Effects of Field of View on Performance with Head-Mounted Displays

by

Kevin Wayne Arthur

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill

2000

Approved by

Advisor: Dr. Frederick P. Brooks, Jr.

Reader: Dr. Vernon A. Benignus

Reader: Prof. Mary C. Whitton

Dr. Gary Bishop

Dr. Russell M. Taylor II

ABSTRACT

Kevin Wayne Arthur

Effects of Field of View on Performance with Head-Mounted Displays

(Under the direction of Frederick P. Brooks, Jr.)

The field of view (FOV) in most head-mounted displays (HMDs) is no more than 60 degrees wide – far narrower than our normal FOV of about 200° wide. This mismatch arises mostly from the difficulty and expense of building wide-FOV HMDs. Restricting a person's FOV, however, has been shown in real environments to affect people's behavior and degrade task performance. Previous work in virtual reality too has shown that restricting FOV to 50° or less in an HMD can degrade performance.

I conducted experiments with a custom, wide-FOV HMD and found that performance is degraded even at the relatively high FOV of 112°, and further at 48°. The experiments used a prototype tiled wide-FOV HMD to measure performance in VR at up to 176° total horizontal FOV, and a custom large-area tracking system to establish new findings on performance while walking about a large virtual environment.

FOV was significant in predicting performance of two tasks: searching for and locating a target by turning one's head, and walking through a simple maze-like environment while avoiding walls. Wide FOV (112° or greater) was especially important for the walking task; for it, performance at 112° was 23% less than at 176°. At 48°, performance was 31% less than at 176°. For the search task, performance at 112° was 12% less than at 176°. At 48°, performance was 24% less than at 176°.

Additional analyses of the data show trends that suggest future investigation. Restricting FOV appears to decrease the user's sense of presence, as measured by a questionnaire. VR sickness, also measured by questionnaire, increased with successive exposures to our system within an hour-long session, but stayed at relatively low levels. FOV appears to alter the occurrence of some sickness symptoms, but the data are inconclusive on whether FOV predicts total sickness. I performed additional measures and analyses, including tests of postural instability, distance memory, spatial memory, head-movement behavior, and comparisons with other HMDs and with real-world performance.

ACKNOWLEDGMENTS

Thanks to

Frederick P. Brooks, Jr. for being my advisor and for his insights and encouragement;

Frederick P. Brooks, Jr., Mary C. Whitton, Vernon A. Benignus, Russell M. Taylor II, and Gary Bishop for serving on my doctoral dissertation committee and for all that they have taught me about doing research;

Gary Bishop for suggesting field of view as a dissertation topic;

Jannick Rolland and Frank Biocca for serving as my research advisor during parts of this work;

Keith Muller for serving on my dissertation committee during part of this work;

Zac Kohn, Wolfgang Stürzlinger, Phil Winston, and the Tracker and nanoManipulator project teams for software support, Kurtis Keller and David Harrison for equipment support, and Darlene Freedman and Janet Jones for administrative support;

Mike Meehan, Zac Kohn and others on the Hybrid Reality/House team for comments on pilot versions of the experiments;

My subjects for their contributions to this work;

Gina Kolasinski for discussions about simulator sickness;

The participants and organizers of the CHI 96 Doctoral Consortium at which I presented a description of my dissertation proposal;

My colleagues from UBC, in particular Kelly Booth, for continued advice and friendship;

The UNC-CH Department of Computer Science and the following projects and P.I.s for financial support through assistantships: Walkthrough and GRIP (Frederick P. Brooks, Jr.), Psychophysics of Optical See-Through HMDs (Jannick P. Rolland), and Telepresence (Henry Fuchs and Gary Bishop);

For financial support for equipment used in this work: Kaiser Electro-Optics, DARPA, NIH National Center for Research Resources number RR02170, and Intel;

For photographs: Kurtis Keller, Todd Gaul, and Kaiser Electro-Optics;

and most importantly,

My family – Rawleigh, Gwen, Steven, and Shelley Arthur – for their belief in me and for supporting me for all these years;

Maureen McClarnon, for her love and support.

CONTENTS

TABLES

FIGURES

Chapter 1
Introduction

1.1 Introduction

The human field of view (FOV) spans approximately 200 degrees horizontally, taking into account both eyes, and 135 degrees vertically (Gibson, 1979; Werner, 1991; Barfield et al., 1995). (Figure 1-1 shows the typical human FOV for one eye.) Virtual reality (VR) systems replace that view with a simulated one, generated in our case by a head-mounted display and a computer graphics system (Sutherland, 1965; Ellis et al., 1993; Brooks, 1999). In VR, to the extent that the technology affords it, we can perform tasks and feel "present" in the virtual environment the same way we feel present in the real world. For VR to be effective, the design or choice of system needs to take into account the characteristics of human perception and performance. A key design variable for a head-mounted display (HMD) is the FOV size. Most HMDs offer limited FOV, often only 40° to 60° horizontally and 30° to 45° vertically. This work reports on studies to determine how such restrictions of FOV affect performance in VR.

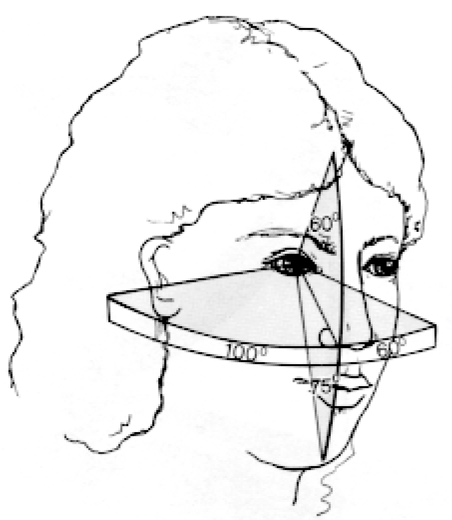

Figure 1-1. Normal human field of view. From Werner (1991).

The effects of FOV on performance have been studied in other contexts. Scientists have tested peoples' behavior while viewing the real world through FOV-restricting goggles (blinders).

Narrow FOV has been shown to degrade performance on locomotion, visual search, and spatial awareness tasks; it causes longer task completion times, disrupted eye- and head-movement coordination, and misperception of size, space, and ego-center (Dolezal, 1982a; Alfano and Michel, 1990). Limited studies have shown some similar effects to occur in VR or flight simulators, in locomotion by flying (Kenyon and Kneller, 1993; de Vries and Padmos, 1997) and in searching tasks (Wells and Venturino, 1990; Piantanida et al., 1992; Cunningham et al., 1996).

Too wide a FOV in VR may also degrade performance; it may cause VR sickness. Researchers have hypothesized that wide FOV in VR will aggravate sickness due to vection and visual-vestibular mismatch (Kolasinski, 1995), which can occur when there are lags in the VR system or other sources of mismatch. Results to date, including those from the present study, have neither confirmed nor refuted this hypothesis, but given how low the levels of VR sickness are that we see with current systems it appears unlikely that wide FOV will cause serious VR sickness under typical circumstances. It may, however, aggravate VR sickness in unusual circumstances with high vection; moving backgrounds, travel by flying, or rotating-scene stimuli (Stern et al., 1990; So and Lo, 1999) may be such cases.

Too wide a FOV may of course be unnecessary for tasks where the 3D region of interest is small, and therefore inefficient given the engineering costs. Thus FOV choice depends on the task as well as on considerations of performance and sickness.

1.2 New approaches

This document presents research investigating the effects of FOV on task performance, sickness, and other factors in VR. In particular I studied the following:

- Performance with a wide-FOV head-mounted display. Previous work comparing FOV sizes in VR has used HMDs with FOV of approximately 120° or less[1]. For the present work I used a 176° tiled display from Kaiser Electro-Optics – the widest FOV available in a HMD at the time of the study.

[1] In studies where *total* FOV was the variable, as it is in the present study, the highest FOV tested in VR is 90° (Piantanida et al., 1992). Piantanida et al. compared performance at total FOV of 90°, 53°, and smaller. They used a VPL EyePhone HMD and reported its FOV as “approximately 100°”; Robinett and Rolland (1992), however, have measured its actual FOV to be 90°. In a different type of FOV study, where total FOV remained constant but stimulus FOV varied, the largest FOV tested in an HMD is 120°, by Wells et al. (1990) using the VCASS HMD for a flight simulator. Total FOV remained constant at 120° and the size of an inner region containing target stimuli was reduced.

- Performance when walking about a large environment. I used the UNC-CH optical ceiling tracker to allow subjects to walk about a 10 m by 4 m environment. This allowed for measuring the effects of FOV on locomotion in VR by real walking, which has not been tested previously.

- Performance of generic spatial tasks chosen to be representative of the tasks typically performed in virtual environments. These generic tasks are: a search task, a walking task, a distance judgment task (distance from observer to object), and a spatial memory task (arrangement of objects with respect to each other).

- Health effects of FOV. VR sickness was measured by questionnaire and by postural instability measures.

These tasks were performed with three different FOV sizes (48°, 112°, and 176° horizontal by 47° vertical, as calibrated in the Kaiser HMD). For comparison, some tasks were tested in the Virtual Research V8 HMD (48° by 36°) and in a real-world blindered condition (48° by 36°).

The following thesis statement summarizes the hypotheses that were verified by the studies.

1.3 Thesis statement

Restricting FOV in a head-mounted display degrades human performance. Performance is degraded even at the moderately high FOV of 112° horizontal. The severity of the effect depends on the task: it affects locomotion (or travel) performance most.

1.4 Summary of findings

I tested task performance at three levels of horizontal FOV: 48°, 112°, and 176° and found the following statistically significant results ($p < .0125$). In all cases the percent-performance comparison is to 176° FOV.

Restricting FOV degrades performance on searching and walking tasks at both 48° and 112°. Gains can be had beyond 112°, which is wider than the FOV available in most commercially available HMDs.

Searching: restricting FOV degrades performance by 12% at 112° and by 24% at 48°. For a headcentric search task with targets whose initial position was outside the visual field, performance decreased by the same amount between the two levels (12% decrease in performance for each drop of 64°). If we consider performance at 176° to be 100%, then restricting FOV to 112° decreased performance to 88%; restricting further to 48° decreased performance to 76%.

Walking: restricting FOV degrades performance by 23% at 112° and by 31% at 48°. For a task of walking through a simple maze-like environment while avoiding walls, performance decreased most in the drop from 176° to 112°, suggesting that a wide FOV benefits locomotion (travel) most. If performance at 176° is 100%, then we saw performance at 112° of 77% (a drop of 23%) and performance at 48° of 69% (a further drop of 8%).

Sickness and FOV. The data were not statistically significant on FOV predicting sickness or not. Opposing (and non-significant) trends were seen, however, on the nausea and disorientation subscores of the SSQ. The levels of sickness seen were low in all conditions.

Visual quality is still important. The vastly better visual quality, the lower weight, and the lower display latency available with the V8 HMD and in the real-world with blinders may have been enough to overcome the performance losses due to restricted FOV for the tasks I tested. Performance in those conditions was slightly better than with the Kaiser HMD at 176°. Practice effects may also have caused some of this improvement.

Future work. Areas for future research are discussed, including the following.

- Augmented HMDs. New HMD designs that use add-on ambient peripheral displays may be the best route towards balancing the trade-offs between FOV and rendering speed. Studies are needed to measure their effectiveness and to choose the best design parameters.

- FOV and presence. The present study found a trend in presence scores, that presence was lower with restricted FOV. Further studies are recommended to establish statistical significance for this hypothesis, using more samples and better measures of presence.

- VR sickness and postural instability. The present study found low sickness levels in general, and found no trend of higher total sickness or postural instability with wider FOV as had been hypothesized. There was a trend of disorientation subscores increasing with wider FOV and nausea subscores decreasing with wider FOV. Finer measures, longer exposure times, and more statistical power would be needed to prove or disprove this trend. It's possible that in practice sickness will not be significant enough that FOV size will change it. The same may be true of postural instability in VR.

1.5 Definition of terms

Field of view (FOV). The angular extent subtended by a display in front of the eyes. The typical human FOV is about 200° horizontal by 135° vertical (Werner, 1991), though it varies with the person. More specifically, this is the **total binocular FOV** – the angular extent seen by both eyes. The **monocular FOV** is the angular extent seen by one eye. The **binocular overlap (or stereo overlap)** is the central region of the two monocular FOVs that overlaps. In this work I varied horizontal FOV only, and therefore in this document I usually write FOV for short, meaning total binocular horizontal FOV size.

Head-mounted display (HMD). A display worn on the head to provide a view of a computer-generated scene. I assume stereoscopic binocular HMDs with appropriate perspective projections, though non-stereoscopic HMDs also exist – monocular HMDs that show one image to one eye, or biocular HMDs that show one image to both eyes.

Virtual reality (VR). The illusion or state of being present and visually immersed in a simulated three-dimensional environment, evoked in our case through use of a head-tracked HMD. I assume that graphics and tracking systems are used to provide dynamic, geometrically correct perspective views from arbitrary viewpoints, and that the user can move his or her head and body to change the view. I will occasionally use the term **virtual environment (VE)**, and take it to mean the same as VR.

VR Sickness. Any sickness arising from exposure to a virtual reality system. VR sickness includes symptoms such as nausea, headache, and disorientation, which are commonly measured by the **simulator sickness questionnaire (SSQ)** (Kennedy et al., 1993). Another VR sickness symptom is **postural instability** (or ataxia), which is an impaired ability to stand up straight, analogous to effects produced by alcohol (Kennedy and Lilienthal, 1995). VR sickness is related to, but is not equivalent to, simulator sickness, which is related to motion sickness (Kennedy et al., 1993; Stanney et al., 1997). VR sickness has also been called cybersickness or VE sickness.

I will use the following terms frequently in describing the experiments:

- **Session** – a visit to the laboratory to perform several tasks during multiple VR exposures. Each session took place on a different day.

- **VR exposure** – one “go” in the VR system, usually lasting 5 or 6 minutes. Subjects had three VR exposures in each session.

- **Task** – there was typically one main task per VR exposure.

- **Trial** – a single instance of the task. Trials were grouped into blocks with breaks between.

1.6 Overview

Chapter 2 reviews the literature on human factors in VR, FOV options for HMDs, VR sickness, and known effects of FOV on performance. Chapter 3 describes the methods and research questions. Chapter 4 presents results and analysis of the main hypotheses. Chapter 5 presents results and analyses on exploratory hypotheses. Chapter 6 discusses areas for future work. The Appendices present additional data summaries and documents.

Chapter 2
Review of the literature

2.1 Human factors of virtual reality from a task-level standpoint

Several researchers have classified and outlined human factors issues of VR. Stanney et al. (1998) review human factors issues for VR from three perspectives: human performance efficiency, health and safety, and social impact. Barfield et al. (1995) take a lower-level approach and review the mismatches between human sensory capabilities and the display capabilities of VR equipment. Melzer and Moffitt (1997) review the issues and relate them specifically to HMD design and selection requirements. They categorize the issues as: visual requirements (derived from understanding of the human visual system), physical requirements (derived from anthropometry and ergonomics knowledge), environmental requirements (according to the physical environment the HMD is to be worn in – requirements for comfort, etc.), and interface requirements (the ease and appropriateness of controls on the HMD).

Recent years have also seen an increase in research activity aimed at quantifying the sense of "presence" in virtual environments (Barfield and Weghorst, 1993; Slater and Usoh, 1993a) and studies to compare human performance with and without HMDs (Henry and Furness, 1993; Pausch et al., 1993).

2.1.1 Tasks for human factors studies in virtual reality

Key to understanding human performance in virtual environments is identifying the tasks that will be performed in them. Lampton et al. (1994) describe a set of tasks (the "Virtual Environment Performance Assessment Battery" or VEPAB) developed for testing training applications of virtual environments. Their primary goal was to develop tasks that "produce cost-effective transfer of training from VE practice to real-world performance." Gerth (1997) discusses task issues in performance-based testing of HMDs.

I consider below tasks in terms of a "frames of reference" system that has been used previously to characterize human task performance (Lee, 1977; Feldman, 1985; Howard, 1993). The following sections describe four common classes of tasks: search, locomotion, judgments of distance and space, and manipulation (reaching and grasping). These are shown in Table 2-1 with reference to the appropriate frames of reference according to Howard's system (1993).

Table 2-1. Classes of tasks in VR and frames of reference.

Task	**Frame of Reference**
Search	Headcentric
Locomotion (Travel)	Bodycentric
Judgments of distance and space	Egocentric or exocentric
Reaching	Handcentric

In the following sections I describe theses classes of tasks and their previous use in VR studies. Later, in Section 2.4 I revisit these tasks and discuss implications of FOV.

2.1.1.1 Search

To some extent, everything to be done in a virtual environment involves visual search – locating a target visually. To travel we need to find the place we wish to move to; to reach for something we usually need to locate it visually; to "take in" an environment and form a mental model of it we need to scan it visually. Visual search is a task that has been widely studied outside of VR (Stark et al., 1992), and has been used in VR studies of field of view (Piantanida et al., 1992) and degradation of the periphery in HMDs (Watson et al., 1997), and comparison of HMD performance with "desktop" (non-tracked) performance (Pausch et al., 1997).

