component in a flexible arena in which the participant could safely explore his functional capabilities. However there are some drawbacks to the current setup.

The current VE system setup is not easily reconfigurable for accommodating other participants. Although it is true that cabinetry and counter tops come in standard sizes, this plywood configuration would be difficult to alter. We are currently investigating new materials to enhance the changeability of the environment before we attempt testing different technologies

and levels of contextualization. Also, the participant's memory impairment singly determined applicability of some features of the HEM.

For example, the HEM has after-action review features that allow the participant to review his or her performance. When performing the after-action review, the participant became extremely emotional when watching himself struggle to perform the task. In addition to his memory impairment and the length of the after-action review process, we did not utilize this feature on subsequent training trials. Although adequate for the task, the HMD caused some visual problems that the participant was unable to report. For example, when reaching for some items he would misjudge the distance and most likely would overshoot the target and collide with mock setup. Although minor, with other patients this misperception of space could be an issue. Unrelated to the HEM another potential problem was the administration of the SSQ.

The SSQ is a pretest posttest type of questionnaire. In our procedure, we removed the HMD and allowed the participant to take his bowl of cereal to a table area to eat. During the time it took him to sit down, memory of the training session was already degrading. He could not remember difficulties that he experienced during the training. For example, during one training set, after retrieving the bowls and setting them on the counter, he did not see them after he retrieved another item. His solution was to retrieve another set of bowls; however, we were unsure of why he was unable to locate the bowls he retrieved initially. We could not determine if he was experiencing visual problems with the display or just forgot where he had placed them. Similarly, although the SSQ can be thought of as a "current state" monitoring assessment, we could not determine if the participant was reporting about his status with respect to his TBI symptoms or the VE exposure. Thus, a cybersickness assessment that did not rely on verbal

241

report may be appropriate for some patient populations. This issue also supports the need for baseline metrics for participants performing any type of training within a VE.

The results of this pilot are encouraging and suggest that the HEM will provide a standard environment within which to perform technology comparison studies along with benchmarking and quantifying multimodal interactions. The goal of such work is to provide guidance on how best to design VE rehabilitation spaces to maximize transfer of training potentials for each individual, regardless of their rehabilitation needs. In addition, the optimized, scaled version of the optimal training system could be offered commercially to reach more rehabilitation providers and consequently more person with TBI.

CHAPTER 10 HMD ASSESSMENT, USABILITY & COGNITIVE REHABILITATION

10.1 Overall Results Summary

This work encompasses many fields of research: HMD Design, Human Factors approaches to VE Design, and VE based Cognitive Rehabilitation. The combined effort of this work is to bring the theoretical best practices of these fields together to create a VE rehabilitation environment that is safe and acceptable to the end user, while ensuring the validity and reliability of the outcome metrics employed by the rehabilitation protocol. The foundational links that unify these currently self-governing fields within a common framework is that of VE technology benchmarking and user performance baselines. These metrics should be obtained early in the VE design cycle to guide and support the decision making process that culminates into a final VE application.

More specifically, we argue that current user-centered design approaches are too high level to support the empirical nature of VE rehabilitation. When building a specific rehabilitation application, it is appropriate to introduce sensory performance assessments after the initial task analysis is performed. However, general sensory baselines attained via a user-in-the-loop methodology are the responsibility of the device manufacturer. The goal of the therapist is not to determine the individual limits of the technology from a user perspective, but to define the integrative effects of technology as they assist or detract from user performance in a rehabilitation scenario. We provide two examples of benchmarking prototype see-through projection displays through the use of resolution visual acuity and stereoacuity measures of human observers.

As outlined in Chapter 4, a modified Landolt C test was designed following accepted standardized guidelines and was used to measure resolution visual acuity of participants who performed the assessment while wearing the HMPD-1. The results confirmed that the microdisplay resolution of 4.1 arc minutes constrained the users' rVA to that value as well. However, in the case of both the Canon and the VR6 displays while the resolutions of 5.83 and 4.78 arc minutes were predicted, we measured resolutions of 11.66 and 9.56 respectively. By assessing the user's rVA under real world viewing conditions we were able to compare for changes in rVA caused by the VE exposure. More importantly, we could directly measure the user's rVA. Thus the modified Landolt C assessment provides additional controls for modifying potential confounding variables that may affect the final outcome of the analysis.

In an extension of the study recapped above, we added optotypes of varying contrast to assess the interplay of the HMPD-1 characteristics with those of two types of retroreflective materials currently used to project the virtual image back to the user as explained in Chapter 5. The low contrast version of the modified Landolt C test allowed us to screen participants prior to performing the assessment in the VE. We found that persons who had LASIK surgery or had a prior injury to the retina of one eye could not perform the real world viewing portion of this test. As noted, these participants had passed the high contrast version of the test, the results of which determined the limiting rVA of the HMPD-1. Additionally this study showed that participants performed poorest when viewing the optotypes on the beaded retroreflective material. This result provided guidelines for optimizing the VE system (e.g., HMD, retroreflective material, and computer software) for applications that require viewing of low contrast targets (i.e., medical). However, the results have application to cognitive rehabilitation as well. Assuming that the

patient can perform the test, understanding their rVA limitations given the VE system setup is paramount to understanding the effectiveness of the VE based protocol.

In the next series of studies, we outlined a methodology for attaining stereoacuity measures for real world (HD-task) and virtual world (V-HD task) viewing. The V-HD test was administered over different HMDs to compare the specifications given in the owner's manual with those measured by the assessment. The metrics derived from the assessment (i.e., CE, |CE|, VE_r, η, % correct for front and back) served the dual purpose of identifying potential user characteristics for future research and profiling the capabilities of each HMD.

User individual characteristic such as dark focus and dark convergence could assist in signaling cybersickness or adaptation effects that may affect the outcome of the assessment. Although the sample sizes used in each experiment were too few to gauge the merit of including these measures, there were indications such as significant shifts in dark focus and clustering of participants within high and low groupings of dark convergence that suggest their importance as test metrics. The variability in individual responding during the V-HD test suggests that higher sample sizes are needed to understand these differences from a usability perspective. However, as an assessment for HMD performance, this methodology is an excellent example of user-centered approaches.

The V-HD assessment was able to quantify stereoacuity limitations in a range of HMDs. Each of which performed better than the predicted resolution based on their respective microdisplays pixel size. As argued in Chapter 3, the VE technology, the computer graphics, and the user are sustained in an interdependent relationship that ultimately defines the quality of the resulting VE application. The fact that all HMDs achieve subpixel resolution at the ½ level or better points to future research on the interdependence of computer graphics techniques with

respect to image resolution. We attribute the improved resolution as a result of antialiasing techniques that improve visibility of the edges of a rendered object. One could test this hypothesis by varying the amount of aliasing within each V-HD stimuli. Such studies will assist in improving the V-HD design and allow us to understand more of what information (e.g., lighting, contrast, and aliasing) most affects the perceptibility of the user.

Figure 79 shows the overall stereoacuity measured for each HMD tested, including the HMPD-2. The HMPD-2 has the best resolution (2.63 arc minutes or 157 arc seconds), and achieves better than 157 arc seconds on the overall stereoacuity measured. Each HMD performs better than the resolution of their respective microdisplay would predict.
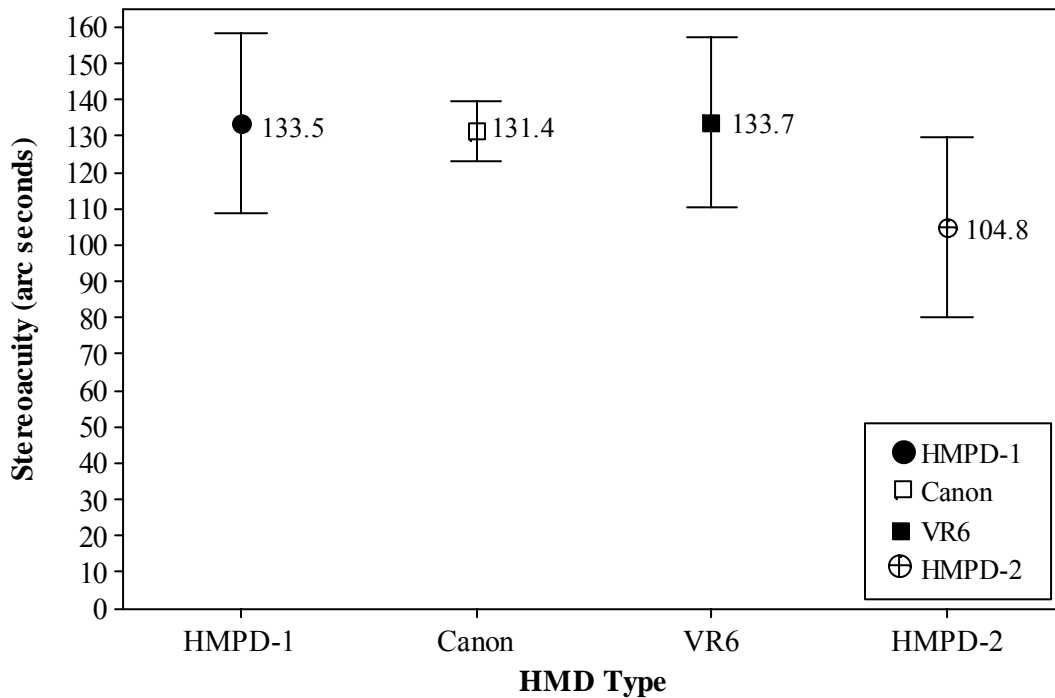
Figure 79 Overall stereoacuity for all HMDs tested