2.1.1.2 Locomotion

Many virtual environments are large enough that the user needs a way to move about them (beyond just moving his or her head). For this reason it is important to provide an effective method for the user to get to where he wants to be – through actual walking, virtual flying (Robinett and Holloway, 1992), walking in place (Slater et al., 1995), or other methods. The effectiveness of a method of navigation, and the suitability of that method for a given display technology, can be evaluated by giving the user a task such as "walk along the path marked on the floor without stepping outside of it" (Alfano and Michel, 1990), or "fly through the tunnel without hitting the walls" (Ware and Osborne, 1990). Recently, Jorgensen et al. (1997) described studies in progress to compare performance and physiological responses when navigating through randomly generated maze scenes.

2.1.1.3 Judgments of distance and space

The issue here is how accurately subjects perceive the spatial layout of a virtual environment, including the distances to and between objects. Correct perception of space and distance is vital for applications such as architectural walkthrough (Brooks, 1986) and medical visualization. I use the term “distance judgments” here to mean judgments of the distance from oneself to an object (an egocentric judgment). By “spatial judgments” I mean judgments of distance between objects, external to oneself (exocentric judgments) (Howard, 1993).

Spatial memory can be evaluated using a task such as that used by Alfano and Michel (1990) in real-world studies, where subjects were asked to look around an office for 60 seconds and were then removed from the room and asked to indicate the positions of objects by positioning icons onto a 2D map.

Dinh et al. (1999) measured spatial memory in VR as a function of multisensory inputs. They added tactile, olfactory and auditory cues to a virtual environment (of an office) and measured presence and memory. They questioned people on their memory of where different objects were located in the environment, and found that with olfactory and tactile cues, subjects performed significantly better in recalling the location of objects. Additional studies have compared memory performance in VR to memory performance in the real world (Billinghurst and Weghorst, 1995; Hoffman et al., 1995).

2.1.1.4 Manipulation (Reaching and grasping)

These are tasks that involve manual manipulation, using one or both hands. Such hand movements can be considered as having two phases: reaching (rapidly moving the hand to the vicinity of a target for manipulation) and grasping (using the fingers and thumb to have an effect on the object) (Sivak and MacKenzie, 1992). Most VR applications involve some type of hand-based interaction, and thus it is important that reaching and grasping can be performed adequately. Various 3D reaching tasks have been applied to study 3D interaction in head-tracked displays (non-HMD) (McKenna, 1992; Ware and Balakrishnan, 1994).

2.2 Field of view and head-mounted display design

2.2.1 Field of view in conventional head-mounted displays

This section describes the FOV sizes of currently available HMDs and some methods that have been suggested to provide wider FOV or extra peripheral cues. Figure 2-1 shows FOV sizes for several HMDs. Most data are from August 1999 and 1998 surveys (Latham, 1998; Latham, 1999). The diamond-shaped point represents the Kaiser tiled HMD, which will be discussed in the next section. Increasing the FOV in an HMD complicates the optical design and increases cost.

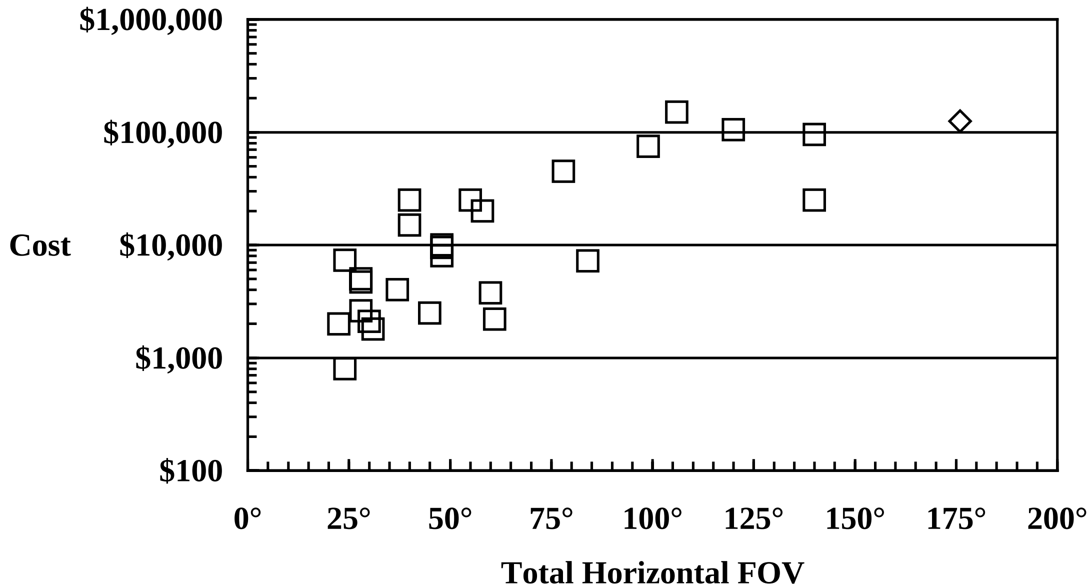

Figure 2-1. FOV and cost of several currently available HMDs.

Some commercially available displays, such as the Virtual Research Flight Helmet or VPL EyePhone, use distorting LEEP (Large Expanse Extra Perspective) wide-angle optics to increase the FOV (Howlett, 1990). These distort the scene to provide a total horizontal field of view of 90 degrees, with a binocular overlap of approximately 60 degrees (Robinett and Rolland, 1992). Other displays, using a conventional lens system with less distortion, have horizontal FOV typically less than 70°.

In a simple lens system for an HMD, FOV interacts with screen size, lens size, eye relief distance, exit pupil size, and focal length (Robinett and Rolland, 1992; Davis, 1997; Fischer, 1997; Task, 1997). Increasing the FOV for a simple lens system requires increasing the size of the display screen (LCD in most cases), increasing the diameter of the lens, or decreasing the focal length, which shortens the eye relief distance. This relationship is expressed by $E = D - L_e S / f$, where D is the lens diameter, S is the screen diameter, f is the focal length, L_e is the eye relief distance (the distance of the eye from the lens), and E is the exit pupil diameter (the diameter of the window through which the eye can move laterally to still see an image at the eye relief distance) (Task, 1997).

Shortening the eye relief distance is undesirable because it makes the display less likely to fit many users. Whitestone and Robinette (1997) discuss design requirements that affect how well HMDs fit multiple users. Increasing lens size is undesirable because it adds cost and weight to the HMD. Perry and Buhrman (1997) recommend that HMDs be symmetric and weigh less than approximately 3 lb., with the center of gravity at most 2 inches in front of the head's *y* axis (center of side-to-side rotation).

FOV also interacts with other system characteristics. Reducing FOV might reduce the size of the binocular overlap FOV, reducing the effectiveness of stereo. An overlap region that is too small can result in brightness artifacts called "luning" (Davis, 1997). FOV varies inversely with resolution for a given frame buffer size. Increasing the frame buffer size (the number of pixels) will help with resolution but will slow down the frame rate. FOV requirements depend also on whether the HMD is see-through or not. Augmented reality systems using see-through HMDs benefit from the fact that the real world is also visible through the display, and thus the user gets peripheral cues for locomotion from the real environment.

2.2.2 Increasing FOV by using other HMD designs

One approach to increasing the FOV available in an HMD is to simulate wider FOV using distortion. Slater and Usoh (1993b) suggest distorting the displayed images in a way that compresses more of the scene into the outer regions of the display area. The central (foveal) region of the display is projected normally (they suggest that this be approximately 90 per cent of the display's field of view), and in the outer 10 per cent, a larger portion of the field of view than normal is imaged (by adding a shear to the projection). Their hypothesis is that since a primary function of peripheral vision is to detect objects and trigger head movements to foveate them, the compressed peripheral display will serve to trigger head movements and make the user aware of a larger space than would be possible with the standard uncompressed display. The distortion that Slater and Usoh propose employs linear perspective onto multiple flat image planes. More general distortion methods (such as fish-eye projection) are commonly used for presenting 2D maps, and these methods may suggest other projections useful for 3D display in HMDs (Leung and Apperley, 1994). Slater and Usoh describe an implementation of their method on a CRT, but have yet to implement the method in an HMD. Draper (1998) tested horizontal compression in an HMD and found that it increased simulator sickness.

The Kaiser FIHMD ("full immersion head-mounted display") achieves wide FOV by tiling multiple screens inside the display. The HMD contains 12 liquid crystal displays placed in a 3 by 2 arrangement in each eye. In our calibration it provides a total FOV of approximately 176° horizontal by 47° vertical. This display is discussed further in Chapter 3.

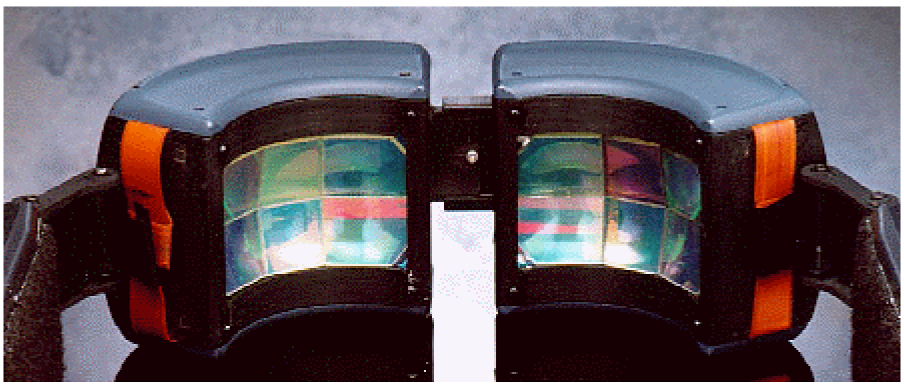

Figure 2-2. Kaiser HMD.

Figure 2-3. Kaiser HMD showing helmet mount.

Others have suggested adding low-fidelity displays to a conventional HMD to provide peripheral information. The use of low-fidelity peripheral display has been demonstrated by designers of helicopter and airplane cockpits who have experimented with the use of an ambient "peripheral vision display" that projects a laser beam representing the horizon across the cockpit, thereby providing an orientation cue that does not interfere with the pilot's primary task, and studies have shown its effectiveness (Malcolm, 1984; Nordwall, 1989). The Ambient HMD project at Monterey Technologies has shown that adding peripheral LCD displays improves performance in a helicopter simulator (Monterey Technologies, 1999).

Researchers have also shown that displaying icons within the visual field can effectively indicate the presence and/or location of objects outside the FOV. Geiselman and Trou (1996) conducted a study to compare five different types of "vector symbology" for indicating the presence of an object outside a pilot's FOV. They used a HMD with a 22° circular FOV. All symbols were vectors extending between the center of the FOV towards the edge, and used different lengths or positions or numbers to indicate how far the target is outside the FOV. Craig et al. (1997) performed a similar study to evaluate symbols for indicating targets outside a pilot's FOV (their term was "target locator cues" or TLCs). Their study used three of the methods tested previously by Geiselman and Trou in addition to others, using a dome screen with FOV restricted by software to 10° circular, and with the Viper II HMD with a 20° FOV. They recommend further studies to determine the effects of clutter in the display on performance with these cues.

To offset the rendering slowdown caused by a wider FOV and correspondingly larger frame buffer, researchers have proposed "area-of-interest" HMDs. Area-of-interest displays are suggested by the fact that, in the human visual system, resolution in the periphery is much coarser than in the fovea (Hood and Finkelstein, 1986; Davson, 1990). Frame rates can be increased by displaying high resolution images only in a small area of interest, and displaying low-resolution images elsewhere. HMDs with built-in eye tracking allow for moving this high resolution insert, either mechanically or electronically, to where the eye is looking at a given time (Toto and Kundel, 1987; Yoshida et al., 1995). Moving area-of-interest displays have yet to be commercialized and it remains to be seen whether the modest frame rate increase merits the extra cost of fast eye tracking. Watson et al. have proposed using a fixed area-of-interest display without the need for eye tracking. Their argument is that the high resolution insert does not need to move, given that typical eye movements do not go farther than 15 degrees from the fovea. They investigated through user studies the use of a display with a fixed higher resolution insert and lower resolution periphery (Watson et al., 1997). The study used a Virtual Research Flight

Helmet display and a visual search task and found no significant difference between performance at full resolution and performance with a low-resolution periphery and full-resolution insert.

2.3 Sickness in virtual reality

In addition to task performance we need to consider other human factors, including health factors (Mon-Williams et al., 1993; Stanney, 1995). Administering simulator sickness questionnaires as part of user studies can help to identify those visual factors in HMDs that lead to sickness symptoms (Kennedy et al., 1993; Kolasinski, 1995).

The Simulator Sickness Questionnaire (SSQ) was developed by Kennedy et al. (1993) and remains the questionnaire most commonly used for testing sickness in flight simulators or virtual reality systems. They adapted the older "Motion Sickness Questionnaire" (MSQ) for use with flight simulators. They found that simulator sickness was in general less severe than motion sickness and therefore required a finer test. Simulator sickness also affects a smaller population than does motion sickness. Simulator sickness differs in the nature of the stimulus – in a simulator the participants can usually close their eyes to make all stimuli stop (except when the motion platform is moving). This is not true with motion sickness (in a car, for example), where the vestibular input continues at all times. Motion sickness occurs primarily because of the vestibular input, whereas with simulators or VR, visual stimulation alone can cause it. (For this discussion I assume that simulators have motion platforms and VR systems do not.)

Kolasinski (1995) summarizes theories of sickness applicable to virtual reality and the factors that contribute to it. She groups these factors into three categories: those associated with the individual, those associated with the simulator, and those associated with the task. Simulator sickness is by nature polysymptomatic, meaning that no single symptom dominates, and polygenic, meaning that no single factor is the cause. The primary theory used to explain simulator sickness is the cue conflict theory: that disparity between or within sensory inputs (primarily visual and vestibular) causes sickness. The link between the perceptual mismatch and the actual sickness that results has not been proven (Oman, 1993), though, according to Kolasinski (1995), Treisman (1977) hypothesized that sickness results from a mechanism designed to protect the body against toxins (which can produce similar sensory disruptions). A second theory (Riccio and Stoffregen, 1991) asserts that, rather than cue conflict, it is postural instability (ataxia) that precedes and causes sickness symptoms; that sickness results when the participant does not have the inputs or strategies for maintaining postural stability.

In addition to sickness symptoms and ataxia, simulator exposure may result in dark focus shifts. The dark focus is the physiological resting point of accommodation; it might be shifted inward after exposure (Kolasinski, 1995). Kolasinski suggests using this phenomenon as another measure for simulator sickness, but points out that it may be confounded with other factors due to the task (a more demanding task tends to result in less dark focus shift).

Factors associated with the simulator that Kolasinski cites include field of view (though she states that it may be an indirect effect, related to the degree of texture and flicker), refresh rate, resolution, time lag, and the calibration accuracy of the display. Kolasinski points to future sickness research being needed in the areas of correlating scene/application elements with sickness, measuring and using eyestrain effects, and testing physiological measures of sickness, including EGG, skin conductance, and heart rate.

Stanney et al. (1998) distinguish between microscopic health effects (tissue-level, such as eyestrain) and macroscopic health effects (such as sickness and trauma) of virtual reality. They state that there are no definitive predictive theories of simulator sickness, but that the vection-induced cue-conflict theory is the most widely accepted. They also state that asynchrony may be worse than lag, and that interactive control may be an important method for reducing sickness effects (Stanney and Hash, 1998).

In their original discussion of the SSQ, Kennedy et al. (1993) discuss the effect of time between exposures on sickness. For a mid-range time of 2-5 days between exposures (or "hops" in the simulator terminology), pilots could adapt to the system and thereby experience less sickness. Sickness was higher in the other cases – when hops occurred on the same day or one day apart, or when hops were more than five days apart.

Kennedy and Lilienthal (1995) state that, although over the years it has been implied and stated in the flight simulator community that sickness due to cue conflicts would eventually go away – that better technology and thus more realism and fidelity would result in less sickness – this has not in fact happened.

Bittner et al. (1997) discuss how sickness may be a confounding factor in many studies of performance, and argue for compensating for the differential effects of simulator sickness on performance by including sickness as a covariate in an analysis of covariance (ANCOVA) when analyzing performance.

2.3.1 Measuring sickness

As mentioned previously, the Simulator Sickness Questionnaire (SSQ), developed by Kennedy et al. (1993) remains the questionnaire used most commonly for testing sickness in simulators or virtual reality. (The questionnaire is reproduced in Appendix B.4.) The questionnaire is a checklist of 16 symptoms to which the participant responds on a scale of 0 through 3 (none, slight, moderate, or severe). Participants usually fill out the questionnaire both before and after exposure to a simulator or VR system, but the authors suggest using only the post-exposure scores for indications of sickness induced by the exposure, rather than difference scores. Difference scores (i.e., the post-exposure score minus the pre-exposure score) are less reliable statistically. (Pre-exposure scores are useful for determining that a subject is fit to use the system to begin with.)

To introduce better "handles" on the 16 symptoms, the researchers did a factor analysis (analogous to finding eigenvectors) (Kennedy et al., 1993). They isolated three independent factors and one general factor. The scores for these four factors are weighted sums of the responses on selected symptoms. The scoring gives an ordinal scale that has no inherent meaning and is only provided to allow comparison between different simulators. The general factor (or "total score") is to be taken as the best indication that a sickness problem exists; the three independent factors are then to be used to diagnose the problem (to determine what part of the experience is causing the problem, not that it is known to exist). These three factors were given the following names, based on symptoms that comprise them: oculomotor discomfort (eyestrain, difficulty focussing, blurred vision, headache, etc.), disorientation (dizziness, vertigo, etc.), and nausea (nausea, stomach awareness, salivation, burping, etc.).