Without knowledge that the performance of the HMPD-1 is compromised by the manner in which we set the focal plane, the results suggest that the HMPD-1, Canon, and VR6 have equivalent stereoacuity. Thus, one may think that users should perform well on depth discrimination tasks regardless of these three HMDs. However, the percent correct front and back for identifying the initial position of the octahedron with respect to the cylinder, as depicted in Figure 80, shows that participants perform below the threshold when wearing the Canon or the VR6. Participants wearing the HMPD-2 are 95% accurate in their responses. This result has not been further investigated. However, the HMPD-2 distinguishes itself with a much higher visual acuity capability, contributing to better sharpness of the image. Also, HMPD-1 has been shown to be superior to Canon and VR6 in resolution visual acuity, while their stereoacuity is equivalent. Therefore the results indicate that resolution visual acuity may be a main contributor to maximizing the percent correct for front and back.
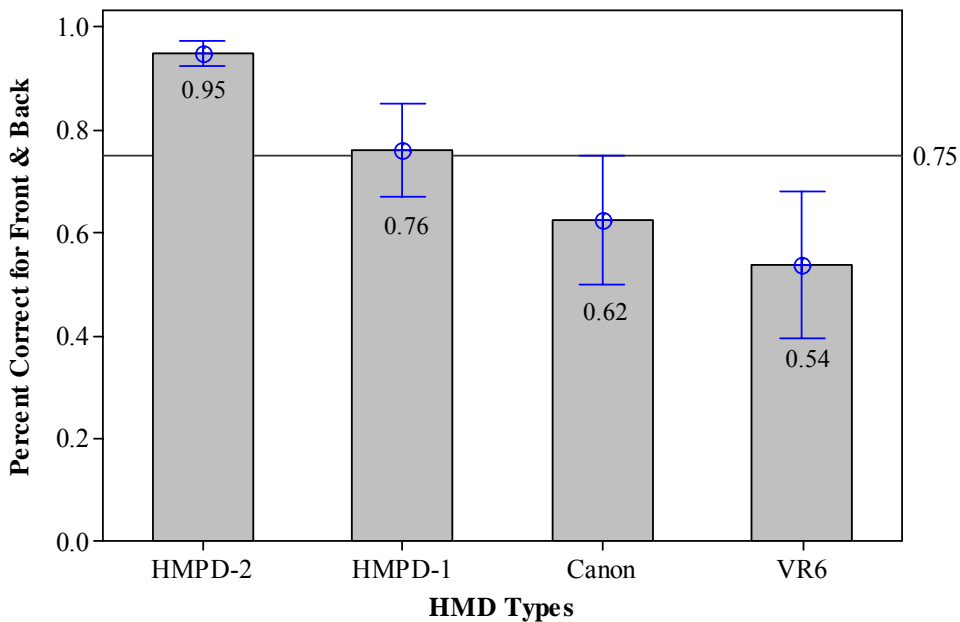
Figure 80 Percent correct front and back reported for all HMDs tested

The calibration results presented in Chapter 5 strongly advocate for improved collaboration between optical, optomechanical, and perception scientists in assembling the displays. Specifically, our experience with user-centered studies lead to the need for the optomechanical design assembly team to measure the FBO of the microdisplay before assembly, so that the center of the viewport can be computed and mechanically localized, centered on the optical axis. Measuring these values cannot be an end user assignment because it will lead to an incorrect optomechnical assembly.

Additionally, we showed how the tilting of the lenses after an alignment not done to specification caused eye strain and inability to fuse the images. We corrected the lens tilt using a customized procedure that only guaranteed that the lenses were facing in the same direction as one another. We may achieve better stereoacuity with the HMPD-2 display with a built in optomechanical adjustment that would allow more precise calibration. Not performing a calibration procedure for the HMD at all, as is commonly done, would introduce errors that although are attributable to the HMD are not part of the HMD design specification. Calibration and proper alignment of the HMD are necessary for ensuring the safety of the persons with TBI and improving the reliability of the outcomes measured within a VE based rehabilitation setting.

Finally, we presented results from a pilot study introducing a new cognitive rehabilitation VE system called the Human Experience Modeler (HEM). This study demonstrated the capability of the integrated HEM technology in delivering a contextualized VE consisting of high level computer graphics overlaid onto a real world fabrication of the participant's kitchen. The goal of this study was to determine the feasibility of constructing such personalized environments and to conduct a pilot study with one participant who presented with anterograde

amnesia (new memories are not transferred to long-term memory) to demonstrate the efficacy of such a system.

A depth camera and a 3D laser scanner captured the volumetric data necessary to reconstruct the physical space of the kitchen within the modeling software 3D Studio Max. Subtleties of the kitchen such as proper reflections off objects were captured with digital photography. For example, when the observer moved to the sink area as pictured in Figure 75a, he or she could look out the window and see the bougainvillea flowers located outside. Because this was our first effort in constructing such a VE, we used this pilot to evaluate the methodology of capturing the real space and converting it into a mixed reality paradigm.

The capture of the individual's home was completed within four hours; however, accurate modeling of the environment was a week long process. Although preliminary, the upfront investment of the reconstruction process may not be cost effective to support individualized training that is less than a week's time in duration. To support more long-term rehabilitation protocols, the physical construction of the environment must be more flexible to accommodate multiple users and their personalized living spaces.

The results of the pilot study suggest that the participant was constructing an internalized script of the breakfast item retrieval sequence and transferring that script to the home environment. Thus, the work potentially shows that persons thought to be unable to store new memories are capable of learning in a contextualized VE. For a patient with anterograde amnesia to exhibit transfer of training to the home is an exciting neuropsychological event and suggests that the MR kitchen paradigm provides an appropriate experimental setting for further evaluation of this result. More research is also needed to understand the level of contextualization necessary to support such learning. The HEMs ability to integrate multimodal inputs will be critical in

determining contextualization levels as well as expanding the project to include more complex personal spaces (i.e., one's favorite fishing hole).

10.2 Looking Forward

This work approaches the question of VE Design for cognitive rehabilitation from a multidimensional point of view. First, we must ensure that the base VE is correctly perceived by the human observer. We provide an HMD assessment to quantify the strengths and weaknesses of the VE system. These constraints allow us to understand limitations in our experimental analysis. However, the two assessments presented, rVA and stereoacuity, only evaluate conditions where the environment is stationary. An additional module to the HMD assessment should evaluate resolvability of elements of a scene in motion.

Dynamic visual acuity (DVA) is the observer's ability to resolve the details of an object when there is relative motion between a fixated target and the observer. DVA has been used as a predictor for performance for tasks where both the objects of interest and the observer are in motion, such as when simply walking or when playing a sport. Given a world built of pixels, understanding how FOV, pixel size, etc. affect DVA of the user may provide more information about the interaction between computer graphics techniques and user interaction. One of the problems in constructing a DVA assessment for VE is the lack of standardization of real world assessments (Banks, Moore, Liu, & Wu, 2004). Thus, elucidating these methodologies is a next step to improving the current HMD assessment battery.

Adding a real versus virtual condition to the experiment performed in chapter 8 is underway. Given that participants can extend space into and off of the plane of the retroreflective material, we would like to know how far off the retroreflective material an object can be

presented and still be accurately perceived in depth. This work will allow us to understand better the limitations of having retroreflective material in the environment. These results can then be compared with the newest mobile HMD which integrates the retroreflective material inside the HMD.

Finally, we plan to integrate the HMD assessment with the HEM testbed to provide sensory baselines for different populations of VE users while performing tasks wearing different types of HMDs. By benchmarking the HEM technology, we can extend the MR rehabilitation paradigm to include new imaging technologies such as functional Near Infrared Brain Imaging (fNIR). Unlike fMRI (functional magnetic resonance imaging), which requires a stationary person to be exposed to a high force field magnet, the fNIR is portable and allows the participant's brain to be imaged while performing a task. The imaging capability of the fNIR comes from an infrared light in the range of 600 to 900 nm that penetrate the skull (Son & Yaziei, 2005). This non-invasive brain imaging technique may provide needed brain functioning data to understand how the HEM supports cognitive rehabilitation through cortical neuronal plasticity. The end result of such a methodological and cohesive plan of study, which enlists multiple field experts, is to expedite a cost effective, reliable, and valid rehabilitation solution to the TBI community.

APPENDIX A: IRB Forms

Informed Consent Form

Please read this consent document carefully before you decide to participate in this study.

Project title: "Analysis and Development of a Full Immersion Projective Head-mounted Display for Dismounted Infantry Training"

Purpose of the research study: The purpose of this study is to measure visual performance in head mounted displays. We shall measure visual acuity as well as depth perception in head mounted displays using standard practice procedures such as the Landolt C visual acuity test.

What you will be asked to do in the study: You will be asked to look into a pupillometer and to adjust a button until you see clearly a circle of light. The pupillometer is used to measure the distance between your two eyes. We shall then measure your visual acuity on a standard monitor. You will be asked to look at a letter C displayed in the center of the display. You will be asked to tell us whether the C is up, down, side right or side left using a 4 button device. Following your answer you will be presented another C and so on for up to 30 trials. You will then be asked to wear a head-mounted display and we shall measure your visual acuity again using the same display of a C. For depth experiments, we shall first measure your ability to discriminate depth on a standard 2 rod device, where you will be asked to pull a string until the two rods are perceived at the same depth. In the head mounted display, you will then use a two-button device to similarly place the two objects (rods, cubes, pyramids) at the same perceived depth. 30 trials will be required.

Time required: One (1) hour maximum per session. A complete experiment may require several hours to be distributed over several weeks (e.g. 10 weeks for 30 hours total).