The basics of SSQ use are summarized by Hettinger and Kennedy (1996). They recommend first looking at the SSQ total score, then, if it is above average, looking at the 3 diagnostic subscores to determine the profile of the sickness score and determine which part of the simulator to be concerned with. For a higher-than-average nausea score, they recommend checking motion base asynchronies, lags, and washout. For high oculomotor scores they recommend checking the lighting, distortion, off-axis projections, textures, and refresh rate. For high disorientation or disequilibrium, they recommend checking the amount of vection, motion base responses for close work, and pseudo-coriolis (gravition motion in simulator).

The relative ordering of the three subscores is referred to commonly as the sickness profile. Stanney, Kennedy and Drexler (1997) compared the profiles of simulator sickness and VR sickness (cybersickness) and found that the total severity for VR was three times that of simulator sickness and that the ordering of symptom clusters was different. In simulator sickness (based on over 3000 reports from more than 30 simulators), oculomotor symptoms predominated, followed by nausea, followed by disorientation; in VR (based on 8 experiments with 4 VR systems), the order was disorientation, followed by nausea, followed by oculomotor. Stanney et al. (1998) report further, regarding the D > N > O profile that “this VE profile is reliable, having been replicated with five different VE systems, using different HMDs” and suggest that a new factor analysis of SSQ clusters may be required to optimize the use of the SSQ for VR systems. It’s not clear, however, that VR systems are similar enough at present to produce a reliable result, or whether this pattern is intrinsic to VR systems. For example, the profile differed across the different locomotion conditions tested in the study by Usoh et al. (1999). So and Lo (1999) reported a consistent O > D > N profile for VR sickness in their system, using a Virtual Research VR4 display with 48° by 36° FOV. Their study compared sickness levels resulting from scene oscillation with the head physically stationary. They used both the SSQ of Kennedy et al., and a nausea self-report question on a scale of 0-6. They found that rotation in any of the three axes had similar effects, and that nausea was reported after approximately 5 minutes of exposure to VR with the scene rotating at 30°/s. They state that without scene oscillation, viewing a stationary virtual environment for 15 minutes or less is not likely to induce nausea.

Physiological measures have also been used to detect simulator sickness or motion sickness. Stern et al. (1990) used the electrogastrogram (EGG) to measure motion sickness (to detect “abnormal gastric myolectric activity” associated with nausea) in addition to subjective reports (for a real-world scenario not using VR).

2.3.2 Postural instability

Kennedy et al. (1995) have used a video-based procedure to measure postural stability of users after exposure to a simulator. The person is asked to stand still for 30 seconds and a videotape records them. People tend to sway after long simulator exposures, much the same as the effect of alcohol, and the magnitude of this oscillation can be taken as a measure of simulator sickness. Hettinger and Kennedy give further sample ataxia results in course notes (Hettinger and Kennedy, 1996).

Kolasinski (1995) suggests using a head tracker, already part of a VR system, to measure postural stability. In later work, Kolasinski (1996) measured postural instability before and after 20-minute exposures but did not find significant instability.

2.4 Effects of field of view

2.4.1 Effects of field of view on sickness

Stern et al. (1990) have shown that in a real-world rotating-drum test, wide FOV leads to increased motion sickness and that restricting the FOV or adding fixation to the task greatly help to reduce the occurrence of vection effects. Vection refers to the illusion of self-motion, when we perceive ourselves to be moving when our body is not, and it has been cited as a key indicator of simulator sickness (McCauley and Sharkey, 1992; Pausch et al., 1992). This type of sickness resulting from cue conflict is evident in virtual environments where navigation is controlled by some means other than walking in the real world, or where delay due to lag and frame rate is large enough that the user's movements lose synchronization with the displayed scene.

Vestibular cues come from the inner ear, from the semicircular canals and the otolith organs (Howard, 1993). The semicircular canals provide information about body rotation about each of the three orthogonal axes. The otolith organs respond primarily to linear acceleration. Vection arises when there are competing visual and/or vestibular displays. It occurs when an individual viewing a moving display perceives that he is moving rather than the display. This occurs commonly in everyday environments, such as when one views a moving train out the window and feels that one's own train is moving, etc. Vection is tied to headcentric visual motion (that is, it occurs with respect to the head's frame of reference, as opposed to the body or the eyes). Vection is totally under control of whichever display is perceived as background, and can occur with small or large stimuli or fields of view. Generally, yaw vection (about the vertical axis) is stronger than pitch vection (about the axis passing through the two ears), which is stronger than roll vection (about the visual axis).

2.4.2 Low-level effects of field of view

This section discusses how vision varies with visual angle, considering first the physiology of the eye, then eye- and head-movement considerations, then behavioral considerations.

Werner defines the *visual field* as "encompassing everything that can be seen while the eye remains fixed. Normally the full extent of the visual field (see Figure 1-1) of each eye extends approximately 60° superiorly, 75° inferiorly, 100° to the temporal side, and 60° to the nasal side." (Werner, 1991). Werner reviews the techniques, known as perimetry, used to measure people's visual field extent.

Cannon (1986) reports on differences in physiology between the fovea and the periphery. The number of cortical cells devoted to seeing a given object size decreases with eccentricity, thus there is effectively better resolution for visual functions at the fovea. Visual acuity varies with eccentricity (Barfield et al., 1995). The retina is made up of rods and cones. The cones are most dense at the fovea (the central 2°) and cone density falls off rapidly with increasing eccentricity, reaching minimum at approximately 10°. Rod density increases from the fovea outward and reaches its maximum at approximately 10°, then decreases. The effective resolution depends on the task, and is determined by rod and cone spacing and by the spacing of receptive fields and retinal ganglion cells, which pool responses from multiple rods and cones. General acuity is approximately 1 minute of arc to 30 seconds of arc (1/60° to 1/120°) at the fovea, and drops with eccentricity – to 50% (1°) at 10°, to 20% (2.5°) at 30°, and to 5% (10°) at 60°. Acuity is known to be less at lower luminance levels (Davis, 1997). FOV also varies for color perception; we do not perceive color outside 100° horizontally (Barfield et al., 1995).

Leibowitz (1986) discusses functional differences between the fovea and the periphery. Peripheral vision is most important for spatial orientation tasks such as locomotion, maintaining body posture, and gaze stability. Object motion sensitivity is optimal in the central field; it degrades with eccentricity but less severely than does object discrimination or recognition task performance. All visual functions are superior in central vision except for absolute sensitivity to light (because of the rod distribution), and perception of self-motion through vection. Vection is greater in the periphery because it is roughly proportional to the size of the stimulus and the periphery is larger than the fovea.

The eyes move in saccades. The reaction time for a saccade is approximately 150 to 300 ms, and the saccade movement can be as fast as 900°/s. Head movements can happen as fast as 500°/s. For images of a moving object to be perceived as continuous, the displacement from frame to frame needs to be less than about 0.2°. Eye movements tend to be less than 15° (Barfield et al., 1995; Davis, 1997). When tracking an object, a person moves both the eyes and the head in coordination (Gauthier et al., 1987). For movements of 5-10°, typically the eyes alone move in saccades. For movements > 10°, the eyes make a large saccade (typically about 30°), and this is followed by a slow head movement to compensate and bring the eyes back to their resting place. This stabilization is under direct control of the vestibular system and is influenced by visual and proprioceptive inputs.

2.4.3 Task-level effects of field of view

The effects of a narrow field of view have been tested in several studies, most involving the subjects wearing restricted-FOV goggles and performing the task in a real environment.

The general effects of FOV are well-described qualitatively by Dolezal (1982b), who wore field-restricting goggles for six days without taking them off at any time and recorded his observations. Among the differences he cited were: no "pan-overlap" – meaning more and smaller head movements were required to scan a scene; moving objects being visible for shorter durations, with less context information available; information about the location of one's body parts is reduced; one can only maintain fixation during slow eye movements; overall scene is darker and more variable in intensity, making higher demands on pupil adjustment and on accommodation. He reported that familiar objects appeared smaller and nearer; he had more difficulty maintaining equilibrium during locomotion; he perceived himself as shorter. He adapted to the restricted FOV over time, and, after removing the goggles, experienced after-effects – familiar objects appeared larger, he perceived himself as taller, and he underreached for objects. He distinguishes between after-effects that are perceptual only (a perception that things are different) vs. those that are performatory and resulted in carry-over behavior, such as under-reaching for objects.

The following sections discuss the effects of FOV seen for the four classes of task described in Section 2.1.1.

2.4.3.1 Search

Peripheral vision is important for searching because, whereas we do not use the periphery to recognize and identify objects, visual events in the periphery serve to trigger changes in gaze (Cannon, 1986; Leibowitz, 1986).

Cunningham et al. discuss performance of a target detection task with different fields of view and give both qualitative and quantitative analyses of head-movement behavior (Cunningham et al., 1996). They compared performance in a $60° \times 40°$ HMD with performance in a $150° \times 70°$ dome. There were three dome-screen conditions, one with the full FOV, one with a software-limited FOV, again to $60° \times 40°$ – the head was tracked and the viewport followed accordingly across the dome, and one with a mechanically-limited FOV to $60° \times 40°$. They also varied auditory conditions – present or not, and localized spatially or not. Subjects wore a helmet in all conditions, to produce consistent weight load on the head. For qualitative head-motion analysis, the researchers plotted head position over time and distinguished characteristic strategies such as "circular," "figure eight," "vertical scan," and "chaotic." In the limited FOV conditions, search paths tended to become wider, except in cases where localized audio was present. For quantitative head-motion analysis, the researchers computed average angular velocity over time or total angular displacement over time. They found that in the localized audio conditions there were reductions in both measures, and suggest that the two measures are equivalent across their viewing conditions.

Piantanida et al. (1992) describe a study conducted with a VPL EyePhone display where subjects' performance was measured on a task that required locating and identifying squares among various distractors. Piantanida estimates the VPL EyePhone's FOV at 100° total horizontal; Robinett and Rolland (1992) traced rays through the EyePhone's optics to determine precisely its FOV and found it to be 90° total horizontal. In Piantanida's study of search with restricted FOV, four FOV conditions were tested: 28°, 41°, 53°, and full-field. The targets were positioned in random positions on a large sphere centered at the subject's position. Results showed significantly higher response time for FOV of 53° or less (Piantanida et al., 1992).

In addition to searching being an important task, search behavior in virtual environments can give us information about eye and head movement coordination. This is described by the vestibulo-ocular reflex (VOR) in the real world. There is evidence that narrow-FOV HMDs will disrupt VOR behavior, but that users may be able to adapt to the change (Gauthier et al., 1987). The adaptation requires weeks rather than hours, however, so such adaptation may not be likely to provide benefits for general VR applications. Gauthier et al. measured differences in eye-head movements and coordination under different field of view conditions, viewing LEDs through field-restricting apertures (Gauthier et al., 1987). A small field of view tended to cause earlier triggering of head movements and less eye-tracking of objects indicating that the usual VOR behavior was disrupted.

Wells et al. (1990) have studied target tracking in a flight simulator environment using their VCASS head-mounted display. In a collection of studies, subjects were asked to track targets while viewing a simulated flying sequence over a terrain. In one study, a wide view of the terrain (120 degrees horizontal) was always visible, but the targets (small triangles or squares) appeared only in a window in the center of the display (I will refer to the size of that window as the stimulus FOV). The stimulus FOV size was the main variable in the experiments. The studies showed that error rates were higher with smaller stimulus FOV. Further, the smaller the stimulus FOV, the less complex the tasks that could be performed adequately (i.e., distractors were more harmful the narrower the stimulus FOV) (Wells et al., 1989; Wells and Venturino, 1990).

Stark et al. (1992) demonstrated the effect of window size on visual search behavior with both head-mounted displays and CRT displays (with a mouse being used to control a window). They found that with small search windows, search patterns were regular (eg., scanning left-to-right in a regular up-and-down pattern), but with larger window sizes, search patterns became irregular – the subjects use the wider field of view to carry out a hierarchical search, meaning they make larger scans followed by progressively smaller ones. Subjects scan and detect possible targets in the periphery and quickly foveate on the target to identify it. The head-mounted display was custom-built with LCD displays and had a field of view of approximately 22°.

2.4.3.2 Locomotion

Peripheral vision is well suited to maintaining self orientation during locomotion. Leibowitz (1986) cites the example of our ability to read while walking; we can fixate on a primary task (reading a book), while using peripheral vision to perform a secondary task (avoid obstacles while walking). Thus one might expect that a narrow field of view would detract from a person's ability to navigate effectively while avoiding obstacles. This has been shown to be true by Alfano and Michel (1990) where subjects performed a walking task while wearing field-restricting goggles. Subjects took twice as long with 9° or 14° FOV compared to with no restriction. At 22°, subjects took 1.5 times as long. Wood and Troutbeck (1994) have shown similar effects on subjects' ability to drive a car down a road in a straight line. Kenyon and Kneller (1993) have shown that subjects' ability to maintain orientation (attitude) while flying in a dome-screen flight simulator is worse with a narrow FOV.

The effects of FOV on performance when steering an unmanned aerial vehicle (UAV) were studied by de Vries and Padmos (1997). The task was to maneuver a simulated UAV through trees. To study FOV they simulated the HMD FOV by restricting it in software and projecting a moving window onto a dome screen. They also tested performance with an actual HMD (Virtual IO I-glasses). Results with the simulated HMD yielded significantly better performance than with the actual HMD (with the same FOV). Performance was significantly better with 34° or 57°

FOV compared to with 17° FOV. They also found that FOV significantly effected head movements as indicated by mean total head speed and other measures of motion. They report that one of their subjects became sick during the experiments when using the small FOV. No other subjects reported sickness symptoms (quantitative measures of simulator sickness were not taken, however). The researchers also found that display lag significantly worsened performance, and stereoscopic presentation was no better than monoscopic presentation for this task. Additional studies have measured effects of FOV on helicopter flight (Edwards et al., 1997) and with night vision goggles (McLean et al., 1997).

Péruch and Mestre describe studies of the effects of functional field of view size on navigation performance using a desktop display (Peruch and Mestre, 1999). They studied performance of a vehicle navigation and control task. Physical field of view was always 60° H × 50° V, with the participant viewing images on a monitor. *Functional* FOV varied, however, through head tracking. In the base condition ("fixed"), vehicle motion was controlled by a joystick alone. In the first "mobile vision" condition, vehicle motion changed with head movements (orientation only), in a one-to-one manner – for example, a head rotation of 30° to the left produced an apparent 30° rotation of the vehicle to the left. Note that this differs from "head-tracked stereo" in that the head tracking rotates the virtual eye rather than shifting the projection to produce a stable perspective effect. The second head-tracked condition, "mobile vision 2" was similar to the mobile vision condition but doubled the apparent motion – here a 30° head rotation would produce a 60° vehicle rotation. Thus, in the author's terms, the "functional visual field" was 60° in the fixed case, but increased to 120° in the first mobile case (because ± 30 ° additional motion was allowed), and to 180° in the second mobile case (where ± 60° motion was allowed). They found that participants were able to take smoother trajectories through the environments when they had larger functional fields of view, and made fewer stops at corners, but completion time did not decrease in the head-tracked conditions.

2.4.3.3 Judgments of distance and space

Several researchers have reported that narrow FOV tends to make objects appear nearer than they really are and to effectively shrink the environment around the user (Dolezal, 1982a; Alfano and Michel, 1990; Henry and Furness, 1993). Hagen, Jones, and Reed also showed that narrow FOV, both in pictures and in real-world views "causes a frontal shift in the localization of the visible field, with a resultant compression of perceived size and distance" (Hagen et al., 1978). In a monocular task, when looking through a "peephole" view (2 mm in diameter) or a "truncated"

FOV (through a 6.5 by 4.5 cm hole placed 5 cm in front of the eye), subjects underestimated distances and sizes compared to looking monocularly with no field restriction.

In addition to FOV effects, researchers have shown that other factors may result in misperceived distances, in particular in see-through head-mounted displays (Ellis and Bucher, 1994). Davis (1997) reports that while both Roscoe (1993), and Rolland et al. (1995) have shown that, through a see-through HMD, virtual objects appear farther away than their real-world counterparts (by a factor of approximately 1.25 according to Roscoe), other researchers claim the opposite – that virtual objects appear closer than their real counterparts.

2.4.3.4 Manipulation (reaching and grasping)

If distance from oneself to an object is underestimated with restricted FOV, then it follows that when reaching for that object one will under-reach. There is evidence, from real-world studies, to show this effect for reaching, whereas grasping ability is not significantly affected by field of view (Davids and Stratford, 1989; Sivak and MacKenzie, 1992). Sivak and Mackenzie (1992) have shown that restricted field of view affects the reaching phase. With a narrow FOV the distance to objects was frequently misjudged. Davids and Stratford (1989) studied the ability of subjects to reach with their hand and catch a ball unseen. Subjects were blocked from seeing the actual catch by a screening mask attached to their helmet. Again, the ability of the subjects to position the catching hand to the correct place (to reach) with a 70° FOV was compromised, although the catching or grasping action was not affected (Davids and Stratford, 1989).

Chapter 3
Methods

The review in Chapter 2 discussed previous findings on the effects of field of view and described four classes of tasks to study in virtual reality. This chapter describes the methods used in the current work to address the following areas.

- Controlled studies of FOV across multiple tasks in VR. Previous studies of FOV in VR have tested tasks individually, either search tasks or travel-by-flying tasks. The present study measures performance on multiple tasks under the same viewing conditions. This allows for comparison of the impact of FOV on different tasks.

- Studies using wide-FOV head-mounted displays. Previous studies of FOV in HMDs for VR have tested FOV under 120° or less (see Chapter 1). The present study used the Kaiser tiled HMD to test a radically wider FOV, up to 176°.

- Studies linking performance with sickness in VR. Previous studies of FOV in VR have dealt with performance only or sickness only. In the present study both are addressed at the same time.

The following sections describe the research questions, tasks, viewing conditions, and apparatus used.

3.1 Research questions

Human subject experiments were undertaken to test effects of field of view on a representative set of tasks. The following two are the primary research questions:

Q1. Task performance (task completion time): How does time to search for objects vary with FOV? How does time to walk through a simple maze-like environment vary with FOV?

Q2. Sickness effects: How does the incidence of VR sickness vary with FOV? How is postural stability affected by FOV?

The following questions were evaluated as secondary and exploratory research questions:

E1. How are judgments of distances to objects affected by FOV?

E2. How are judgments of spatial layout in the virtual environment affected by FOV?

E3. How is the participant's sense of presence in VR affected by FOV?

E4. How is head movement affected by FOV?

E5. Do the findings on FOV generalize to other HMDs and compare well to real-world viewing?

From these research questions, hypotheses and tasks were formed and data were collected, as described in Section 3.2, and these hypotheses were tested statistically (see Chapters 4 and 5). The following sections describe the tasks, other general experiment methods, and the experiment apparatus.

3.2 Tasks

The experiment tasks were chosen to be similar to common physical actions performed in virtual environments, and to correspond to the four classes of tasks described in Chapter 2. To establish their coverage more systematically, I describe them here with direct reference to the frames-of-reference system described by Howard (1993). Frames may be classified first as either egocentric or exocentric, as follows.

Egocentric frames are those defined with respect to some part of the observer, e.g., headcentric, bodycentric, and handcentric frames. Additional egocentric frames are the retinocentric and station-point frames, which correspond primarily to actions made through eye movements or fixation. In this study I did not include retinocentric, station-point or handcentric tasks. I ignored the retinocentric and station-point frames because our head-mounted displays do not allow for eye tracking. Handcentric tasks are common to virtual reality, but it is likely that handcentric performance is not directly affected by FOV; it is affected by FOV only indirectly, through errors in judgments of distance and space.

Exocentric frames are those defined with respect to points external to the observer. An example of an exocentric judgment is to report an absolute distance between two points in a room. A similar but egocentric judgment would be to report the distance between a point in a room and your own body position (this would be a bodycentric judgment).

The four primary tasks used in this study correspond to frames of reference as follows:

- An egocentric (headcentric) *search* task, where the subject is asked to turn his head to point at a target. This task, described below, is similar to tasks used in other VR studies (Piantanida et al., 1992; Watson et al., 1997).

- An egocentric (bodycentric) *walking* task, where the subject is asked to walk to a target at the opposite side of an environment while avoiding walls. This task is inspired by the walk-along-a-path task used by Alfano et al. (1990), the fly-through-maze task used by Ware (1990), and the fly-through-tunnel task used by Jorgensen (1997).

- An egocentric (bodycentric) *distance estimation* task, where the subject is asked to walk to the position of an object seen previously in the virtual environment but now unseen. This task was done in the same virtual environment as the walking task. The subject is asked to walk through the environment as before, noting the position of a target, then afterwards walk back directly to the target, which is now unseen. This distance judgment task is suggested by the distance task from Hagen et al. (1978).

- An exocentric *spatial memory* task, where the subject is asked to remember the positions of multiple objects previously seen throughout the virtual environment and afterwards report their positions on a 2D map layout. This was done in the same environment as for the walking task. This task is inspired by the spatial memory tasks used by researchers in real-world restricted-FOV experiments (Hagen et al., 1978; Alfano and Michel, 1990).

An originally proposed task, reaching, was later dropped from the study. I decided that the effects for reaching (underestimated distance) could be obtained through the distance estimation task described above. In addition, the low resolution and brightness artifacts of the Kaiser HMD would have seriously compromised results for a small-scale task such as reaching (where the 3D region of interest is small).

The following secondary tasks were also part of the experiment sessions. Subjects were told informally how to perform these tasks, which were only auxiliary to the main experiment.

- A *postural stability* task measures the participant's instability (or ataxia) at the start and end of each HMD session. I used the "heel-to-toe with arms crossed" position referred to by Kennedy et al. (1995). After the display is calibrated, the experimenter asks the participant to stand in this heel-to-toe position. A visual reminder to begin this task is also provided on the screen (a yellow circle with the word “Stand”). When the participant is standing properly the experimenter presses a button and the screen blanks; the participant continues standing for 20 seconds, at which time the experimenter again presses the button to end the postural stability test.

- *Display calibration*: At the beginning of each session with the Kaiser HMD a simple calibration task ensures that the display is fitted properly. The system displays a calibration object that consists of a horizontal bar, a circle, and crisscrossing diagonal lines. The calibration object arises from the tiled nature of the Kaiser display, and is similar to calibration software provided from the manufacturer. The purpose is to adjust the display so that the person’s eyes are in the correct position to see a single seamless image across all tiles. Figure 3-1 shows the image as drawn into the frame buffer. When calibrated properly, the circles and lines overlap as described below. The procedure (spoken to the subject and demonstrated) is:

  1. *Grasping the sides of the helmet, move the helmet up or down until the white horizontal bar is in the center and stretches all the way across.* (Aligns top and bottom rows of tiles.)

  2. *Turn the IPD dial at the front to move the left and right sides in or out, until the circles overlap.* (Aligns inside columns for correct stereo.)

  3. *Turn the dial at the back of the helmet to bring the screens towards your eyes until the grid of diagonal lines meets up all the way across.* (Aligns middle and outside columns.)

  Before they used the Kaiser HMD the first time, subjects were shown the calibration grid on a workstation monitor, to see what it looked like when properly adjusted. All subjects were able to achieve good calibration of the display. Typically this took five minutes the first time a participant wore the display and only one or two minutes in later sessions.

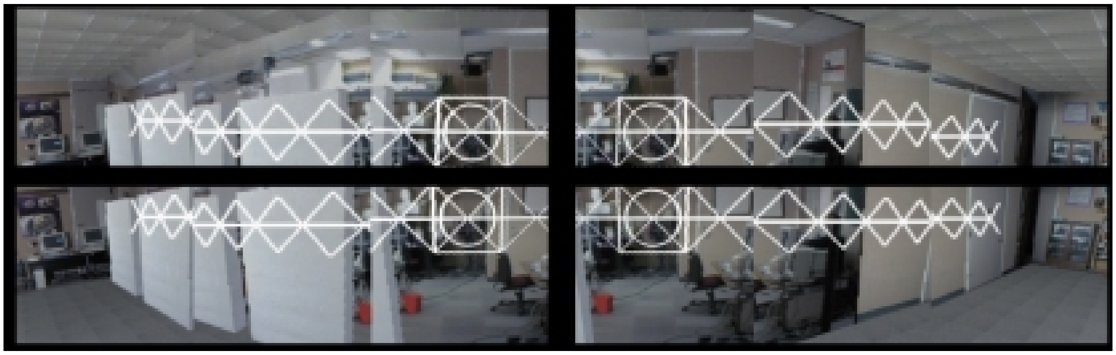

Figure 3-1. Calibration grid and background scene drawn for Kaiser HMD.

The participants also filled out the simulator sickness and presence questionnaires. Participants were given copies of these questionnaires and a general information sheet (Appendix B.2-B.6) during their initial screening sessions and were asked to familiarize themselves with them before their first experiment session.

The tables below summarize the tasks and variables for the primary and secondary tasks, and their relation to the research questions.

Table 3-1. Summary of primary tasks.

	Independent variables	Dependent variables	Controlled variables
Search	FOV Target angle	Search time (Q1) Head movements (E4)	Target misses minimized
Walk	FOV	Walk time (Q1) Head movements (E4)	Collisions with walls minimized
Distance	FOV Target position	Distance error (E1)	Collisions with walls minimized, Time not a factor
Spatial memory	FOV Object positions	Position error (E2)	Collisions with walls minimized, time not a factor

Table 3-2. Summary of secondary tasks.

	Independent variables	Dependent variables	Controlled variables
Sickness Questionnaire	FOV Exposure number	Total severity, sickness subscores (Q2)	
Presence Questionnaire	FOV	Total presence (E3)	
Postural Stability	FOV Exposure number	Head movements – mean velocity (Q2)	
Calibration	N/A	N/A	

3.2.1 Search

The search task measures the time it takes for a subject to turn his or her head to locate visually an object that is initially outside the FOV. The primary hypothesis, **H1**, is that search time is shorter with wider FOV. This hypothesis is suggested by previous findings discussed in Section 2.4.3.1, e.g. (Piantanida et al., 1992).

Task and stimuli. The scene for the search task is diagrammed in Figure 3-3. The subject is situated inside a background environment that is a model of the lab where the experiment actually took place (Figure 3-2 shows the background environment, which is discussed further in Section 3.5.3.). In the virtual environment, a green circle on the floor indicates the position where the subject stands for this task (not shown in the figure). The subject is instructed to move by turning only, and not to step away from this spot.

Figure 3-2. View from inside the background virtual environment.

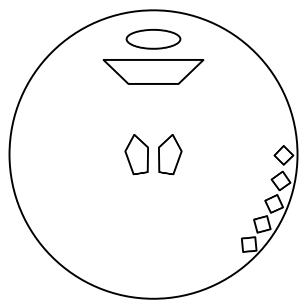

Figure 3-3. Search task: diagram of scene.

Two important 3D objects appear in the scene at different times: the “target,” a green double-pyramid (8 sides) with radius 10 cm (the distance from the center to each corner) (Figure 3-5), and the “Start” symbol (Figure 3-4). The task is to go to the start position, click a button, then turn to your right until you get to the target position, click the button again, then return to the start position and repeat. The subject is “at” the target or home position when the object is within the white square (the cursor) that is fixed 1 m in front of the user. The cursor moves with the user, like a viewfinder on a camera. The start symbol appears 1 m forward of the subject at the starting position (angle 0), as shown in Figure 3-2. It is placed at a height equal to 0.8 times the subject’s height. The target position is also at height equal to 0.8 times the subject’s height, and also 1 m from the subject’s position (i.e., the targets and the start symbol are constrained to a circle with radius 1 m, centered about the subject). The angle the target appears at varies among five angles, 90, 100, 110, 120, and 130 degrees to the right of the home position. The target angle is chosen randomly from these five without replacement. That is, for trials one through five, the sequence is a random permutation of the five angles. For trials six through 10 it is another random permutation, and so on throughout. The target and start objects were smooth-shaded (modeled using 3D Studio Max and radiosity-lit using Lightscape).

A dark background region (a portion of a cylinder) is placed in the scene to reduce distraction from elements of the background environment. This is the black band in Figures 3-4 and 3-5 and the circle in the top view shown in Figure 3-3.

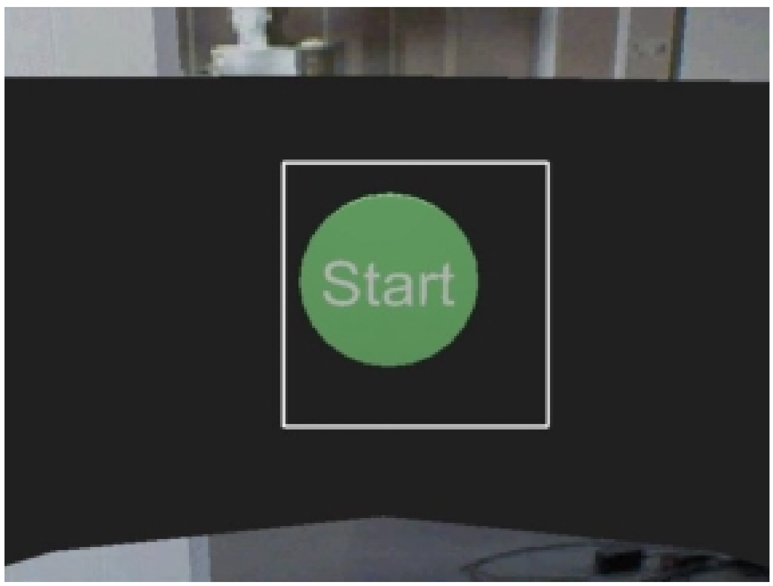

Figure 3-4. Search task: HMD view from home position.

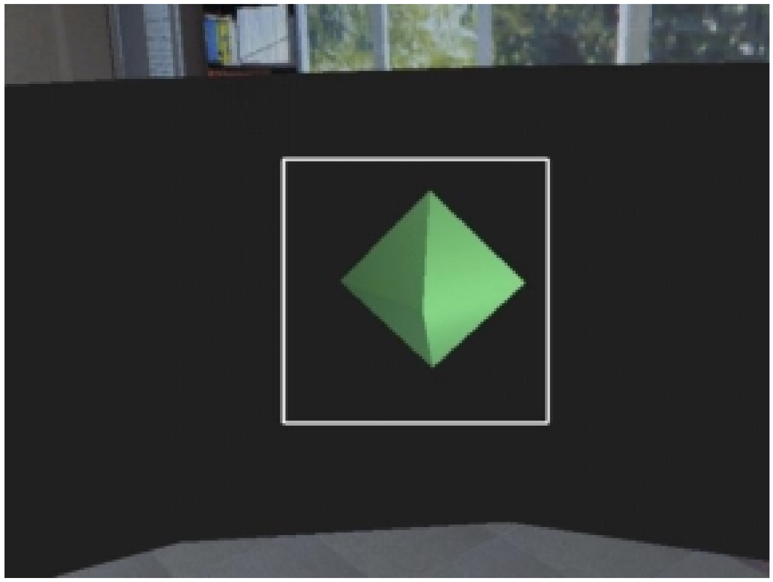

Figure 3-5. Search task: HMD view from target position.

For a given display condition, the subject performs the search task 80 times: 4 blocks of 20 trials each. Data were recorded for all blocks, but the first block was considered practice and only the last three were used for analysis.

Variables. The independent variables for the search task are target angle (90 through 130), and FOV (described below in Section 3.3). The dependent variable is search time, recorded as the time elapsed between the button clicks at the start position and at the target position.

3.2.2 Walking

The walking task measures a subject's ability to travel through an environment while avoiding obstacles. In this case the travel is by real walking. That is, physically walking through the real environment of the lab results in corresponding travel through the virtual environment. The primary hypothesis, **H2**, is that walking time will be shorter with wider FOV. This hypothesis is based on earlier assertions that larger FOV allows subjects to move faster because they can see more of the environment or because more peripheral vision gives them a better sense of their movement and of their own position and orientation in the environment (Alfano and Michel, 1990; Sampson, 1993).

Task and stimuli. A top view of the virtual environment used for this task is shown in Figure 3-6. The background portion of the environment is the same as in the other tasks; the background is a model of the actual lab where the experiments took place. Several virtual walls have been added to the model. These are models of 6-foot high white styrofoam blocks. The subject's task is to walk from one side of the environment to the opposite side at a quick, comfortable pace, while trying not to touch the virtual walls. The start- and end-points are marked on the floor by red and blue circles. (The smaller, green circle in Figure 3-6 indicates a subject's position in a replay of the trial. This was not present during the actual experiment.) Subjects start at the red circle, click a button, then walk to the blue circle, and click the button again. This constitutes one trial. The subject then walks back – starts at the blue circle, clicks the button, walks to the red circle, and clicks the button to finish. The time between starting and ending button clicks is recorded as the walking time to be used for analysis. Subjects can pause between trials to rest. A rest pause is explicitly encouraged between each block of 10 trials, when a yellow circle with the word "Break" appears in front of the subject. Each subject performed three blocks of ten trials for each display condition. Each of these was one VR exposure and occurred on a separate day. The first block of ten from each condition was dropped as practice.

During the experiment, collisions with the virtual walls were detected and recorded. When the subject hits a virtual wall the screen turns reddish to indicate the collision. Collision detection was performed in the software by testing an axis-aligned bounding box, representing the subject's virtual body, against all virtual walls in the environment. When this test indicated a collision, a large semi-transparent red polygon was drawn in front of the subject's eyes. This provided an effective visual cue that did not obscure the original scene because it could still be seen through the red filter. Our system did not provide audio signals to the HMD; otherwise an audio cue would have been considered as an alternative to this visual cue.

The walking task was also performed in a "restricted-real" condition, in which the subject viewed the actual lab environment through a narrow FOV, and had to avoid *real* styrofoam walls that matched the positions of the virtual walls. This was performed using the restricted-real HMD described in Section 3.3.

Variables. The independent variable for the walking task is FOV. The dependent variable is walking time.