Risks: You may experience eye fatigue. Because of the see-through nature of the head mounted display. No other effects are expected and eye fatigue is also rare. There will be rest periods during the experiment to further lessen eye strain. If eye fatigue is experienced, you may stop the experiment immediately.

Benefits / Compensation: There is no compensation for participation in this experiment.

Confidentiality: Your identity will be kept confidential. Your information will be assigned a code number. The list connecting your name to this number will be kept in a locked file in my faculty supervisor's office. When the study is completed and the data have been analyzed, the list will be destroyed. Your name will not be used in any report.

Voluntary participation: Your participation in this study is voluntary. There is no penalty for not participating.

Right to withdraw from the study: You have the right to withdraw from the study at any time without consequence.

Whom to contact if you have questions about the study: Cali Fidopiastis, Graduate Student, School of Optics, Building 53, Orlando, FL 32816; (407) 823-6854. Dr. Jannick Rolland, Faculty Supervisor, School of Optics. The telephone number is: (407) 823-6870.

Whom to contact about your rights in the study: UCFIRB Office, University of Central Florida Office of Research, Orlando Tech Center, 12443 Research Parkway, Suite 207, Orlando, FL 32826. The phone number is (407) 823-2901.

               I have read the procedure described above.

               I voluntarily agree to participate in the procedure.

_____ / _____

Participant   Date

               I would like to receive a copy of the final "interview" manuscript submitted to the instructor.

               I would not like to receive a copy of the final "interview" manuscript submitted to the instructor.

_____ / _____

Principle Investigator   Date

Informed Consent Form

Please read this consent document carefully before you decide to participate in this study.

Project title: "Analysis and Development of a Full Immersion Projective Head-mounted Display for Dismounted Infantry Training"

Purpose of the research study: The purpose of this study is to measure visual performance in head-mounted displays. More specifically, we shall measure depth perception in a projective head-mounted display (HMPD) using standard vision tests such as the Howard-Dolman depth perception test.

In this study you will be asked to perform a series of vision tests under normal viewing conditions and while wearing the HMPD. There are four vision tests that you will perform: Dark Focus, Dark Convergence, Visual Acuity, and Depth Perception. The first test is the Howard-Dolman Depth perception test under normal viewing conditions. While performing this test, you will sit in a chair 5.5 meters away from the Howard-Dolman apparatus. Within the apparatus are two rods. One rod will remain stationary, while the other will be moved randomly in the front or the back of the stationary rod. Your task is to align the rods using two strings that you will hold in each hand. You will be given 10 presentations of this task. For the Dark Focus measure, you will be given a stigmatoscope, which looks like a hand held telescope. Place the view piece up to one of your eyes. Within the scope, you will see a small circular light. Your task is to adjust the barrel of the telescope until the light comes into clear focus. This procedure will be repeated 4 times. For the Dark Convergence test, you will stand 6 meters from a flashing red circular light that is located on the wall in front of you at eye height. Over your preferred eye, you will hold a horizontal grating. You will see the red light as a dot in your uncovered eye and as a straight line in your lens covered eye. You will be asked to walk toward the light until the dot and line are over one another. We will measure the distance that you are away from the light at this point.

Next, you will perform a visual acuity test in the virtual environment. However, before being fitted with the HMPD, you will be asked to look into a pupillometer. The experimenter will measure the distance between your two eyes, interpupillary distance (IPD). You will then be fitted with the HMPD. You will be asked and to look at a string of 11 letter C's displayed in the center of the display. You will be asked to tell us whether the gap in the C is facing up, down, side right or side left. Following your answers you will be presented another string of C's of a different size. This procedure will repeat for 4 strings. For the virtual depth experiments, you will perform a similar test as the Howard-Dolman depth perception test, except a scroll device will be used to place the two objects (rods, cubes, pyramids) at the same perceived depth. There will be 64 presentations of the stimulus presented in a single trial. You will perform two trials. At the end of each trial, we will remove the HMPD and you will perform the Dark Focus, Dark Convergence, and Howard-Dolman tasks again. At the conclusion of the experiment, you will be given a simulator sickness questionnaire. We will also setup another time for you to return at another time to perform the experiment again.

Time required: One (1) hour maximum per session. A complete experiment requires two sessions separated by at least 3 hours.

Risks: You may experience eye fatigue. Because of the see-through nature of the head mounted display. No other effects are expected and eye fatigue is also rare. There will be rest periods during the experiment to further lessen eye strain. If eye fatigue is experienced, you may stop the experiment immediately.

Benefits / Compensation: Extra credit will be given once you have completed both experimental days. Otherwise, there is no compensation for participation in this experiment.

Confidentiality: Your identity will be kept confidential. Your information will be assigned a code number. The list connecting your name to this number will be kept in a locked file in my faculty supervisor's office. When the study is completed and the data have been analyzed, the list will be destroyed. Your name will not be used in any report.

Voluntary participation: Your participation in this study is voluntary. There is no penalty for not participating.

Right to withdraw from the study: You have the right to withdraw from the study at any time without consequence.

Whom to contact if you have questions about the study: Cali Fidopiastis, Graduate Student, School of Optics, Building 53, Orlando, FL 32816; (407) 823-6854. Dr. Jannick Rolland, Faculty Supervisor, College of Optics and Photonics. The telephone number is: (407) 823-6870.

Whom to contact about your rights in the study: UCFIRB Office, University of Central Florida Office of Research, Orlando Tech Center, 12443 Research Parkway, Suite 207, Orlando, FL 32826. The phone number is (407) 823-2901.

_____I have read the procedure described above.

_____I voluntarily agree to participate in the procedure.

_____/_____

Participant   Date

_____I would like to receive a copy of the final "interview" manuscript submitted to the instructor.

_____I would not like to receive a copy of the final "interview" manuscript submitted to the instructor.

_____/_____

Principle Investigator   Date

Informed Consent Form

Please read this consent document carefully before you decide to participate in this study.

Project title: "Transfer of Training Outcomes When Using a Virtual Reality Trainer For Brain Injury Rehabilitation"

Purpose of the research study: The purpose of this study is to demonstrate feasibility of the Human Experience Modeler, a virtual environment for cognitive rehabilitation currently under development at the Media Convergence Laboratory here at UCF. In this pilot study, we will study different methodologies for constructing a 3D graphical display of your kitchen and the effects of an After Action Review on learning.

What you will be asked to do in the study: On day one of the study, you will meet with Dr. Jeff Bedwell, a Clinical Psychologist at UCF, who will perform a series of neuropsychological evaluations to establish more information about your cognitive deficits. On day two of the study, you will familiarize yourself with the HEM technology. For the familiarization process, you will enter the staging area where the HEM (multi-modal display, along with audio and sensory technology) is setup. You will be fitted with either the video or optical see through HMD. Prior to fitting, the interpupillary distance (IPD) of your eyes will be measured using a pupillometer. The HMD eye pieces will be adjusted to that IPD to minimize the potential for eye strain. Within the HEM, you will see a generic kitchen setup. You will be asked to perform different tasks such as locating appliances (e.g., the stove or the refrigerator). Each task will utilize a different a visual cuing mode. Once each task is completed, you will verbally report your comfort level on interacting within the HEM. On your next visit, the training sessions will begin where you will view a replica of your own kitchen. In your kitchen, you will perform the task of making your favorite breakfast, which will include organizing the ingredients, cooking the food, and serving it to your family.

You will receive HEM therapy two times a week for four weeks. During each training session, you will be given a task list for making your favorite breakfast, eggs over easy. Following each training session, your therapist, Dr. Whiteside, will review your performance with the After Action Review.

In order to assess transfer to the home environment, following the completion of 4 weeks of training, we will conduct a brief in home observation to videotape execution of the breakfast making task. Also, you and your family members will be asked to keep a diary of your ability to execute the task during the subsequent four weeks. Diaries will then be collected and reviewed for the post testing assessments.

Time required: Each session will be 35 minutes maximum per session. The training will be conducted twice per week for 4 weeks. We will also do a 4 week post training follow up.

Risks: One drawback of virtual environment research has been simulator sickness. Simulator sickness has some similar bodily sensations as motion sickness. For example, you may experience nausea or dizziness. Current research has shown that simulator sickness is related to the amount of time that you spend performing a task in the virtual environment. To prevent you from feeling simulator sickness we will monitor your sickness level using a Simulator Sickness Questionnaire. In addition, you will be required to take 5-minute breaks throughout the training session. Further, your HEM training sessions will be limited to a maximum of 35 minutes. The

Simulator Sickness Questionnaire will be given to you following the HEM training to continue monitoring for any adverse effects.

You may also experience eye fatigue. Because of the see-through nature of the head mounted display. No other effects are expected and eye fatigue is also rare. There will be rest periods during the experiment to further lessen eye strain. If you experience eye fatigue, motion sickness, or discomfort, you may stop the experiment immediately.

In addition, we recognize that the simulation setup is identical to your home environment; however, this similarity does not mean that all tasks performed in the study can be immediately performed at home competently and without supervision from a family member. As with any learning, transferring training to the actual task takes time and patience. There is a risk of bodily injury and property damage if you try to perform the kitchen tasks presented in your training without first demonstrating that you can complete these tasks in your own kitchen independently using safe procedures.

Benefits / Compensation: There is no compensation for participation in this experiment.

Confidentiality: Your identity will be kept confidential. Your information will be assigned a code number. The list connecting your name to this number will be kept in a locked file in the supervisor's office. When the study is completed and the data have been analyzed, the list will be destroyed. Your name will not be used in any report. Further any video taped material will also be erased upon completion of the entire study.

Voluntary participation: Your participation in this study is voluntary. There is no penalty for not participating.

Right to withdraw from the study: You have the right to withdraw from the study at any time without consequence.