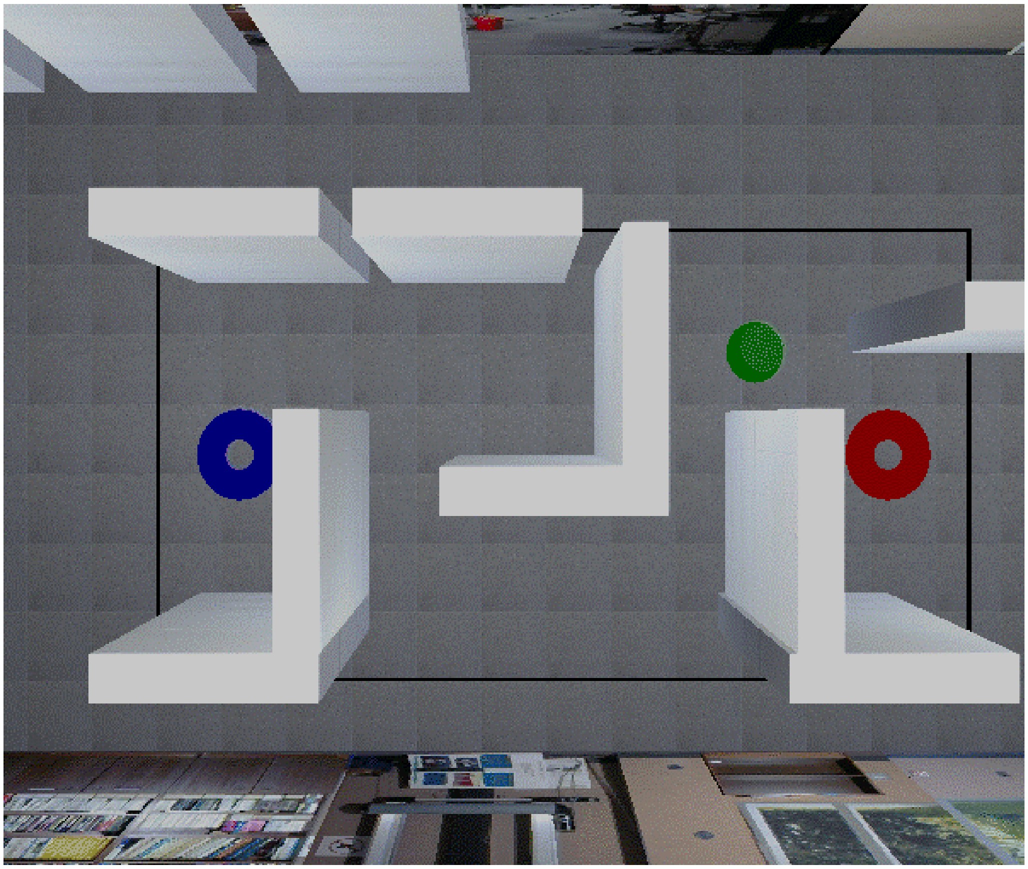

Figure 3-6. Walking task: top view of scene.

3.2.3 Distance estimation

The distance estimation task measures the ability of a subject to perceive distances between places in a virtual environment and to recall those distances while inside the virtual environment. This task is inspired by previous real-world restricted-FOV work by Hagen (1978) and Dolezal (1982b). The hypothesis, as suggested by that work, is that restricted FOV will result in underestimation of the distance to objects.

The distance estimation task was administered using the same VR exposure as the spatial memory task, which is described in Section 3.2.4. The key difference between the two tasks is that the distance estimation task is egocentric whereas the spatial memory task is exocentric.

Task and stimuli. A top view of the virtual environment for the distance estimation and spatial memory tasks is shown in Figure 3-7. The background environment is the same as for the walking task, except that more circles have been placed on the floor in the virtual environment. The subject walks across the room from the red circle to the blue circle, as in the walking task, but in this case time is not important. The subject is told to walk at a comfortable pace and that, unlike in the walking task, time is not a factor.

The key objects in the scene for the estimation and memory tasks are the small colored circles. There are six small colored circles, 10 cm in diameter, placed on the floor in the virtual environment (shown in Figure 3-7, compare with Figure 3-6). The white circle is used for the distance estimation task. The five other circles are used for the spatial memory task, which is described in the next section.

The distance estimation task proceeds as follows. The subject starts at the red circle, clicks the button, walks to the blue circle, and clicks the button again. Then, starting at the blue circle, the subject clicks the button, walks to the red circle, and clicks the button again. The task so far is the same as for the walking task except that the subject is instructed to take their time and to take note of the position of the white circle that they see on the floor. The next part of the task is the recall portion. After the subject returns to the red circle, the virtual environment changes automatically from the standard environment with the background lab scene and virtual walls to a blank environment that consists of only a gray ground-plane and a black background. The subject is instructed to walk directly to the place where they saw the white circle. As before, the subject pushes a button before walking and pushes it again when they reach the estimated white circle position. It was emphasized that the subject should walk *directly* to the white circle position and not along the path they had previously taken. Their previous path would not have been a straight line because the subject had to avoid the virtual walls. In the recall portion of the task the walls are not present. At this point, one measure of distance estimation has been gathered. The subject now walks back to the red circle position, which is now shown in the blank environment. The entire sequence then repeats for a new white circle position. The position was chosen randomly without replacement from 5 possible positions in the virtual environment that were predetermined.

Two blocks of 5 trials each were performed for each subject and viewing condition.

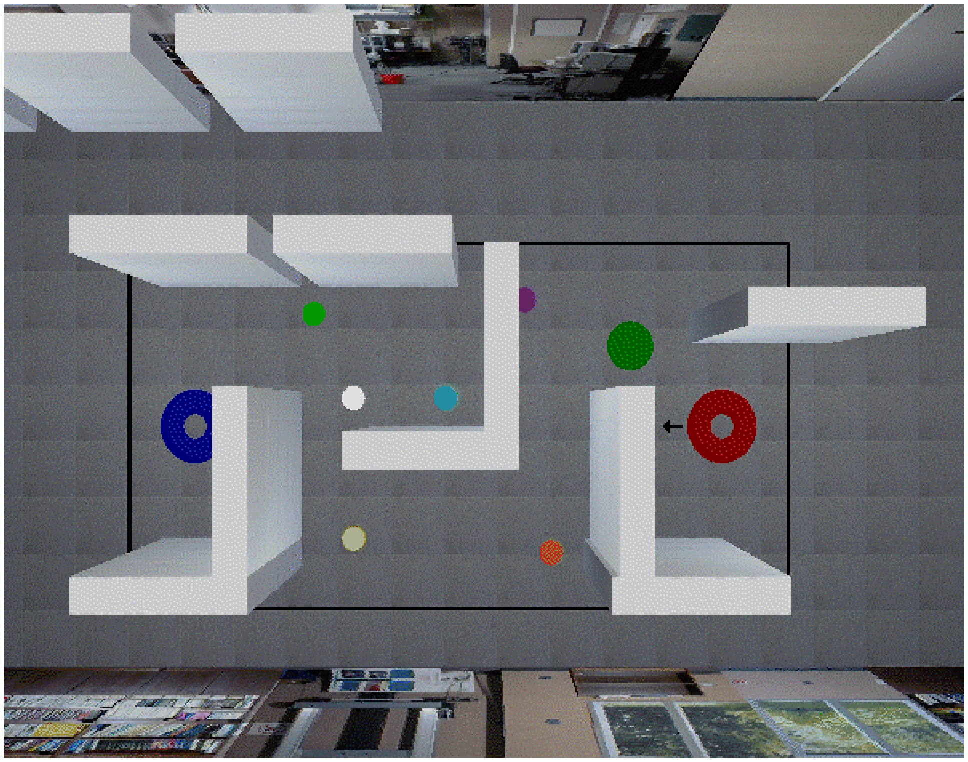

Figure 3-7. Distance estimation and spatial memory tasks: top view of scene.

Variables. The main independent variable for the distance estimation task is FOV. The position of the white circle is also an independent variable and was considered in the analysis. The dependent variable is distance accuracy, taken as (estimated distance) / (actual distance), where distance is the distance from the red circle to the white circle.

3.2.4 Spatial memory

The spatial memory task measures a subject's ability to perceive the spatial arrangement of objects in a virtual environment and to recall that arrangement afterwards on a 2D map, outside the virtual environment. The hypothesis is that with a restricted FOV in the virtual environment, subjects will make more errors afterwards in recalling the object positions. Further, the objects will be positioned closer together than they actually were. These hypotheses are suggested by previous real-world restricted FOV studies (Hagen et al., 1978; Alfano and Michel, 1990).

The spatial memory task was administered in the same VR exposures as the distance memory task.

Task and stimuli. A top view of the virtual environment for the spatial memory task is shown in Figure 3-7. As described in Section 3.2.3, the environment is the same as for the walking task except that 6 small circles have been added. The white circle is used only for the distance estimation task. The remaining five circles (green, yellow, cyan, purple, orange) are used for the spatial memory task as follows. There are five positions on the floor where the circles appear. Only the color of the five circles changes, not their position. For each administration of the task (i.e., for each FOV and subject), an arrangement of the five colors is chosen randomly, and this arrangement remains fixed for the duration of the session. The subject performs 10 trials of the distance estimation task, as described in Section 3.2.3.

After completing those 10 trials, the subject removes the VR headset and performs the recall portion of the spatial memory task. The subject was instructed at the beginning of the experiment to remember the positions of the five colored circles in order to recall them afterwards on a 2D map. The recall is done on a screen using a mouse (see Figure 3-8). The subject clicks and drags the circles from the left side of the screen across to the position where they remember seeing the circles in the environment. There is no time constraint; the subject clicks on “done” when finished.

Variables. The independent variable for the spatial memory task is FOV. The dependent variable is the recall accuracy, taken as the distance between the estimated and actual positions of each colored circle. Thus there were five measures of recall accuracy per FOV per subject.

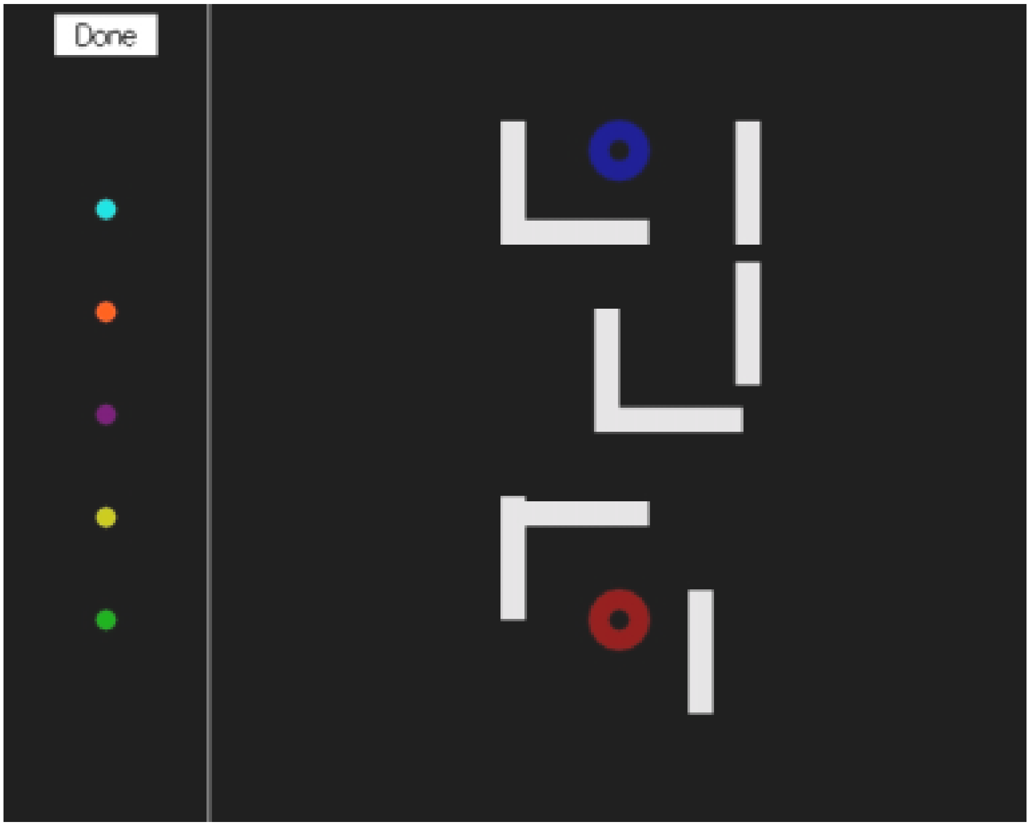

Figure 3-8. Spatial memory task: screen capture of recall portion of the task.

3.3 Viewing conditions

The primary independent variable for the studies was FOV. There were three primary viewing conditions corresponding to three values of total horizontal FOV size in the Kaiser HMD. Table 3-3 lists the three primary viewing conditions.

Table 3-3. Primary viewing conditions.

	Horizontal FOV	Vertical FOV	Horizontal overlap
Kaiser low	48°	47°	48°
Kaiser medium	112°	47°	48°
Kaiser high	176°	47°	48°

The difference between the three primary viewing conditions is that the horizontal FOV varies. Vertical FOV is constant, as is the size of the horizontal overlap region for stereo vision. The Kaiser HMD contains six tiled screens per eye (three columns of two tiles). Figures 2-2 and 2-3 show the Kaiser HMD. The three values of horizontal FOV correspond to the columns of tiles as follows. In the 48° condition, only the inner columns were used (two tiles per eye). In the 112° condition, the inner two columns were used (four tiles per eye). In the 176° condition all three columns were used (six tiles per eye). The software FOV was computed using the calibration software provided by Kaiser Electro-Optics.

For comparison, some tasks were replicated under the two secondary display conditions listed in Table 3-4. The first of these used the Virtual Research V8 HMD, which has a 48° by 36° FOV. The second comparison condition used a mock HMD that provides a restricted-FOV real-world view. This condition will be referred to as "restricted-real."

Table 3-4. Secondary viewing conditions.

	Horizontal FOV	Vertical FOV	Horizontal overlap
V8 low	48°	36°	48°
Restricted-Real	48°	36°	48°

The table below summarizes the different viewing conditions and their use for the different tasks. All three HMDs are described further in Section 3.5.2.

Table 3-5. Tasks and viewing conditions.

	Kaiser low 48°×47°	Kaiser medium 112°×47°	Kaiser high 176°×47°	V8 low 48°×36°	Restricted -Real 48°×36°
Search	✓	✓	✓	✓	
Walking	✓	✓	✓	✓	✓
Distance estimation	✓	✓	✓	✓	
Spatial memory	✓	✓	✓	✓	

3.4 General methods

Experiments were conducted in the Head-Mounted Display Laboratory at the UNC Chapel Hill Department of Computer Science. The following sections describe the selection and scheduling of subjects and the methods used to implement the experiments and collect measures of human task performance.

3.4.1 Scheduling and subjects.

For pilot experiments, two subjects who were members of computer graphics projects in the UNC-CH CS department performed all tasks under selected conditions. This gave initial data that I used to debug and improve the tasks.

Five subjects were recruited for the full experiments. Each of these subjects visited the lab six times. Their first visit was an information session only. During this visit I gave the subject a consent form and a written description of what they would be doing, screened the subject for stereo vision and color blindness, and showed the subject the VR equipment.

To screen for color blindness I had subjects view reproductions of two color plates from a standard color blindness test (Birch, 1993) and tell me what they saw. The correct response was a number that would appear in the figure. All subjects passed this test. To screen subjects for correct stereo vision, I showed them a stereo test image on a screen using Stereographics CrystalEyes shutter glasses. The images consisted of two squares presented with binocular disparity only and no perspective. The squares are drawn separately for the left eye and the right eye. On the top half of the screen a square is presented with positive disparity (i.e., the square is shifted to the right in the right-eye image); it should appear behind the screen surface. On the bottom half of the screen a square is presented with negative disparity (the square is shifted to the left in the right-eye image); it should appear in front of the screen surface. The subject wears the stereo glasses and is asked to describe what they see. A correct response is, for example, "the top square looks like it's farther away and the bottom square looks like it's closer." All subjects passed the stereo test.

After the screening tests, I showed subjects both the Kaiser and the V8 HMDs and had them to put on the Kaiser HMD briefly to view an environment. The purpose of this initial fitting was to ensure that the subject was comfortable with the display and that it could be properly fitted for them to see an accurate image. The virtual environment they saw for this test was the same background virtual environment that was used in the experiments.

The final step in the initial screening visit was to give each subject a copy of the simulator sickness questionnaire (SSQ) along with a brief description of its terms, the presence questionnaire, the consent form, and a 2-page information sheet describing the experiment. These documents are reproduced in Appendix B. I wanted to make the subject familiar with all terms they would encounter in the questionnaires before they filled them out. The subject was asked to read these information sheets before their next visit to the lab. On the next visit I briefly discussed these documents again to ensure that the subject had read and understood them, and had the subject sign the consent form.

The next five experiment sessions each lasted approximately 75-90 minutes for each subject. Each session took place on a different day. In the first three of these sessions, the subject performed experiments using the Kaiser HMD. On each day they performed the search task first, followed by the walking task, followed by the distance estimation and spatial memory tasks. These three sessions with the Kaiser HMD corresponded to the three Kaiser viewing conditions (Table 3-3). On the fourth day the subject performed the same sequence of tasks but this time using the Virtual Research V8 display. On the fifth day the subject wore the restricted-real display and performed the walking task twice. I scheduled the sessions with the Kaiser display to occur before the sessions with the V8 display because switching between the displays more frequently was impractical (it requires approximately 20 minutes to reconnect cables and tracker equipment, and adds wear to the Kaiser display, which is fragile). The restricted-real condition happened last because it required real styrofoam walls to be placed throughout the lab, and placing those more than once was impractical.

Approval for the studies was obtained from the UNC Academic Affairs Institutional Review Board in March 1999. A call for subjects was posted both electronically, on newsgroups, and on paper, on bulletin boards on campus. The advertisement called for people 18 years of age or older with good vision without eyeglasses (because the Kaiser HMD cannot accommodate eyeglasses). Payment was $8/hour. Subjects who responded to the call were invited to the lab for the screening session. Six subjects passed the screening test and were scheduled to begin the experiments. One subject cancelled before the next session and was dropped from the study, leaving five.

3.4.2 General measures of performance on primary tasks

Typically in human performance experiments of this type, error rate and response time are related and either one can be isolated only by controlling the other. For the search and walking tasks, task completion time was the primary measure; I controlled error rates by instructing the subject to make as few errors as possible. I removed the first block of trials as practice.

For the distance estimation and spatial memory tasks, time was not a factor. Subjects were told to take as much time as they needed. The primary dependent variables for these tasks were measures of recall accuracy.

3.4.3 Measures of VR sickness

The third main hypothesis of the study, **H3**, was that VR sickness is greater with wider FOV. For this main test, VR sickness was taken as the Total Severity score from the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993). At several points during the experiment sessions, subjects filled out a version of the questionnaire. The questionnaire is reproduced in Appendix B.4. This questionnaire was administered on-line in hypertext format using Netscape Navigator. In one experiment session for one subject, the form was inaccessible because of network problems, and so it was administered on paper. The questionnaire is a list of 16 symptoms. Each question's response gives a number between 0 and 3 ("none" through "severe"). The "Total Severity" score is obtained as a linear combination of responses to all questions. Three diagnostic subscale scores are obtained similarly, incorporating smaller sets of questions. The three subscales were named by Kennedy et al. as "Nausea," "Oculomotor discomfort," and "Disorientation." The calculations are given in Appendix B.4.

Another health questionnaire (Appendix B.3) was administered once at the start of each session to verify that the participant was in adequate health to begin, and was not ill or under the influence of medication or alcohol. In all cases subjects were in adequate health.

3.4.4 Measures of presence

To measure presence, I used a subset of the Witmer and Singer Presence Questionnaire (Witmer and Singer, 1998). The questionnaire and scoring instructions are reproduced in Appendix B.6. The Witmer and Singer questionnaire was chosen based on the recommendations of Stanney et al. (Stanney and Salvendy, 1998). Others have indicated its pitfalls (Slater, 1999) and have suggested alternative methods for measuring presence subjectively through questionnaires or objectively through means such as counting "breaks in presence." Dinh et al. (1999) describe an alternate presence questionnaire with 14 questions including an "overall presence" question, on which the participant simply responds on a scale of 0 to 100.

The Witmer & Singer questionnaire (version 3.0 dated November 1994, obtained from its authors on February 17, 1999) includes 32 questions, each scored on a seven-point scale. Scoring the entire questionnaire is done by adding scores from selected questions, as indicated by the authors: the "Total Presence" score involves most of the questions; several subscales involve smaller sets of questions. These subscales are named by the authors "Involved/Control," "Natural," "Interface Quality," "Auditory," "Haptic," and "Resolution." For the present FOV studies, no auditory or haptic display was given in the VE system and so I dropped those questions that contributed only to the Auditory or Haptic subscales.

Two questions were added to the questionnaire. Question 33 asks the participant about their perception of the relative sizes of the virtual and real environments. Question 34 asks for any additional comments the participant wishes to make. The questionnaire was administered on paper.

The presence questionnaire was administered once per session (once per display condition), after all three VR exposures. The participant was asked to respond "with regard to the virtual environments that you experienced *today*." This is a slight modification of the standard instruction given in the Witmer and Singer questionnaire, which assumes only one VR exposure.

3.4.5 Measures of postural instability

The fourth main hypothesis of the study, **H4**, was that postural instability is greater with wider FOV. Tracker records were logged during the experiment to allow for analyzing head motion for postural instability and for playing back the experiments. This was done using the logging facility provided with the Virtual Reality Peripheral Network software library (VRPN, 1999). All tracker records and button press records are saved with time stamps; the application can later run in "replay" mode to repeat the experiment for viewing or analysis.

For the postural instability test I computed average head movement velocity (translational). I also computed average translation velocity along the three axes and average angular velocity about the three Euler angles. The head-movement analysis software was implemented by Zac Kohn.

3.5 Virtual reality equipment

3.5.1 Image generation

The experiments were run on one R12000 CPU of an SGI Onyx2 Reality Monster using one InfiniteReality2 graphics pipeline. The display software consisted of a custom application and a modified version of the UNC-CH Vlib library. Vlib was customized to perform special operations for the experiments, such as collision detection and drawing the cursor for the search task, and to display scenes in the Kaiser HMD. The Kaiser HMD code was optimized using the "projective textures" feature of OpenGL, using the multisurface-display optimization method described by Raskar et al. (1998), implemented by Wolfgang Stürzlinger.

The frame rate throughout the experiments was fixed at 30 Hz stereo. That is, all scenes could be drawn with a frame rate of at least 30 Hz, and I kept the frame rate from increasing beyond 30 Hz by checking the clock during each frame update and delaying the application if necessary. Double-buffering was used, and total display latency is estimated approximately to be 100 ms for the V8 HMD, and 116 ms for the Kaiser HMD.

3.5.2 Displays

All tasks were performed using the Kaiser 12-tile FIHMD ("full immersion head-mounted display"). To compare with performance in a commercial display with better visual quality, the tasks were also performed in a Virtual Research V8 HMD. To test how the findings compare with performance in the real world with a restricted FOV, the walking task was performed in the restricted-real condition using a mock HMD.

Kaiser HMD. The Kaiser HMD is shown in Figures 2-2 and 2-3. The total field of view of the Kaiser display, as calibrated for the experiments, is 48° vertical by 176° horizontal. The stereo overlap region extends over approximately 48° horizontal. For each eye, the display has six tiled LCD screens, arranged in three columns of two. Each screen has approximately 254 by 227 color pixels (the numbers vary slightly from tile to tile). This gives an average resolution of approximately 11 minutes of arc. Normal human visual acuity is 1 to one-half minutes of arc.

To create images for the Kaiser HMD, the software draws into an 800×600 window for each eye. This signal is scan converted by the Kaiser electronics into the individual tiles. This process adds an extra field-time of latency (~16.7 ms). Images rendered for the Kaiser HMD are shown in Figure 3-9. The 12 sub-images each have a projection and viewport extent that was determined using a calibration program provided by Kaiser Electro-Optics. The calibration involves manually adjusting several calibration images until they align and overlap well at the tile boundaries. Those adjustments determine the software parameters for viewports, projections, and FOV. When the Kaiser HMD is calibrated and adjusted properly, the user sees a continuous image across both eyes.

The stereo overlap region in the display is contained within the inner two tiles. For the restricted FOV conditions, the outer tiles of the Kaiser display were blacked out (that portion of the viewport was rendered as black).

To ensure that subjects had the Kaiser HMD properly adjusted I had them view a simple calibration scene before beginning. The scene was the background environment described earlier, with a calibration grid added. The grid consisted of a horizontal bar to indicate up-down alignment, a circle directly in front to indicate left-right stereo alignment, and a grid of crossing diagonal lines extending out in front of the eyes to indicate alignment of the tile columns. All subjects were able to adjust the display to achieve good, though not perfect calibration. Subjects did not see a perfect seamless image because the tile boundaries were still somewhat visible, and because there are differences in brightness from tile to tile.

The Kaiser HMD with the HiBall tracker sensor mounted on it weighs 5.0 pounds. It has high moment of inertia, though it is relatively well-balanced.

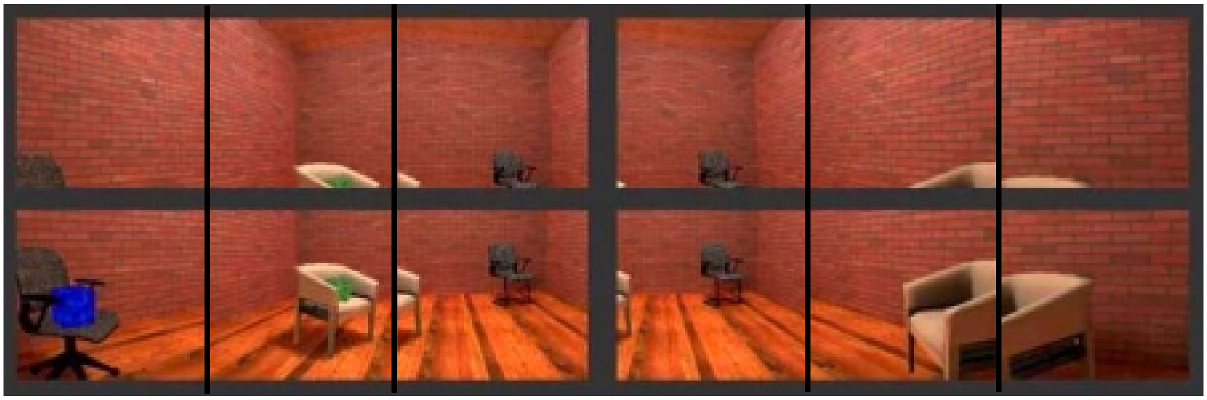

Figure 3-9. Images for Kaiser HMD.

V8 display. To evaluate the impact of artifacts specific to the Kaiser display (such as the boundaries between tiles), the primary tasks were also run using a commercially available HMD that does not suffer from these artifacts. The Virtual Research V8 display has a FOV of 48° horizontal by 36° vertical, and displays 640 by 480 color pixels. The stereo overlap is 100%. Its average resolution is 4.5 minutes of arc. The V8 with the HiBall tracker mounted on it weighs 2.5 pounds. The HiBall tracker was mounted on the front of the HMD, so the center of gravity is shifted forward.

Restricted-real display. The walking task was performed in a real world setting with the subject wearing a mock-up HMD that mimics the FOV of the V8 HMD (48° × 36°). The restricted-real display is shown in Figure 3-10. The subject walked through a physical reconstruction of the walking scene, consisting of walls made of styrofoam blocks (Figure 3-11). This real-world comparison gave information useful for assessing the severity of artifacts that occur in both the Kaiser and V8 displays, including latency and low resolution. The weight of the restricted-real display with the HiBall tracker mounted on it was approximately 1 pound.

Figure 3-10. Mock HMD for restricted-real condition.

Figure 3-11. Photograph of scene for restricted-real condition.

Verifying FOV. To verify the actual physical FOV of the V8 display I viewed an object of known size at a known distance in the real world and compared it with a similar virtual object. I viewed a plank with a precise checkerboard calibration pattern, and displayed similarly a black-and-white checkerboard pattern (10 squares across) in the display (a 2D image of alternating squares). I wore the HMD to view the virtual checkerboard, and repeatedly lifted the HMD to view the real checkerboard. I adjusted my distance until the boundary of 10 squares matched (i.e., the endpoints matched in their location in both the image and the real scene). Then, knowing the physical extent of the 10 squares on the plank, and measuring my distance from the plank I was able to compute the FOV seen in the display as the arc-tangent. FOV as computed was the same as that specified by the manufacturer to within the error tolerance for this method (approximately 5°). For the Kaiser display I relied on FOV as computed through Kaiser Electro-Optic's calibration program using computer generated images. Viewing the real checkerboard pattern with the Kaiser HMD was impractical because the Kaiser cannot be taken off and on quickly while maintaining its calibration. To take it off requires moving its screens outward. Rinalducci et al. (1996) describe a method for calibrating physical FOV in an HMD using after-images. This type of method may be useful for calibrating the physical FOV of the Kaiser HMD.

3.5.3 Virtual reality software and modeling

The experiments were implemented in custom code that used the UNC-CH Vlib and VRPN libraries (VRPN, 1999). A basic overview of VR application structure and transformations as used in the Vlib library is given by Robinett and Holloway (1995). Head tracking was done using the UNC Opto-electronic ceiling tracker with HiBall sensor (Ward et al., 1992; Welch and Bishop, 1997). Communication with tracker and button devices was done using the UNC-CH VRPN library.

Models for all objects and for the background lab environment were created using 3D Studio Max. The models were shaded using precomputed radiosity lighting computed with Lightscape radiosity software. The background lab environment consisted was a simple textured-box model of Sitterson Hall's HMD lab. All experiment tasks took place within this virtual environment. An avatar (virtual body) model was included for the user. It moved along the floor according to head position; limbs were not tracked. Subjects were also given a hand-held controller with buttons to click to start and end trials.

Chapter 4
Main Results and Analysis

The experiments were carried out between April 8 and 20, 1999. I describe the results and analyses in two parts. This chapter considers the *a priori* formal hypotheses on the four main dependent variables. These are variables for which hypotheses were stated in Chapter 3 and for which statistically significant results were expected. The next chapter considers the informal hypotheses. These are hypotheses for which only trends (and not significance) were expected, or are *post hoc* hypotheses suggested by the data. The first set of results contains the main findings of this work; the second set contains exploratory findings intended to suggest future work.

4.1 A priori hypotheses

Hypotheses for the four main dependent variables were stated in Chapter 3. Those variables are *search time*, *walking time*, *VR sickness*, and *postural instability* (or ataxia). I hypothesized that these four variables could be predicted by FOV, at the three FOV levels tested in the Kaiser HMD. The four hypotheses restated are:

H1. Search time is shorter with wider FOV,

H2. Walking time is shorter with wider FOV,

H3. VR sickness is greater with wider FOV,

H4. Postural instability is greater with wider FOV.

The following sections describe the statistical methods used, the values observed for the relevant independent and dependent variables, and the results of the statistical tests.

4.2 Methods of analysis

I chose an overall significance level of $p < .05$, and performed Bonferroni correction, which is required because there is more than one hypothesis. For four variables one should accept their hypotheses as significant if the *p*-value is less than one-quarter of the overall value, i.e., $p <$

.0125 for an overall level of $p < .05$. This level (.0125) applies to both the overall model tests for each hypotheses and the tests on the independent terms (predictors) in the models.

The analysis performed was a univariate repeated-measures analysis of variance (ANOVA), i.e., ANOVA with subject as a predictor variable (Stevens, 1996). This takes into account the effects of variation between subjects. Sample code for SAS 7.0 is given in the appendices. The ANOVA tests a linear model with one dependent variable (for example, search time), and multiple independent terms (for example, the independent variables subject and FOV, and the subject-by-FOV interaction term). It produces the following statistics of interest:

- The model R^2, which gives the proportion of variance in the dependent variable accounted for by the independent terms,

- The F statistic for the model and probability p of obtaining a larger F value, indicating whether the model is significant or not.

- The F and p values for each independent term in the model, indicating whether the term is significant or not.

The following analysis sequence was followed for each dependent variable and corresponding hypothesis:

- View plots of the data for each relevant independent and dependent variable for each subject. View summary statistics (mean, standard deviation, median, minimum, maximum). Remove outliers, considered as any data falling outside plus or minus 3 times the standard deviation from the median, as computed for the appropriate subset of the data (at a given FOV level for a given subject, for example). The median is used instead of the mean because it is less affected by outliers and thus gives a better estimate of the average value (Stevens, 1996). The dataset with outliers removed is used for the following steps.

- Test a repeated measures ANOVA model with main effects and all two-way interaction effects. ANOVA is a special case of multiple regression analysis, which fits a linear model to a set of data. For example, a simple linear regression model to test how well a dependent variable, Y, is predicted by two independent variables, X_1 and X_2, is $Y = b_0 + b_1X_1 + b_2X_2$. We use a least-squares method to find the best-fit values for the regression coefficients, b_i. Then we compute the F and p statistics for the overall model and for each independent term. We can add interaction terms, which are products of independent variables, e.g., $b_3(X_1 \times X_2)$. This is a simple example for linear regression with continuous independent variables. ANOVA is a special case for dealing with categorical (discrete) independent variables and is equivalent to regression through a transformation of the variables (Cohen and Cohen, 1983).

- The next step is to refine the model iteratively by dropping terms that are not significant (i.e., a "step-down" analysis), or whose contribution to the model is small, as indicated by the proportion of the R^2 accounted for by that term (the amount that R^2 drops when that term is removed from the model). Terms that are significant but are small and readily explained by the data may be dropped from the model. This process is repeated to obtain a final ANOVA model.

- For significant terms in that refined model, perform *post hoc* pairwise Bonferroni t-tests to determine where the differences lie and the directions of the differences between each level of the independent term.

All analysis was done using statistical software, SAS version 7.0.

4.3 Independent variables

4.3.1 Subject

For the repeated measures analysis, "subject" is considered a categorical variable for the ANOVA. Six subjects passed the screening tests for stereo vision and color blindness, and were signed up to begin the experiment. One subject withdrew before the first session, leaving five subjects (three female and two male), all of whom completed the experiment. Each came to the lab for five sessions, one for each of the three main viewing conditions, and twice more for the two comparison conditions. Each session was on a different day, with typically two or three days between sessions. Subjects were between 21 and 30 years old (21/M, 21/M, 22/F, 22/F, 30/F). All subjects were right-handed. Other descriptive data for the subjects is given in Appendix A.1 (experience with VR, video games, etc.). The three main viewing conditions were presented first, in a unique order for each subject; the assignment was done using a random number generator. There are six possible orderings; only five were used because one subject withdrew from the study. The fourth and fifth sessions, using the V8 and restricted-real displays, will be discussed in Chapter 5. Table 4-1 shows the order of presentation for the subjects.