If you believe you have been injured during participation in this research project, you may file a claim with UCF Environmental Health & Safety, Risk and Insurance Office, P.O. Box 163500, Orlando, FL 32816-3500 (407) 823-6300. The University of Central Florida is an agency of the State of Florida for purposes of sovereign immunity and the university's and the state's liability for personal injury or property damage is extremely limited under Florida law. Accordingly, the university's and the state's ability to compensate you for any personal injury or property damage suffered during this research project is very limited.

Whom to contact if you have questions about the study: Cali Fidopiastis, Graduate Student, School of Optics, Building 53, Orlando, FL 32816; (407) 823-6854. Dr. Janet Whiteside, Communicative Disorders Clinic. The telephone number is: (407) 249-4770.

Whom to contact about your rights in the study: UCFIRB Office, University of Central Florida Office of Research, Orlando Tech Center, 12443 Research Parkway, Suite 207, Orlando, FL 32826. The phone number is (407) 823-2901.

_____I have read the procedure described above.

_____I voluntarily agree to participate in the procedure.

_____/_____

Participant   Date

_____ I would like to receive a copy of the final "interview" manuscript submitted to the instructor.

_____ I would not like to receive a copy of the final "interview" manuscript submitted to the instructor.

_____ / _____

Principle Investigator   Date

APPENDIX B: Questionnaires

Simulator Sickness Questionnaire (SSQ)

Developed by Robert S. Kennedy & colleagues under various projects. For additional information contact:
Robert S. Kennedy, RSK Assessments, Inc., 1040 Woodcock Road, Suite 227, Orlando, FL 32803 (407) 894-5090

Subject Number: _____   **Date:** _____

PRE-EXPOSURE BACKGROUND INFORMATION

1.   How long has it been since your last exposure in a simulator? _____ days
      How long has it been since your last flight in an aircraft? _____ days
      How long has it been since your last voyage at sea? _____ days
      How long has it been since your last exposure in a virtual environment? _____ days
2.   What other experience have you had recently in a device with unusual motion?

PRE-EXPOSURE PHYSIOLOGICAL STATUS INFORMATION

3.   Are you in your usual state of fitness? (Circle one)   YES   NO
      If not, please indicate the reason:
4.   Have you been ill in the past week? (Circle one)   YES   NO
      If "Yes", please indicate:
      a)   The nature of the illness (flu, cold, etc.):
      b)   Severity of the illness: Very _____ Very
                                       Mild   Severe
      c)   Length of illness: _____ Hours / Days
      d)   Major symptoms:
      e)   Are you fully recovered?   YES   NO
5.   How much alcohol have you consumed during the past 24 hours?
      _____ 12 oz. cans/bottles of beer   _____ ounces wine   _____ ounces hard liquor
6.   Please indicate <u>all</u> medication you have used in the past 24 hours. If none, check the
      first line:
      a) NONE
      b) Sedatives or tranquilizers
      c) Aspirin, Tylenol, other analgesics
      d) Anti-histamines
      e) Decongestants
      f) Other (specify):
7.   a) How many hours of sleep did you get last night? _____ hours
      b) Was this amount sufficient? (Circle one)   YES   NO
8.   Please list any other comments regarding your present physical state which
      might affect your performance on our test battery.

Baseline (Pre) Exposure Symptom Checklist

Instructions: Please fill this out BEFORE you go into the virtual environment. Circle how much each symptom below is affecting you <u>right now</u>.

#	Symptom	Severity			
1.	General discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Boredom	None	Slight	Moderate	Severe
4.	Drowsiness	None	Slight	Moderate	Severe
5.	Headache	None	Slight	Moderate	Severe
6.	Eye strain	None	Slight	Moderate	Severe
7.	Difficulty focusing	None	Slight	Moderate	Severe
8a.	Salivation increased	None	Slight	Moderate	Severe
8b.	Salivation decreased	None	Slight	Moderate	Severe
9.	Sweating	None	Slight	Moderate	Severe
10.	Nausea	None	Slight	Moderate	Severe
11.	Difficulty concentrating	None	Slight	Moderate	Severe
12.	Mental depression	None	Slight	Moderate	Severe
13.	"Fullness of the head"	None	Slight	Moderate	Severe
14.	Blurred Vision	None	Slight	Moderate	Severe
15a.	Dizziness with eyes open	None	Slight	Moderate	Severe
15b.	Dizziness with eyes closed	None	Slight	Moderate	Severe
16.	*Vertigo	None	Slight	Moderate	Severe
17.	**Visual flashbacks	None	Slight	Moderate	Severe
18.	Faintness	None	Slight	Moderate	Severe
19.	Aware of breathing	None	Slight	Moderate	Severe
20.	***Stomach awareness	None	Slight	Moderate	Severe
21.	Loss of appetite	None	Slight	Moderate	Severe
22.	Increased appetite	None	Slight	Moderate	Severe
23.	Desire to move bowels	None	Slight	Moderate	Severe
24.	Confusion	None	Slight	Moderate	Severe
25.	Burping	None	Slight	Moderate	Severe
26.	Vomiting	None	Slight	Moderate	Severe
27.	Other				

*   Vertigo is experienced as loss of orientation with respect to vertical upright.
**   Visual illusion of movement or false sensations of movement, when <u>not</u> in the simulator, car, or aircraft.
*** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

STOP HERE! The test director will tell you when to continue.

POST 00 Minutes Exposure Symptom Checklist

Instructions: Circle how much each symptom below is affecting you <u>right now</u>.

#	Symptom	Severity			
1.	General discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Boredom	None	Slight	Moderate	Severe
4.	Drowsiness	None	Slight	Moderate	Severe
5.	Headache	None	Slight	Moderate	Severe
6.	Eye strain	None	Slight	Moderate	Severe
7.	Difficulty focusing	None	Slight	Moderate	Severe
8a.	Salivation increased	None	Slight	Moderate	Severe
8b.	Salivation decreased	None	Slight	Moderate	Severe
9.	Sweating	None	Slight	Moderate	Severe
10.	Nausea	None	Slight	Moderate	Severe
11.	Difficulty concentrating	None	Slight	Moderate	Severe
12.	Mental depression	None	Slight	Moderate	Severe
13.	"Fullness of the head"	None	Slight	Moderate	Severe
14.	Blurred Vision	None	Slight	Moderate	Severe
15a.	Dizziness with eyes open	None	Slight	Moderate	Severe
15b.	Dizziness with eyes closed	None	Slight	Moderate	Severe
16.	*Vertigo	None	Slight	Moderate	Severe
17.	**Visual flashbacks	None	Slight	Moderate	Severe
18.	Faintness	None	Slight	Moderate	Severe
19.	Aware of breathing	None	Slight	Moderate	Severe
20.	***Stomach awareness	None	Slight	Moderate	Severe
21.	Loss of appetite	None	Slight	Moderate	Severe
22.	Increased appetite	None	Slight	Moderate	Severe
23.	Desire to move bowels	None	Slight	Moderate	Severe
24.	Confusion	None	Slight	Moderate	Severe
25.	Burping	None	Slight	Moderate	Severe
26.	Vomiting	None	Slight	Moderate	Severe
27.	Other				

*   Vertigo is experienced as loss of orientation with respect to vertical upright.
**   Visual illusion of movement or false sensations of movement, when <u>not</u> in the simulator, car or aircraft.
*** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

POST-EXPOSURE INFORMATION

1.   While in the virtual environment, did you get the feeling of motion (i.e., did you experience a compelling sensation of self motion as though you were actually moving)? *(Circle one)*

YES   NO   SOMEWHAT

2.   On a scale of 1 (POOR) to 10 (EXCELLENT) rate your performance in the virtual environment:

3.   a. Did any unusual events occur during your exposure? *(Circle one)*   YES   NO
       b. If YES, please describe

Participant Information Questionnaire

Please circle or fill in the appropriate answer

1) Gender: Female   Male

2) Age (in years): _____

3) How many hours of sleep did you get last night? _____

                        Do you feel rested?   YES   NO

4) Indicate all medications/substances you have used in the past 24 hours:

        CIRCLE ALL THAT APPLY

        0 - None
        1 - Sedatives or tranquilizers
        2 - Aspirin, Tylenol, other analgesics
        3 - Antihistamines
        4 - Decongestants
        5 - Other Please list _____

5) Have you had caffeine in the last 2 hours?   YES   NO
       If yes, about how many ounces? _____

6) Have you ever experienced motion sickness or car sickness?   YES   NO

7) How susceptible are you to motion sickness or car sickness?
     0   1   2   3   4   5
  Not at all   Mildly   Average   Highly

8) How many hours per day do you use a computer? _____

9) How many hours per day do you watch television? _____

10) How many hours per week do you play video games? _____
11) How many hours per week do you spend reading? _____

12) How much experience do you have with virtual environments?
     0   1   2   3   4   5
  Not at all   Very Little   Average   A Lot

13) What is your occupation? _____

14) How often during an average work day do you feel eye strain?

0	1	2	3	4	5
Not at all	Very Little		Average		A Lot

15) During an average work day, do you feel that you focus on near objects (about 2 meters away) more than objects that are far away (6 meters or more)?

1	2	3	4	5
Strongly Disagree		Agree		Strongly Agree

16) How much of your day (in percent) is spent viewing

Near objects (about 2 meters away)? _____

Far Objects (6 meters or more away)? _____

Total should equal 100%

17) Do you have normal or corrected to normal 20/20 vision?   YES   NO

If yes, do you wear glasses or contact lenses? _____

Are you near or far sighted? _____

18) Have you ever had eye surgery? YES   NO

If yes, please explain (e.g., LASIK surgery) _____

19) Do you have a history of epilepsy or seizures?   YES   NO

LIST OF REFERENCES

Angel, E. (2003). Interactive computer graphics: A top-down approach to OpenGL (3rd ed.). Boston, MA: Addison-Wesley.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In W. K. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 1, pp. 89-195). New York, NY: Academic Press.

Bach, M. (1996). The Freiburg Visual Acuity test: Automatic measurement of visual acuity. *Optometry and vision science, 73*(1), 49-53.

Bach, M., Schmitt, C., Kromeier, M. & Kommerell, G. (2001). The Freiburg Stereoacuity Test: Automatic measurement of stereo threshold. *Graefe's Archive for Clinical and Experimental Ophthalmology, 239*(8), 562-566.

Backus, B. T. (2000). Stereoscopic vision: What's the first step? *Current Biology, 10*, R701-R703.

Backus B. T., Fleet D. J., Parker A. J., and Heeger D. J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology, 86*, 2054–2068.

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *The Journal of Neuroscience, 24*(9), 2077-2089.

Banks, P. M., Moore, L. A., Liu, C., & Wu, B. (2004). Dynamic visual acuity: A review. *South African Optometrist, 6*(2), 58-64.

Baumann, S.; Neff, C., Fetzick, S., Stangl, G., Basler, L., Vereneck, R., & Schneider, W. (2003). A Virtual Reality System for Neurobehavioral and Functional MRI Studies. *CyberPsychology & Behavior, 6*(3), 259-266.

Biocca, F. A. & Rolland, J. P. (1998). Virtual eyes can rearrange your body: Adaptation to visual displacement in see-through, head-mounted displays. *Presence, 7*(3), 262-277.

Boake, C. (1991). History of cognitive rehabilitation following head injury. In J. S. Kreutzer & P. H. Wehman (Eds.) *Cognitive rehabilitation for persons with traumatic brain injury: A functional approach (pp. 3-12)*. Baltimore, MD: Paul H. Brookes Publishing Co.

Bowman, D. A., Gabbard, J. L. & Hix, D. (2002). A survey of usability evaluation in virtual environments: Classification and comparison of methods. *Presence, 11*(4), 404-424.

Bowman, D. A., Johnson, D. B. & Hodges, L. F. (2001). Testbed evaluation of virtual environment interaction techniques. *Presence: Teleoperators and Virtual Environments, 10*(1):75-95.

Brooks, B. M., McNeil, J. E. Rose, F. D., Greenwood, R. J., Attree, E.A., Leadbetter, A.G. (1999). Route Learning in a Case of Amnesia: A Preliminary Investigation into the Efficacy of Training in a Virtual Environment. *Neuropsychological Rehabilitation, 9*, 63-76.

Brown, D. J., Standen, P. J., & Cobb, S. V., (1998). Virtual Environments Special Needs and Evaluative Methods. In G. Riva, B. K.Wiederhold, & E. Molinari (Eds.). *Virtual Environments in Clinical Psychology and Neuroscience.* Amsterdam, Netherlands: Ios Press.

Bullimore, M. A., Howarth, P. A., & Fulton, J. (1995). Assessment of visual performance. In J. R. Wilson & E. N. Corlett (Eds.). *Evaluation of human work: a practical ergonomics methodology*. (2cd ed., pp. 804-839). Nottingham, U.K.: Taylor & Francis.

Burdea, G. C. & Coiffet, P. (2004). *Virtual Reality Technology* (2^{cd} ed., p. 3). New Jersey: John Wiley & Sons, Inc.

Cakmakci, O. & Rolland, J. P. (in press). Head-worn displays: A review. *IEEE/OSA Journal of Display Technology.*

Capdevielle, V. Martins, R. & Rolland, J. P. (2003). Optical properties of retroreflective material. Retrieved on September 15, 2004, from http://www.creol.ucf.edu/Research/ REU/StudentProjectDetail.asp?PeopleID=770&Yr=2003

Carney, N., Chestnut, R. M, Maynard, H. Mann, N. C., Paterson, P.,& Helfand, M. (1999). Effect of cognitive rehabilitation on outcomes for persons with traumatic brain injury: a systematic review. *Journal of Head Trauma Rehabilitation, 14*, 277-307.

Centers for Disease Control and Prevention, CDC (2001). *Injury fact book 2001-2002*. Atlanta ,GA: Centers for Disease Control and Prevention.

Chase, W. G., Ericsson, K. A. (1981). Skilled memory, In J. R. Anderson (Ed.) *Cognitive skills and their acquisition* (pp.141-189). Hillsdale, NJ: Erlbaum.

Chen, S., Thomas, J., Glueckauf, R., Bracy, O. (1997). The effectiveness of computer-assisted cognitive rehabilitation for persons with traumatic brain injury. *Brain Injury, 11,* 197–209.

Christensen, J. (1993). The nature of systems development. In R. W. Pew & P. Green (Eds.), *Human Factors Engineering Short Course Notes*, (34th ed., Vol. 1). Ann

Arbor, MI: The University of Michigan, Chrysler Center for Continuing Engineering Education.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684.

Cruz-Niera, C, Sandin, D. J, & DeFanti, T. A. (1993). Surround screen projection-based virtual reality: The design and implementation of the CAVE. In J.T. Kajiya (Ed.), Proceedings of SIGGRAPH 93, Computer Graphics Proceedings, Annual Conference Series (pp.135-142). New York, NY: ACM.

Cumming, B. G. & Parker, A. J. (1999) Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *The Journal of Neuroscience, 19*(13), 5602–5618.

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.

Cutting, J.E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, & Computers, 29*(1), 27-36.

Davis, E.T. (1997). Visual requirements in HMDs: What can we see and what do we need to see? In J. Melzer and K. Moffitt (Eds.) *Head Mounted Displays: Designing for the User* (pp. 207-284). New York: McGraw-Hill.

Davis, L., Rolland J. P., Hamza-lup, F., & Ha, Y. (2003, March/April). Enabling a continuum of virtual environment experiences. *IEEE Computer Graphics and Applications*, 2-4.

DeAngelis G. C. (2000). Seeing in three dimensions: The neurophysiology of stereopsis. *Trends in Cognitive Sciences, 4*(3), 80-90.

Duh, H. B., Lin, J. J., Kenyon, R. V., Parker, D. E. & Furness, T. A. (2002). Effects of characteristics of image quality in an immersive environment. *Presence: Teleoperators and Virtual Environment, 11*(3), 324-332.

Eggelston, R. G. (1997). User-centered design in the trenches: Head-mounted display system design and user performance (Chp. 2). In J.E. Melzer & K. Moffit (Eds.) *Head Mounted Displays: Designing for the user* (pp. 17-54). New York, NY: McGrawhill.

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

Ellis, S. R. (1993). Seeing. In S. R. Ellis, M. K. Kaiser, & A. J. Grunwald (Eds.) *Pictorial communication in virtual and real environments* (2^{cd} ed., pp. 419-424). Washington, DC: Taylor & Francis.

Engum, E. S., Cron, L., Hulse, K. C., Pendergrass, T. M., Lambert, E. (1988). Cognitive behavioral driver's inventory. *Cognitive Rehabilitation, 6*, 5, 34–50.

Fencott, C. (1999). Towards a design methodology for virtual environments. Poster presented at the User Centered Design and Implementation Workshop September 30, York England. http://www.cs.york.ac.uk/hci/kings_manor_workshops/UCDIVE/.

Fidopiastis, C.M., Meyer C, Fuhrman K, Rolland J.P. (2003). Quantitative assessment of visual acuity in projection head-mounted displays. In: Rash CE, Colin ER (Eds.) Proceedings of the SPIE Aerosense: Helmet- and Head-Mounted Displays VIII: Technologies and Applications 5079, 399-406.

Fidopiastis, C.M., Furhman, C. Meyer, C. & Rolland, J.P. (2005). Methodology for the

    Iterative Evaluation of Prototype Head-mounted Displays in Virtual Environments:

    Visual Acuity Metrics, *Presence: Teleoperators and Virtual Environments, 14*(5), 550-

    562.

Fidopiastis, C. M., Stapleton, C. B., Whiteside, J. D., Hughes, C. E. Fiore, S. M. Martin, G. M.

    Rolland, J. P., Smith, E. M. (September, 2005), Human Experience Modeler: Context-

    Driven Cognitive Retraining to Facilitate Transfer of Learning, *Proceedings of the 4th*

    International Workshop on Virtual Rehabilitation (IWVR), September 19-21, Catalina

    Island, CA.

Fidopiastis, C. M., Stapleton, C. B., Whiteside, J. D., Hughes, C. E. Fiore, S. M. Martin, G. M.

    Rolland, J. P., Smith, E. M. (April, 2006). Human experience modeler: Context-driven

    cognitive retraining to facilitate transfer of learning. *CyberPsychology & Behavior,*

    9(2), 183-187.

Finney, D. J. (1980). *Probit analysis* (3rd Ed.). Cambridge: Cambridge University Press.

Fiorentini, A., Baumgartner, G., Magnusson, S., Schiller, P. H. and Thomas, J. P. (1990). The

    perception of brightness and darkness. In L. Spillman and J. S. Werner (Eds.), *Visual*

    perception: The Neurophysiological Foundations (pp. 129-161). San Diego, CA:

    Academic Press, Inc.

Foley, J. D., Van Dam, A., Feiner, S. K. & Hughes, J. F. (1995). Computer graphics: Principles

    and practice (2^{cd} ed.). Reading, MA: Addison and Wesley Publishing Company.