Table 4-1. Order of FOV presentation.

Subject	Session 1	Session 2	Session 3	Session 4	Session 5
A	Kaiser 48°	Kaiser 176°	Kaiser 112°	V8 48°	Restricted-real 48°
B	Kaiser 48°	Kaiser 112°	Kaiser 176°	V8 48°	Restricted-real 48°
C	Kaiser 176°	Kaiser 48°	Kaiser 112°	V8 48°	Restricted-real 48°
D	Kaiser 112°	Kaiser 48°	Kaiser 176°	V8 48°	Restricted-real 48°
E	Kaiser 112°	Kaiser 176°	Kaiser 48°	V8 48°	Restricted-real 48°

In each session, all tasks were performed (search, walk, distance estimation, spatial memory) in three VR exposures each lasting approximately 5 minutes. The distance estimation and spatial memory tasks were performed together during the third VR exposure.

4.3.2 Field of view

Field of view (FOV) had three values, corresponding to horizontal FOV of 48°, 112°, and 176° in the Kaiser HMD. Vertical FOV was always 47°. The order of presentation was initially balanced as described above, but after removal of the sixth subject was only partially balanced. FOV was treated as a categorical variable in the ANOVA.

4.3.3 Additional independent variables

The following incidental variables are task-specific and appear in the ANOVA for the main hypotheses.

- Search target angle – 90, 100, 110, 120, or 130°,
- VR exposure number (for SSQ) – 1, 2, or 3,
- Posture test type – pre-test or post-test (before or after the VR exposure).

4.4 Dependent variables

4.4.1 Search time

Data. Each subject performed the search task 80 times for each FOV – four blocks of 20 trials. The first block of 20 was considered practice and is not included in the analysis. Within each block of 20, the search target angle varied between 90°, 100°, 110°, 120°, and 130°, as discussed in Section 3.2.1.

There were no missing values for the search task, thus there were 900 measurements of search time, 60 each for 5 subjects for 3 FOV levels. Seven outliers were removed, leaving 893 observations. Points were considered outliers if the search time was outside 3 standard deviations from the median value for that subject at that FOV and target angle. These outlier values arose from trials in which subjects took an unusually long time because they became confused or forgot to press a button. Points were also considered outliers if the search time was very close to zero. This happened when a subject accidentally double-clicked the button and ended a trial by mistake.

Figure 4-1 shows average search times for all subjects at each FOV and target angle. Average search time increases as target angle increases, and decreases as FOV increases. Appendix A.2 contains the corresponding plots for each subject.

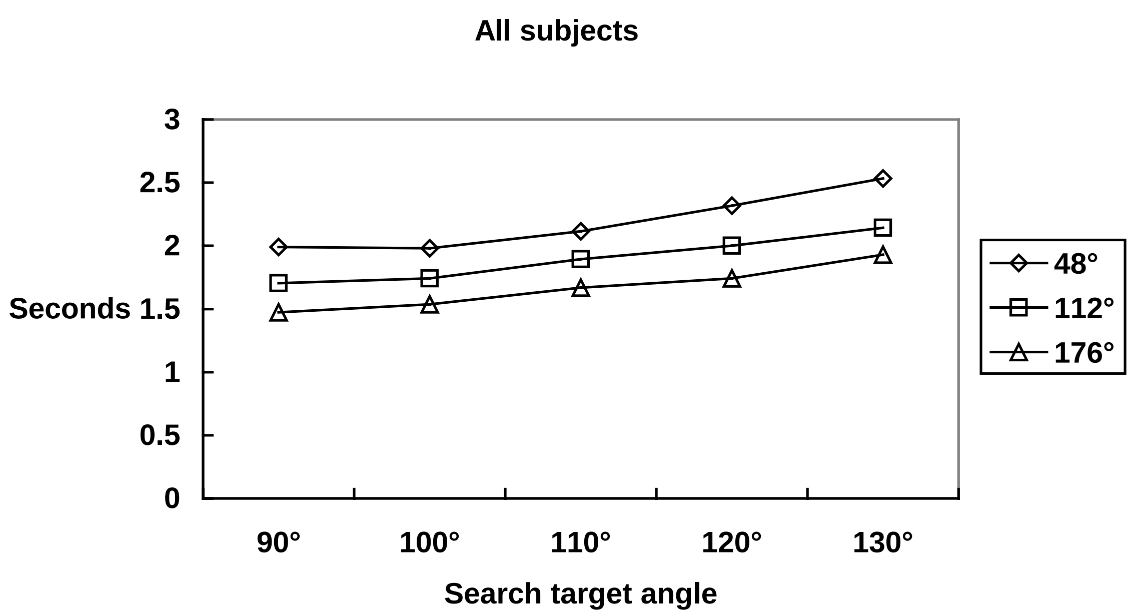

Figure 4-1. Average search time across FOV and angle.

Figure 4-2 shows the search time averaged across target angle and subject for each FOV. The error bars in this plot and in all subsequent plots with error bars indicate plus or minus one standard deviation from the mean for data across all subjects. Note that the within-subject standard deviation is typically much smaller.

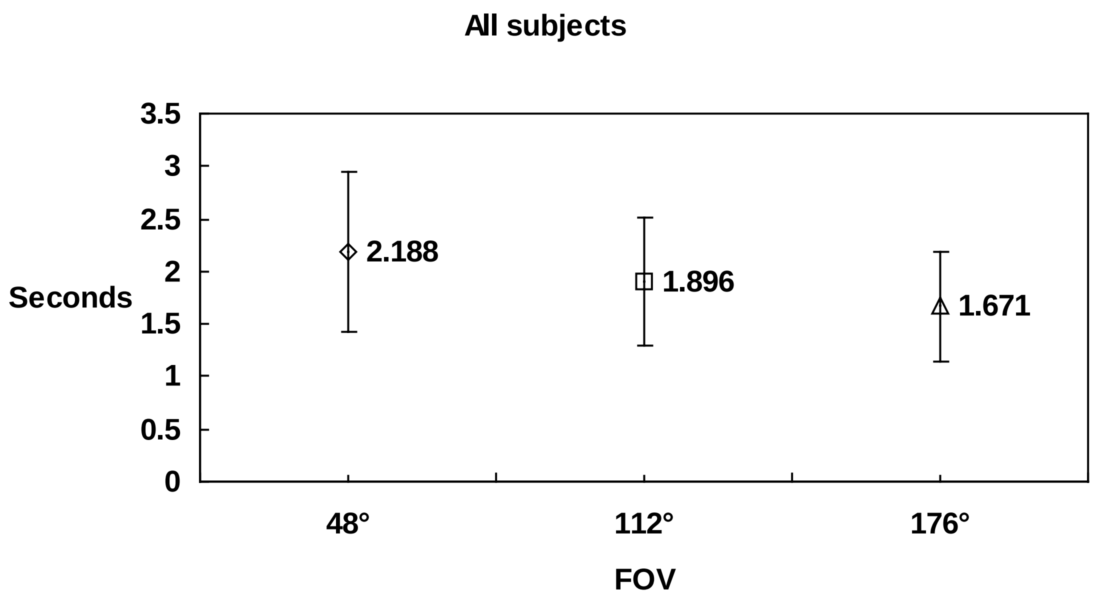

Figure 4-2. Average search time across FOV.

Analysis. In a first repeated measures ANOVA on the search data, 89.1% of the variance was accounted for. The dependent variable was search time in seconds. The independent variables were subject, FOV, and target angle, and all possible two-way interaction terms were included in the model (subject-by-FOV, subject-by-angle, and FOV-by-angle). The overall model fit was significant ($p < .0001$); the statistics are: $R^2 = 0.891$, $F(42, 850) = 166.14$, $p < .0001$. The following independent terms were significant: subject ($F = 1104.41$, $p < .0001$), FOV ($F = 385.16$, $p < .0001$), target angle ($F = 135.82$, $p < .0001$), subject-by-FOV ($F = 144.16$, $p < .0001$), and subject-by-angle ($F = 4.90$, $p < .0001$).

To improve the model I successively dropped terms that were not significant or that accounted for little of the variance in search time. First I removed the one non-significant term (the FOV-by-angle interaction) and did another analysis. This new model accounted for 89% of the variance ($R^2 = .890$).

Next I removed the subject-by-angle interaction term. This new model accounted for 87.9% of the variance ($R^2 = .879$). While the subject-by-angle term was significant, its effect is very small. Its significance means that there was a difference in how subjects performed across target angle.

This shows up in the plots in Appendix A as differences in the curve of search time versus target angle; for some subjects the slope is near constant while for others the increase from 90° to 100° to 110° is smaller than the increase from 110° to 120° to 130°. The effect, however, is small, given the small change in R^2, and therefore I judged that its removal from the model was appropriate.

Next I removed the subject-by-FOV interaction term. This new model accounted for 73.2% of the variance ($R^2 = .732$). The significance of the subject-by-FOV interaction term means that FOV had a different effect for some subjects. As indicated in the plots shown in Appendix A, only subject B's means follow the group means, with performance consistently improving from 48° to 112° to 176°. The differences may arise from two things. The order of the sessions (see Table 4.1) may have affected results – subjects usually improved over time, but the amount of improvement due to session order is smaller than the amount of improvement due to FOV. For example, subjects C and E still performed better at 176° than at 48° even though 176° was presented prior to 48°. For no subject was performance better at 48° than at 176°, though 112° sometimes was better or worse than either, or in-between. I judged the size of the interaction's effect to be small enough to permit its removal from the final model, giving a final model without interaction effects.

In this final model for search time, the statistics are $R^2 = .732$, $F(10,882) = 241.1$, $p < .0001$. The significant predictors are subject ($F = 464.51$, $p < .0001$), FOV ($F = 162.03$, $p < .0001$), and target angle ($F = 57.14$, $p < .0001$).

Pairwise Bonferroni t-tests on the means were done and showed significant differences between all levels of FOV and all subjects at the $p = .05$ level. Between all pairs of FOVs, the higher FOV gave the smaller search time. All levels of target angle were also significantly different except for 90° and 100°. Between all other pairs of target angles, the larger angle gave the higher search time.

Sample SAS code for the analysis of search time is given in Appendix A.3.

Summary. The data confirmed the hypothesis that FOV predicts search time. Wider FOV gives higher performance (shorter search time). If we consider performance at 176° to be 100%, then performance at 112° falls to 88%. Performance at 48° falls to 76%. Thus performance decreases the same at the two steps – a 12% decrease for each 64° decrease in FOV.

4.4.2 Walking time

Data. Each subject performed the walking task 30 times for each FOV – three blocks of ten trials. The first block of 10 was removed as practice and is not included in the analysis. Thus each subject contributed up to 60 measurements of walking time (20 measurements for 3 FOV sizes), giving a possible total of 300 measurements for five subjects.

Data for one subject for FOV of 176° were not collected because of equipment problems. Seven outliers were removed. These were values more than three standard deviations from the median for that subject and FOV. This left a total measurement count of 273.

Figure 4-3 shows the mean of walking time for all subjects for each FOV. As FOV increases, mean walking time decreases.

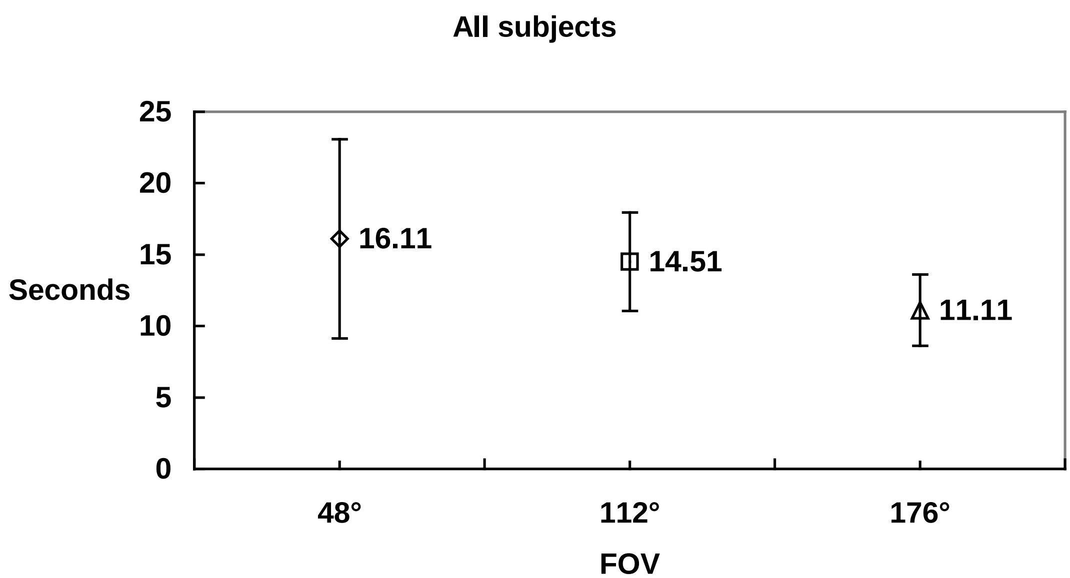

Figure 4-3. Average walking time across FOV.

Analysis. A first ANOVA gave a model accounting for 86% of the variance. This model had walking time as the dependent variable and FOV and subject as the independent variables. The interaction term subject-by-FOV was also included in the model. The overall model fit was significant (R^2 = .86, F(13,259)=122.47, p < .0001). FOV and subject were significant predictors (for FOV, F = 128.97, p < .0001, for subject, F = 235.96, p < .0001). The subject-by-FOV interaction was also significant (F = 52.91, p < .0001). Thus a model with these three terms accounts for 86% of the variance.

I tested a second model without the subject-by-FOV interaction to determine that term's importance. This model had only FOV and subject as the independent variables and accounted for 66% of the variance (R^2 = .66, F(6,266)=86.07, p < .0001). In this model, FOV was a significant predictor (F = 58.74, p < .0001), as was subject (F = 99.73, p < .0001).

The subject-by-FOV interaction is likely to have occurred mostly because of subject E's variation. Subject E's mean walking time was 9.75 at 48°, 16.03 at 112° and 10.01 at 176°. Scatter plots of the walking data for all subjects are presented in Appendix A.4. Subject E's high performance at 48° was likely due to practice or a change in strategy.

Pairwise Bonferroni t-test comparisons on this second model (with R^2 = .66) gave significance at the p < 0.05 level for all pairs of FOV. In all pairs, the wider FOV gave the shorter walking time.

An alternate analysis was tried on the walking data using one-way ANOVA on means instead of repeated measures. The data for this ANOVA were simply the mean walking time for each subject at each FOV. There were 14 data points. The ANOVA, with walking time as the dependent variable and FOV as the independent variable did not yield a significant fit (R^2 = 0.175, F(2,11) = 1.17, p = .347). It was considered likely that this analysis did not yield significance because of the small number of data points. Thus we rejected this analysis method and kept the previous repeated measures ANOVA model using all the data.

Summary. The data show that FOV size predicts walking time. FOV is significant, and wider FOV results in shorter walking time at all levels tested. If we consider performance at 176° to 100%, then performance at 112° is 77%, and performance at 48° is 69%. The higher performance drop occurs in the drop to 112°.

4.4.3 VR sickness

Data. The subjects answered the SSQ five times during each experiment session: once before starting, once after each of the three VR exposures, and once just before leaving the lab (after a break to fill out the presence questionnaire). The purpose of the first and last tests was only to verify that the subject was not sick to begin with and was not sick afterwards and at risk when leaving the lab. The first and fifth tests were not included in the analysis. Serious sickness was not seen in any of the first or fifth SSQ tests; in most cases no more than three symptoms, all "slight" were reported.

There are up to three SSQ scores for each subject at each FOV, one each after the first, second, and third VR exposures. Two measurements are missing because of equipment problems that caused the subject to end the session early. There are 43 measurements in total for the five subjects.

When scored, the SSQ gives four numbers: a “total severity” and three diagnostic subscale scores (nausea, disorientation, and oculomotor discomfort). I used Total Severity as the main indicator of sickness, following Kennedy (1993) and Kolasinski (1996).

Table 4-2 shows summary statistics of post-exposure sickness scores over the entire experiment. These values give an overall indication of the “health” of our VR system. The figures that follow compare our system to others previously reported.

Table 4-2. Average sickness scores.

Score	Mean	Standard Deviation	Low	High
Total severity	11.65	10.89	0	37.40
Nausea	8.431	8.381	0	28.63
Oculomotor Discomfort	10.05	10.56	0	37.90
Disorientation	12.63	16.02	0	41.76

I compared these values to SSQ scores that have been reported for other systems. The chart in Figure 4-4 shows mean Total Severity from the SSQ as reported from five different sources. Figure 4-5 shows means of the three SSQ subscores to allow comparison of the “sickness profiles” (Stanney and Salvendy, 1998). The data marked Arthur are from the current study. The Kolasinski data are from 20 minute exposures using virtual i-o glasses (HMD) to play a video game (Kolasinski, 1996). The Usoh data are from a study of the effects of locomotion technique on presence in VR (Usoh et al., 1999). Exposure times in that study were approximately 5 minutes. The Stanney data are averages from five VR systems (using HMDs) (Stanney and Salvendy, 1998). The Kennedy data are averages from five flight simulators (presumably using dome screens, though the systems were not described) (Kennedy et al., 1993).