Fowlkes, J. E., Kennedy, R. S., Hettinger, L. J., & Harm, D. L. (1993). Changes in the dark focus

    of accommodation associated with simulator sickness. *Aviation, Space, and*

    Environmental Medicine, 64(7), 612-618.

Furhman, C., Davis, L., Fidopiastis, C. M., & Rolland, J. P. (2006). ODALab HMD Visual Test Battery Manual. ODALab University of Central Florida.

Gabbard, J. L., Hix, D., Swan, J. E. (1999). User-centered design and evaluation of virtual environments. *IEEE Computer Graphics and Applications, 19*(6), 51-59.

Gilchrist, A. (1994). Absolute versus relative theories of lightness perception. In A. Gilchrist (ed.) *Lightness, brightness, and transparency* (pp. 1-33). Hillsdale, NJ: Erlbaum.

Goodale, M.A. & Haffenden, A. (1998). When vision is not sight. In L.R. Harris & M. Jenkins (Eds.) *Vision and action* (pp. 270-294). Cambridge, UK: Cambridge University Press.

Goss, D. A. & West, R. W. (2001). *Introduction to the optics of the eye*. Boston, MA: Butterworth-Heinemann.

Gourlay, D., Lun, K. C., Lee, Y. N., & Tay, J. (2000). Virtual reality for relearning daily living skills. *International Journal of Medical Informatics, 20*, 255-261.

Griffiths, G., Sharples, S. & Wilson, J. R. (2006). Performance if new participants in virtual environments: The Nottingham tool for assessment of interaction in virtual environments (NAÏVE). *International Journal of Human-Computer Studies, 64*, 240-250.

Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing-and test-field luminances. *Journal of Experimental Psychology, 50*(2), 89-96.

Hershenson, M. (2000). *Visual space perception: A primer* (p. 2). Cambridge: MIT Press.

Hix, D. & Gabbard, J. L. (2002). Usability engineering in virtual environments. In K. Stanney (Ed.) *Handbook of virtual environments: Design, implementation, and applications* (pp. 681-699). Mahwah, NJ: Lawrence Erlbaum.

Hoffman, H. G., Patterson, D. R., Carrougher, G. J., & Sharar, S. (2001). The effectiveness of virtual reality based pain control with multiple treatments. *Clinical Journal of Pain, 17*(3), 229- 235.

Hölscher, C., Schnee, A., Dahmen, H., Setia, L. & Mallot, H. A. (2005). Rats are able to navigate in virtual environments. *The Journal of Experimental Biology, 208*, 561-569.

Howard, H. J. (1919). A test for the judgment of distance. *American Journal of Ophthalmology, 2*, 656-675.

Howard, J. P. & Rogers, B. J. (1995). *Binocular Vision and Stereopsis*, Oxford, England: Clarendon Press.

Howarth, P. A. (1995). Assessment of the visual environment. In: Wilson, J. R. and E. N. Corlett (Eds.) *Evaluation of Human Work* (2nd ed., pp. 445-482). Bristol, PA: Taylor & Francis, Inc.

Howarth, P. A. & Costello, P. J. (1997). The development of a visual test battery for virtual reality users. In S. A. Robertson (Ed.) *Contemporary Ergonomics* (pp. 109-116). London, England: Taylor and Francis, Inc.

Howarth, P. A. (1999). Oculomotor changes within virtual environments. *Applied Ergonomics, 30*, 59-67.

Ha, Y. & Rolland, J. P. (September, 2002). Optical assessment of head-mounted displays in visual space. *Applied Optics, 41*(25), 5282-5289.

Hu, H. H., Gooch, A. A., Thompson, W. B., Smits, B. E., Rieser, J. J., Shirley, P. (2000): Visual cues for imminent object contact in realistic virtual environment. *IEEE Visualization*, 179-185.

Hua, H., Girardot, A., Gao, C., & Rolland, J. P. (2000). Engineering of head-mounted projective displays. *Applied Optics, 39*(22), 3814-3824.

Hua H., Krishnaswamy, P. & Rolland, J. P. (in press). "Video-based eyetracking methods and algorithms in head-mounted displays, *Optics Express*.

Hughes, C. E., Stapleton, C. B., Hughes D. E., & Smith, E. (2005). Mixed Reality in education, entertainment and training: An interdisciplinary approach, *IEEE Computer Graphics and Applications, 26*(6), 24-30.

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

Johansson, B. B. (2003). Environmental influence on recovery after brain lesions- experimental and clinical data. *Journal of Rehabilitation Medicine, 41(S)*, 11-16.

Kaber, D. B., Draper, J. V., & Usher, J. M. (2002). Influence of individual differences and application design for individual and collaborative immersive virtual environments (pp. 379-402). In K. Stanney (Ed.) *Handbook of virtual environments: Design, implementation, and applications*. Mahwah, NJ: Lawrence Erlbaum.

Kaernbach, C. (2001). Slope bias of psychometric functions derived from adaptive data. *Perception & Psychophysics, 63*(8), 1389-1398.

Kalawsky, R. S. (1993). *The science of virtual reality and virtual environments: a technical, scientific and engineering reference on virtual environments*. Wokingham, England: Addison-Wesley.

Kandel, E. R., Schwartz, J. H. & Jessell, T. M (1995). *Essentials of Neural Science and Behavior* (p. 37). Norwalk, CT: Appleton & Lange.

Kaufman, J. E. & Christensen, J. F. (1989) *IES Lighting Ready Reference*. New York, NY: Illuminating Engineering Society of North America.

Kaur, K., Maiden, N., Sutcliffe, A. (1999). Interacting with virtual environments: an evaluation of a model of interaction. *Interacting with Computers. 11*, 403-426.

Kaur, K., Sutcliffe, A., Maiden, N. (1999). A design advice tool presenting usability guidance for virtual environments. Paper presented at the User Centered Design and Implementation Workshop September 30, York England. http://www.cs.york.ac.uk/hci/kings_manor _workshops/UCDIVE/

Keller, K. & Colucci, D. (1998). Perception in HMDs: what is it in head-mounted displays (HMDs) that really make them all so terrible? In R.J. Lewandowski, L. A. Haworth, & H. J. Girolamo (Eds.) Proceedings of SPIE Volume 3362 Helmet and Head-Mounted Displays III, 46-53.

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.C., Kennedy, R. S., & Drexler, J. M. (2006, April). Are motion sickness symptoms normally distributed? Evidence for the Pareto Effect. Paper presented at the symposium of Human Performance at Sea: Influence of Ship Motions on Biomechanics and Fatigue, Panama City, Florida.

Kirsch, N.L, Levine, S.P., Lajiness-O'Neill, R., Schnyder, M. (1992). Computer-assisted interactive task guidance: Facilitating the performance of a simulated vocational task. *Journal of Head Trauma Rehabilitation, 7*(*3*), 13-25.

Kolb, H., Fernandez, E., Nelson, R., & Jones, B.W. (2006). Webvision: The organization of the retina and visual system. Retrieved on April 24, 2004, from http://webvision.med.utah.edu/index.html

Lambert, E. & Engum, E. S. (1992). Construct validity of the Cognitive Behavioral Driver's Inventory: age, diagnosis, and driving ability. *Cognitive Rehabilitation, 10*(3), 32–45.

Lampton, D., Knerr, B., Goldberg, S., Bliss, J., Moshell, J., & Blatt, B. (1994). The Virtual Environment Performance Assessment Battery (VEPAB): Development and evaluation. *Presence: Teleoperators and Virtual Environment, 3*(2), 145-157.

Larson, W. L. (1985). Does the Howard-Dolman really measure stereoacuity? *American Journal of Optometry & Physiological Optics, 62*(11), 763-767.

Lieberman, H. R. & Pentland, A. P. (1982). Computer technology microcomputer-based estimation of psychophysical thresholds: the best pest. *Behavior Research Methods & Instrumentation, 14*(1), 21-25.

Levin, W. (1991). Computer applications in cognitive rehabilitation. In J. S. Kreutzer & P. H. Wehman (Eds.) *Cognitive rehabilitation for persons with traumatic brain injury: A functional approach* (pp. 163-179). Baltimore: Paul H. Brookes Publishing Co.

Liepert, J., Bauder, H., Wolfgang, H.R., Miltner, W.H., Taub, E., & Weiller, C. (2000). Treatment-induced cortical reorganization after stroke in humans. *Stroke, 31*, 1210-1216.

Loomis, J. M. & Knapp, J. M. (2003). Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger and M. W. Haas (Eds.), V*irtual and Adaptive Environments* (pp. 21-46). Mahwah NJ: Erlbaum.

Lynch, W. (2002). Historical review of computer-assisted cognitive retraining. *Journal of Head Trauma Rehabilitation, 17*, 5, 446-457.

Lynch, W. (1982). The use of videogames in cognitive rehabilitation. In L. Trexler (Ed.) *Cognitive Rehabilitation: Conceptualization and Intervention*. New York: Plenum.

Martins, R. & Rolland, J. P. (Apr 2003). Diffraction of Phase conjugate material in a new HMD architecture. Paper presented at SPIE's 17th Annual International Symposium on Aerospace/Defense Sensing Simulation and Controls, Orlando, FL.

May, J. G. and Badcock, D. R. (2002) Vision and Virtual Reality. In K. M. Stanney, (Ed.), *Handbook of Virtual Environments Technology* (pp. 29-63), Lawrence Erlbaum Associates, Inc.

Max, W., MacKenzie, E. J., Rice, D. P. (1991). Head injuries: costs and consequences. *Journal of Head Trauma Rehabilitation*, *6*, 76-91.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York, NY: Cambridge University Press.

Meister, D. (1985). *Behavioral analysis and measurement methods*. New York, NY: Wiley.

Melzer, E., Moffit, K. (1997). *Head Mounted Displays: Designing for the User*. New York, NY: McGraw-Hill.

Milgram, O. & Kishino, F, (1994, December). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems, E77-D*(12):1321-1329.

Milliez, A. J. (2006). Up-conversion in rare-earth doped micro-particles applied to new emissive 2D display. Unpublished doctoral dissertation, University of Central Florida, Orlando, FL.

Moreira da Costa, R.M. & Vidal de Carvalho, L.A. (2000). "Virtual reality in cognitive retraining". *Proceedings of IWALT 2000, the International Workshop on Advanced Learning Technology*, Palmerston North, New Zealand, 4-6 December, Los Alamitos, CA: IEEE Computer Society, 221-224.

Myers, R. L. & Bierig, T. A. (2000). Virtual reality and left hemifield neglect: A technology for assessment and therapy. *CyberPsychology & Behavior, 3*(3), 465-468.

National Academy of Sciences. (1980). Recommended standard procedures for clinical measurement and specification of visual acuity. *Advances in Ophthalmology*, *41*, 103-148.

National Center for Injury Prevention and Control (2001). State *Injury Indicators Report*. Retrieved from the CDC website http://www.cdc.gov on July 15, 2004.

National Research Council (1987). *Work, aging, and vision.* Washington D.C.: National Academy Press. http://darwin.nap.edu/books/POD252/html/2.html.

Nature Editorial (2006, April). More than the money. *Nature, 440*(7086), 845-846.

Nemire, K., & Ellis, S. R. (1993). *Calibration and evaluation of virtual environment displays. IEEE 1993 Symposium on Research Frontiers in Virtual Reality*. San Jose, California: IEEE Computer Society Press, 33-40.

Neri, P. (2005). A stereoscopic look at the visual cortex. *The Journal of Neurophysiology, 93*, 1823-1826.

Newman, M. (1970). Visual acuity. In R. Moses (Ed.) *Adler's Physiology of the Eye: Clinical Applications*, (5th ed., pp. 561-583). St. Louis, MO: The C.V. Mosby Company.

O'Conner, M. & Cermack, L. S. (1987). Rehabilitation of organic memory disorders. In M. J. Meier, A. L. Benton, & L. Diller (Eds.) *Neuropsychological rehabilitation*, New York, NY: Guilford Press.

O'Connor, M. & Hughes, C. E. (2005). Authoring and delivering mixed reality experiences. *Proceedings of 2005 International Conference on Human-Computer Interface Advances in Modeling and Simulation (SIMCHI'05)*, New Orleans, LA, January 23-27, 33-39.

Ogle, K. N. (1962). Visual optics and the optical space sense. In H. Davidson (Ed.) *The Eye* (Vol. 4, pp. 271–324). New York, NY: Academic Press.

Owens, D. A. & Leibowtiz, (1976). Oculomotor adjustments in darkness and the specific distance tendency. *Perception and Psychophysics, 20*, 159-67.

Owen, C. B., Zhou, J., Tang, A. & Xiao, F. (2004). Display-relative calibration for optical see-through head-mounted displays. Third IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2004, 70-78.

Owens, D. A. (1984). The resting state of the eyes. *American Scientist, 72*, 378-387.

Parker A. J. (2004). From binocular disparity to the perception of stereoscopic depth. In L. M. Chalupa and J. S. Werner (Eds.) *The Visual Neurosciences* (pp. 779-792), Cambridge, MA: MIT Press.

Patterson, R. (1992). Human stereopsis. *Human Factors, 34*(6), 669-692.

Pedrotti, L.S. & Pedrotti, F.L. (1998). *Optics and Vision*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Pinkus, P. & Task, H. L. (1998). Measuring observers' visual acuity through night vision goggle*s. ASC98-1884.* Retrieved on June 21, 2003, from www.hec.afrl.af.mil/publications/VisAcu.pdf

Pounds, S. (June, 13, 2004), Florida universities cashing in on science. Palm Beach Post, Retrieved on May 1, 2006, from http://www.palmbeachpost.com/news/content/news/special_reports/scripps/a1f_techtrans_0613.html

Preece, J. Rogers, Y. & Sharp, H. (2002). Interaction design: Beyond human-computer interaction. New York, NY: John Wiley & Sons.

Rabbetts, R. B., (1998). *Bennett's & Rabbett's Clinical Visual Optics* (3rd ed.). Oxford: Butterworth Heinemann.

Reading, R.W. (1983). *Binocular Vision: Foundations and Applications*. Boston, MA: Butterworth Publishers.

Ricker, J. H. (1998). Traumatic brain injury rehabilitation: is it worth the cost? *Applied Neuropsychology, 5*, 184-193.

Rizzo, A. A. and Buckwalter, J. G. (1997) Virtual reality and cognitive assessment and rehabilitation: The State of the Art. In: G Riva (Ed), *Virtual Reality in Neuro-Physiology*, Amsterdam, Netherlands, IOS Press.

Rizzo, A. A., Buckwalter, J. G., & Van der Zaag, C. (2002). Virtual environment applications in clinical Neuropsychology (Chp. 50). In K. Stanney (Ed.) *Handbook of virtual environments: Design, implementation, and applications* (pp. 1027-1063). Mahwah, NJ: Lawrence Erlbaum.

Rizzo, A. A & Kim, G. J. (2005). A SWOT Analysis of the field of virtual rehabilitation and therapy. *Presence: Teleoperators and Virtual Environments, 14*(2): 119-146.

Rizzo A. A. (2003). SWOT analysis of the field of virtual reality. Keynote Address, *Proceedings of the Second International Workshop on Virtual Rehabilitation,* IWVR 2003, Piscataway, NJ, September 21-22.

Rizzo, A. A., Kim, G. J., Yeh S. C., Thiebaux, M., Hwang, J., Buckwalter, G.(2005).Development of a Benchmarking Scenario for Testing 3D User Interface Devices and Interaction Methods. *The Proceedings of the 11th International Conference on Human Computer Interaction.* Las Vegas, Nevada, July 22-27.

Rizzo A. A., Bowerly T., Buckwalter J.G., Klimchuk D., Mitura R, & Parsons T. D. (2006, January). A virtual reality scenario for all seasons: the virtual classroom. *CNS Spectrum, 11*(1), 35-44.

Rinalducci, E. J. (1996). Characteristics of visual fidelity in the virtual environment, *Presence: Teleoperators and Virtual Environment: Teleoperators and Virtual Environments, 5*(3), 330-345.

Rolland, J. P. (1990). Factors influencing lesion detection in medical imaging, Unpublished Ph.D. Dissertation, University of Arizona.

Robinett, W. & Rolland, J. P. (1992). A computational model for the stereoscopic optics of a head-mounted display. *Presence: Teleoperators and Virtual Environments,* 1, 1, 45-62.

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

Rolland, J. & Fuchs, H. (2001). Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoperators and Virtual Environments, 9*, 3, 287-309.

Rolland, J. P., Kreuger, M. W., & Goon, A. (2000). Multifocal planes head-mounted displays. *Applied Optics, 39*(19), 3209-3215.

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

Rolland, J. P., Meyer, C., Arthur, K., & Rinalducci, E. (2002). Methods 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*(6), 610-625.

Rolland, J. P., Ha, Y., & Fidopiastis, C. F. (2004). Albertian errors in head-mounted displays: I. Choice of eye-point location for a near- or far-field task visualization, *Journal of the Optical Society of America A, 21*(6), 901-911.

Rolland, J. P. & Hua, H. (2005). Head-mounted display systems. In R.G. Driggers (Ed.) *Encyclopedia of Optical Engineering* (pp. 1-13). New York, NY : Marcel Dekker.

Roorda, A. (2002). Human visual system: Image formation. In J.P. Hornak (Ed.) *The Encyclopedia of Imaging Science and Technology* (Vol. 1, pp. 557), New York, NY: John Wiley & Sons.

Roscoe, S. N. (1991). The eyes prefer real images. In S. R. Ellis, M. K. Kaiser, & A. J. Grunwald (Eds.) *Pictorial communication in virtual and real environments* (2[cd] ed., pp. 577-585).Washington, DC: Taylor & Francis.

Rose, F. D., Brooks, B. M., & Rizzo, A. A. (2005). Virtual reality in brain damage rehabilitation: Review. *CyberPsychology & Behavior, 8*(3), 241-262.

Rose, F. D., Attree, E. A., Brooks, B. M., Parslow, D. M., Penn, P. R., & Ambihapahan, N., (2000). Training in Virtual Environments: transfer to real world tasks and equivalence to real task training. *Ergonomics, 43*, 494-511.

Rose, F. D., Attree, E. A., Brooks, B. M., & Johnson, D. A. (1998). **V**irtual reality in brain damage: A rationale from basic neuroscience**.** In G. Riva, B. Weiderhold, & E. Molinari (Eds.) *Virtual Environments in Clinical Psychology and Neuroscience,* Amsterdam, Netherlands: IOS Press.

Rothbaum B. O., Hodges L. F. (1999). The use of virtual reality exposure in the treatment of anxiety disorders. *Behavior Modification, 23*(4), 507-525.

Rutledge, R. & Stutts, J. (1993) The association of helmet use with the outcomes of motorcycle crash injury when controlling for crash/injury severity. *Accident Analysis & Prevention, 25*(3), 347-353.

Saladin, J. J. (1998). Phorometry and stereopsis. In W.J. Benjamin (Eds.) *Boorish's Clinical Refraction* (pp. 724-773), Philadelphia, PA: W.B. Saunder's Company.

Schrauf, M. & Stern, C. (2001). The visual resolution of Landolt-C optotypes in human subjects depends on their orientation: the 'gap down' effect. *Neuroscience Letters, 299*(3) 185-188.

Schultheis, M. T. & Rizzo, A. A. (2001). The application of virtual reality technology for rehabilitation. *Rehabilitation psychology, 46*(3), 296-311.

Schiffman, H. R. (2000). *Sensation and Perception: an integrated approach* (5th *ed*.), New York, NY: Wiley.

Schuster, S., Strauss, R., & Gotz, K. G. (2002, September). Virtual-Reality techniques resolve the visual cues used by fruit flies to evaluate object distances. *Current Biology, 12*, 1591–1594.

Schutz, R.W. & Roy, E.A. (1973). Absolute error: The devil in disguise. *Journal of Motor Behavior, 5*, 141-153

Schutz, R.W. (1977). Absolute, constant, and variable error: Problems and solutions. In D. Mood (Ed.), *Proceedings of the Colorado Measurement Symposium* (pp. 82-100). Boulder, CO: University.

Szekeres, S. F., Ylvisaker, M. & Holland, A. L. (1985). Cognitive rehabilitation therapy: A framework for intervention. In M. Ylvisaker (Ed.) Head injury rehabilitation: Children and adolescents. Austin, TX: Pro-Ed, Inc.

Shapley, R. & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Progress in Retinal Research*, *3*, 263-346.

Sinai, M. J., Krebs, W. K., Darken, R. P., Rowland, J. H., and McCarley, J. S. (1999). Egocentric distance perception in a virtual environment using a perceptual matching task. *Proceedings of the 43rd Annual Meeting Human Factors and Ergonomics Society, 43*, 1256-1260.

Sloan, L. L. & Altman, A. (1954). Factors involved in several tests of binocular depth perception. *A.M.A Archives of Ophthalmology, 52*, 524-544.

Sloane, A. & Gallagher, J. R. (1945). Evaluation of stereopsis: A comparison of the Howard-Dolman and the Verhoff test. *Archives of Opthalmology, 34*(5), 357-369.

Sohlberg, M. M., & Mateer, C. A. (2001). *Cognitive Rehabilitation: An integrative neuropsychological approach*. New York: The Guilford Press.

Sosin, D. M., Sniezek, J. E., & Thurman, D. J. (1996). Incidence of mild and moderate brain injury in the United States, 1991. *Brain Injury, 10*(1), 47-54.

Speck, R. P. & Herz, N. E. (2000, April). Impact of automatic calibration techniques on HMD life cycle costs and sustainable performance. Paper presented at SPIE's 14th Annual International Symposium on Aerospace/Defense Sensing Simulation and Controls, Orlando, FL.

Stanney, K. M., Mollaghasemi, M., Leah, R., Breau, R., & Graeber, D.A. (2003). Usability engineering of virtual environments (Ves): identifying multiple criteria that drive effective VE system design. *International Journal of Human-Computer Studies, 588*, 447-481.

Stanney, K. M., Mourant, R. R., & Kennedy, R. S. (1998). Human factors issues in virtual environments: A review of the literature. *Presence: Teleoperators and Virtual Environments, 7*(4), 327-351.

Stanney, K.M. (1995). Realizing the full potential of virtual reality: human factors issues that could stand in the way. In the *Proceedings of IEEE Virtual Reality Annual International Symposium '95*, (pp. 28-32). Research Triangle Park, North Carolina, March 11-15.

Steinman, S.B., Steinman, B. A., & Garzia, R.P. (2000). *Foundations of binocular vision: A clinical perspective*. New York, NY: McGraw-Hill.

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.

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

Son, I. & Yazici, B. (2005). Near infrared imaging and spectroscopy for brain activity

      monitoring. Advances in sensing with security issues. Paper presented at NATO

      Advanced Study Institute, Il Ciocco, Italy, July 17-30.

Sugar A., Rapuano C. J., Culbertson W. W., Huang, D., Varley, G. A., Agapitos, P. J., de Luise,

      V. P., & Koch, D. D. (2002). Laser in situ keratomileusis for myopia and astigmatism:

      safety and efficacy. *Ophthalmology, 109*(1), 175–187.

Task, H. L. (1997). HMD Image source, optics, and the visual interface. In J.E. Melzer & K.

      Moffit (Eds.) *Head Mounted Displays: Designing for the user* (pp. 55-82). New York,

      NY: McGraw-Hill.

Tarr, M. J. & Warren, W. H. (2002). Virtual reality in behavioral neuroscience and beyond.

      Nature, Neuroscience Supplement, 5, 1089-1092.

Todorov, E., Shadmer, R., & Bizzi E. (1997). Augmented feedback presented in a virtual

      environment accelerates learning of a difficult motor task. *Journal of Motor Behavior,*

      29, 147-158.

Travis, D., Watson, T. & Atyeo, M. (1994). Human psychology in virtual environments. In L.

      MacDonald & J. Vince (Eds.) *Interacting with virtual environments* (pp. 43-59).

      Chichester, United Kingdom: John Wiley & Sons.

Tuceryan, M. and Navab, N. (2000). Single Point Active Alignment Method (SPAAM) for

      optical see-through HMD calibration. *Proceedings of IEEE and ACM International*

      Symposium on Augmented Reality, Munich, Germany, 149-158.

Tyrrell, R. & Leibowitz, H. (1990). The relation of vergence effort to reports of visual fatigue

      following prolonged near work. *Human Factors, 32*(3), 341-357.

Uttal, W. R. & Randall, W. G. (2001). On the psychophysics of night vision goggles. In   R. R. Hoffman & A. B. Markman (Eds.). *Interpreting remote sensing imagery: human factors*. (pp. 117-135). Boca Raton: Lewis Publisher.

Ylvisaker, M. (1998). *Traumatic Brain injury Rehabilitation, Children and Adolescents, (2nd ed.)*. Boston, MA: Butterworth-Heinemann.

Ylvisaker, M., Hanks, R., & Johnson-Greene D. (2002). Perspectives on rehabilitation of individuals with cognitive impairment after brain injury: Rationale for reconsideration of theoretical paradigms. *Journal of head trauma rehabilitation, 17*(3), 191-209.

Videotopia (1998). Home games. Retrieved on July 20, 2004, from http://www.videotopia.com/games2.htm

Vaissie, L., & Rolland, J. P. (2002). Eye-tracking integration in head-mounted displays, U.S. Patent 6,433,760 B1, August 13, 2002.

Wanger, L. R., Ferwerda, J. A., & Greenberg, D. P. (1992). Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics & Applications, 12*(3), 44-51, 54-58.

Walraven J., Enroth-Cugell C., Hood D. C., MacLeod D. A., Schnapf J. L. (1990) The control of visual sensitivity: receptoral and postreceptoral processes. In L. Spillmann & J.S. Werner (Eds.) *Visual perception: The neurophysiological foundations* (pp 53–101). San Diego, CA: Academic Press, Inc.

Wann, J. P., Rushton, S., & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision Research, 35*(19), 2731-2736.

Wann, J. P., Rushton, S., & Mon-Williams, M. (1996). What does virtual reality NEED?: Human factors issues in the design of three-dimensional computer environments. *International Journal of Human-Computer Studies, 44*, 829-847.

Warren, N. (1940). A comparison of standard tests of depth perception. *American Journal of Optometry, 17,* 208-211.

Watson, A.B. (1983). Quest: a Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113-120.

Web, N. A. & Griffin, M. J. (2002). Optokinetic stimuli: motion sickness, visual acuity, and eye movements. *Aviation, space, and environmental medicine, 73*(4), 351-8.

Weiss, P. L., Rand, D., Katz, N., & Kizony, R. (2004, December). Video capture virtual reality as a flexible and effective rehabilitation tool. *Journal of NeuroEngineering and Rehabilitation*, 1-12.

Watson, A.B. (1983). Quest: a Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113-120.

Westheimer, G. (1986). The eye as an optical instrument. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.) Handbook of Perception and Human Performance: Sensory Processes and Perception (Volume I, pp. 1-20), New York, NY: John Wiley and Sons.

Weymouth, F. W. & Hirsch, M. J. (1945). The reliability of certain tests for determining distance discrimination. *American Journal of Psychology, 58*(3), 379-390.

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

Willemsen, P. & Gooch, A. (2002, February). An experimental comparison of perceived egocentric distance in real, image-based, and traditional virtual environments using direct walking tasks, Technical Report UUCS-02-009, School of Computing, University of Utah.

Wilson, J. R., Eastgate, R. M., & D'Cruz, M. (2002). Structured development of virtual environments. In K. M. Stanney, (Ed.), *Handbook of Virtual Environments Technology* (pp. 353-378), Lawrence Erlbaum Associates, Inc.

Wilson, B. (2000). Compensating for cognitive deficits following brain injury. *Neuropsychological Rev*iew*, 10*, 233–243.

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

Witmer, B. G. & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225-240.

Whitney, S. L., Sparto, P. J., Brown, K., Furman, J. M., Jacobson, J. L., & Redfern, M. S. (2002, September). The potential use of virtual reality in vestibular rehabilitation: Preliminary findings with the BNAVE. Neurology Report. Retrieved on April 22, 2006, from http://www.findarticles.com/p/articles/mi_qa3959/is_200206/ai_n9113429.

Wong, B. P., Woods, R. L., & Peli, E. (2002). Stereoacuity at distance and near. *Optometry and Vision Science, 79*(12), 771-778.

Zhang, L., Abreu, B., Seale, G., Masel, B., Christiansen, C., & Ottenbacher, K. (2003). A virtual reality environment for evaluation of a daily living skill in brain injury rehabilitation: reliability and validity. *Archives of Physical Medicine and Rehabilitation, 84,* 1118-1124.