While the level of sickness is our system was not zero, it is far less than the level of scores that typically induce serious symptoms or vomiting. Kolasinski reported her most serious effects occurring in a subject with total sickness score of 138.38, nausea 114.48, oculomotor discomfort 90.96, and disorientation 180.96. The subject had to withdraw from the exposure after just over 7 minutes and vomited three times (Kolasinski, 1996).

Sickness levels found in the present study are close to those found by Usoh et al., which is not surprising given that those experiments also used the low-latency UNC-CH Opto-electronic tracker system used in the current study. The Kolasinski system could be expected to induce more sickness because visual motion was not tied to user motion (it was controlled by joystick). Much less sickness is typically found in flight simulators, as indicated by the Kennedy data, and it is important to note that those data were collected after much longer exposures (an entire flight) than are the SSQ scores reported for VR systems.

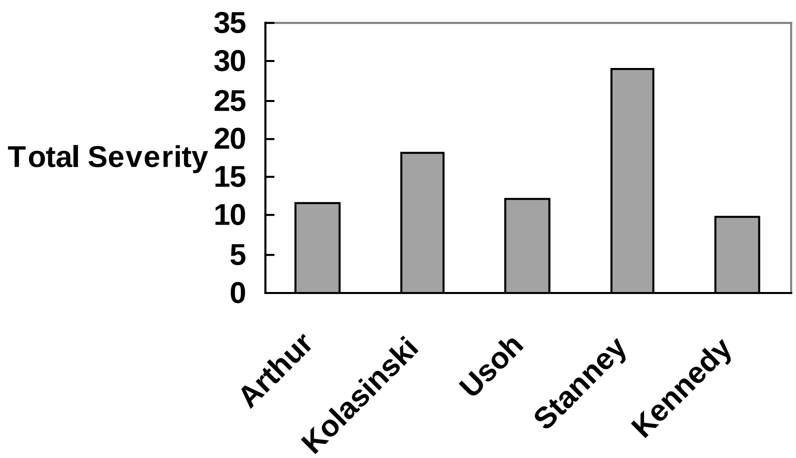

Figure 4-4. Total Severity comparison of 5 reports of sickness.

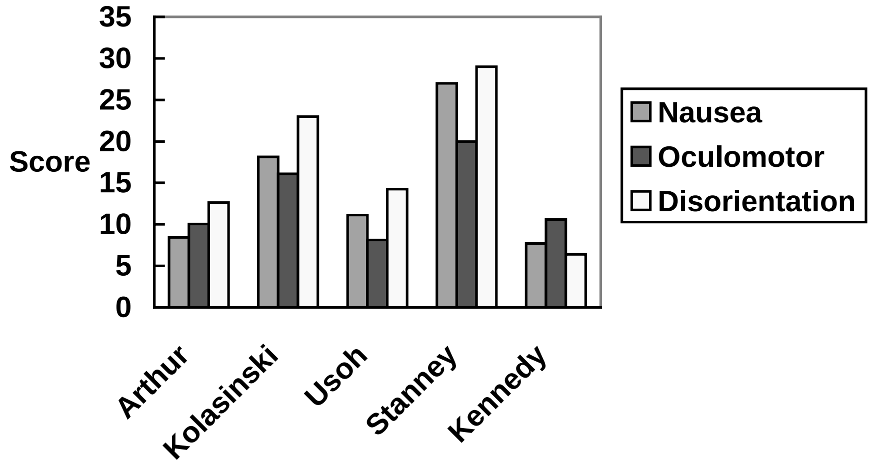

Figure 4-5. Profile comparison of 5 reports of sickness.

In the next chapter I consider the SSQ subscores obtained in the present experiment and their variation with FOV. The remainder of this section discusses only total severity of sickness.

Figure 4-6 shows averages of total severity for the three FOV conditions, broken down by exposure number. Recall that there were three VR exposures per session (per display condition). On average, sickness was greater for the second exposure (this difference was significant), and yet greater for the third exposure (but this difference was not significant), under all FOV conditions.

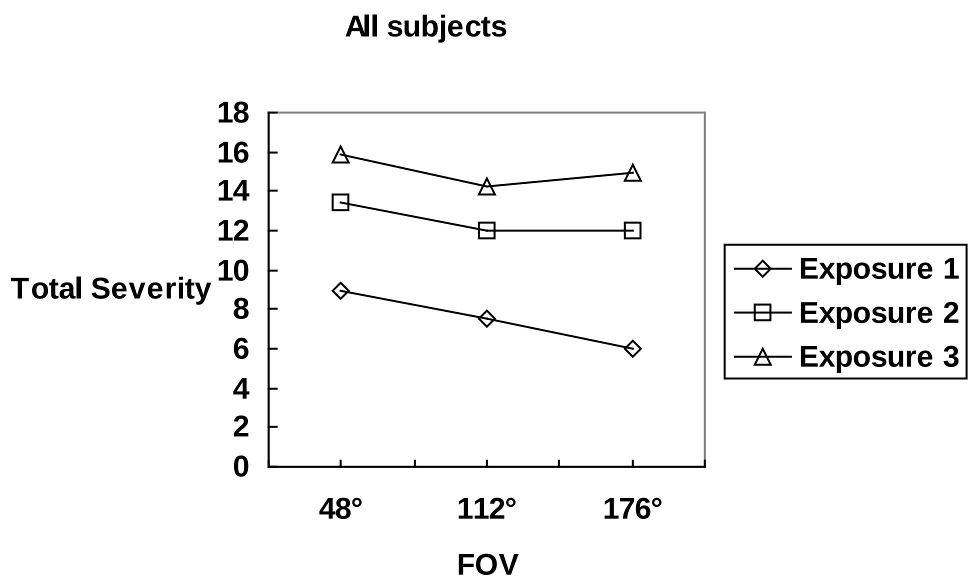

Figure 4-6. Total Severity of sickness across FOV and exposure number.

Averaging over exposure number gives means of total severity for each FOV of 12.56, 11.49, and 10.97 for 48°, 112°, and 176° respectively, indicating a slight decrease in total sickness with increasing FOV, though this difference was not significant.

Analysis. A first ANOVA with all factors included gave a model accounting for 97% of the variance. This model had total sickness as the dependent variable, and subject, FOV, and exposure number as independent variables. All two-way interaction terms were also included in the model: subject-by-FOV, subject-by-exposure, and FOV-by-exposure. The overall model had $R^2 = .965$, $F(28, 14) = 13.87$, $p < .0001$. The significant predictors were subject ($F = 71.58$, $p < .0001$), time ($F = 16.42$, $p = .0002$), and the subject-by-exposure interaction ($F = 6.08$, $p = .001$).

I dropped all non-significant interaction terms from the model, giving a new model with $R^2 = .918$, $F(16,26) = 18.13$, $p < .0001$. Dropping the subject-by-exposure interaction from this model gave $R^2 = .797$, $F(8, 34) = 16.66$, $p < .0001$. I kept the previous model as the final model. Its overall statistics were $R^2 = .918$, $F(16,42) = 18.13$, $p < .0001$. The following were the significant independent terms: subject ($F = 56.2$, $p < .0001$), exposure number ($F = 12.89$, $p < .0001$), and subject-by-exposure ($F = 4.78$, $p < .001$).

The subject-by-exposure interaction arose from the fact that subjects A and B did not follow the trend of greater sickness with successive exposures, whereas subjects C, D, and E did follow the trend. The sickness averages for each subject are presented in Appendix A.5.

Post-hoc Bonferroni t-tests on variables in this model showed that total severity was significantly different between exposures 1 and 2 but not between exposures 2 and 3.

Summary. The data showed no significant effect of FOV on VR sickness. Section 5.2 contains a power analysis to determine the sample size needed for significance. I found significance for exposure time; sickness increased with number of VR exposures within one session on a single day.

4.4.4 Postural instability

Data. The posture test was done before and after each VR exposure, giving six tests per session. During the posture test the system recorded a log of head position and orientation. We computed average translational head velocity (meters per second) during the posture test (from position values only).

The mean pre-exposure instability was .0288 m/s, and the mean post-exposure instability was .0301 m/s (Figure 4-7). The difference is .0013. A first ANOVA test with only subject and test type (pre- or post-exposure) as variables gave significance only for subject, and not for test type. Thus our data are inconclusive on whether instability increased after VR exposure. Section 5.2 contains a power analysis to determine the number of samples required to find a significant effect of this size.

Kolasinski also found that instability did not consistently increase following exposure to a virtual environment, and in fact there was a decrease on average (Kolasinski, 1996).

Figure 4-8 shows the average post-exposure postural instabilities by FOV for all subjects. Instability was lowest for 112°. It was higher for 48° and higher again for 176°. These differences were not statistically significant.

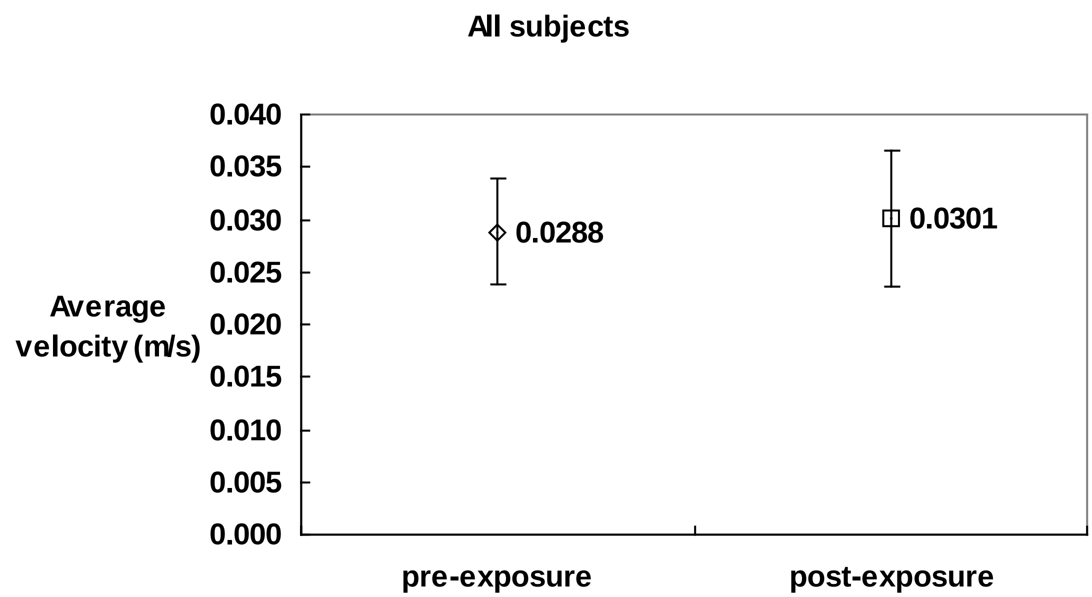

Figure 4-7. Average postural instability pre- and post-exposure.

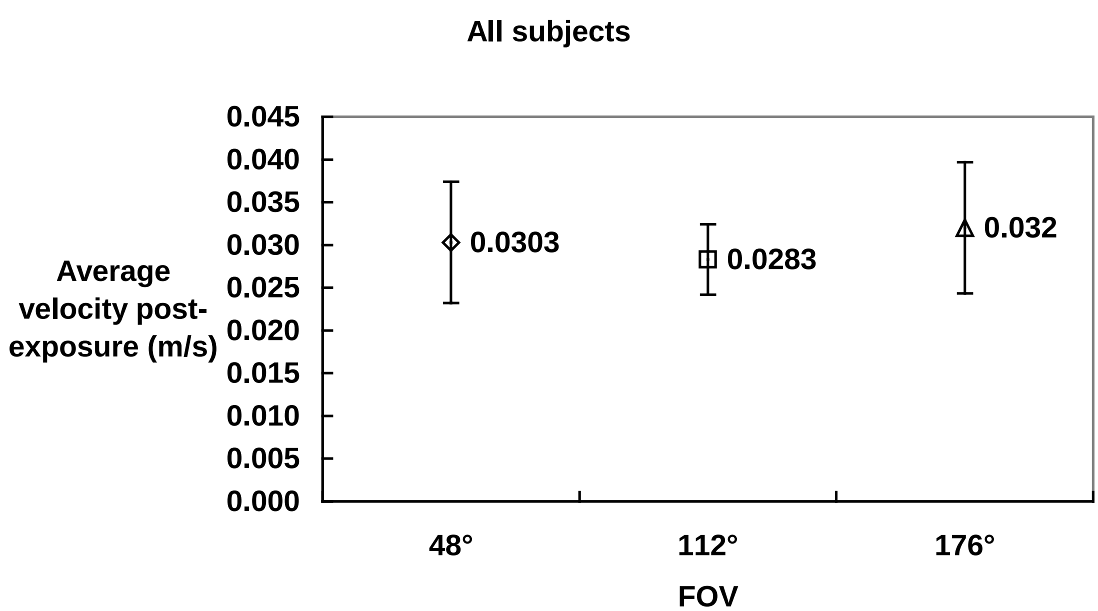

Figure 4-8. Average postural instability post-exposure across FOV.

Analysis. A first ANOVA was tested with post-exposure postural instability as the dependent variable and with subject, FOV, and posture test type (pre- or post-exposure) as the independent variables, and with all two-way interactions. The analysis gave overall model R^2=.33, $F(17, 65) = 1.86$, $p = .0391$. Subject was a significant predictor ($p = .0017$), but no other terms were significant.

A second ANOVA without the interaction terms was tested. This model, with post-exposure postural instability as the dependent variable and with subject, FOV, and posture test type (pre- or post-exposure) as the independent variables gave overall model R^2=.23, $F(7, 75) = 3.18$, $p = .005$. Subject was a significant predictor ($p = .0014$), but FOV and type were not significant.

Summary. The data showed no effect of FOV on postural instability. The hypothesis that wider FOV would increase postural instability was not supported. Nor did the data support the hypothesis that VR exposure increases postural instability; on average, instability did increase slightly after exposure to VR, but this difference was not statistically significant.

Chapter 5
Exploratory Results and Analysis

Additional analyses were performed based on exploratory hypotheses suggested by the literature review and by the data. The analysis methods in this chapter are similar to those used in the previous chapter. The main statistical method was univariate repeated measures ANOVA. Analysis was performed on several new dependent variables. All exploratory hypotheses were tested at the $p < .05$ level.

New tests on dependent variables are presented in the Section 5.1. Section 5.2 discusses the statistical power of analyses performed in Chapters 4 and 5, and recommends sample sizes for studies extending this work. Section 5.3 presents analysis of the secondary viewing conditions that were tested for comparison, the V8 HMD and the restricted-real conditions. Section 5.4 summarizes the written comments obtained from subjects.

5.1 Dependent variables

5.1.1 Distance estimation

Data. Subjects performed the distance estimation task in a virtual environment containing virtual styrofoam walls and a white circle placed on the floor (for details on the methods see Section 3.2.3). The subject's task was to estimate the distance from the red starting circle to the white circle. They did this within the virtual environment by walking from the red circle to the place where they think they saw the white circle. (Recall that "Distance estimation" is the egocentric memory task with recall *within* VR; "Spatial memory" is the exocentric memory task with recall *outside* VR, discussed in Sections 3.2.4 and 5.1.2.) The distance estimation task was performed 10 times per FOV per subject, giving up to 150 measurements in total. Data for one subject at two levels of FOV were lost due to equipment problems, leaving 130 measurements in total.

The accuracy of the distance estimate is taken here as a ratio, estimated distance divided by actual distance. "Estimated distance" is the distance from the red circle to the estimated position of the white circle. "Actual distance" is the distance from the red circle to the actual position of the white circle. An accuracy ratio of 1 means the subject estimated the distance exactly. An accuracy less than 1 means underestimation, and greater than 1 means overestimation. I used this distance measure because it emphasizes over- or under-estimation, which is the interesting quantity that I hypothesized would vary with FOV. The angle or path they took to get to the white circle is less interesting for FOV.

The actual position of the white circle was used as an independent variable in the analysis. The white circle varied among five potential positions in the environment. The variable, circle position, was assigned corresponding values of 0, 1, 2, 3, or 4. The assignment was different for each subject.

Analysis. An ANOVA was performed with accuracy as the dependent variable and subject, FOV, and circle position as independent variables. All two-way interaction terms were included in the model – subject-by-FOV, subject-by-position, and FOV-by-position. The overall model fit was significant and accounted for 61.7% of the variance (R^2 = .617, $F(40, 90) = 3.62$, $p < .0001$). The following independent terms were significant predictors: subject ($F = 10.92$, $p < .0001$), circle position ($F = 8.93$, $p < .0001$), subject-by-FOV ($F = 2.42$, $p = .032$), and subject-by-position ($F = 2.55$, $p < .002$).

Summary. No significant main effect was seen for FOV, so the hypothesis that distance estimation would be worse for restricted FOV was not supported by the data. Figure 5-1 shows average distance estimation accuracy across FOV. There is no apparent trend, nor were the means significantly different.

Subjects performed differently from each other, and FOV had different effects for different subjects. Circle position affected accuracy and did so in different ways for different subjects. Figure 5-2 shows averages of distance estimation accuracy across FOV and circle position, with circle positions ordered according to increasing distance from the red circle. There was no apparent pattern in how accuracy varied with circle position. Appendix A.7 shows data summaries for each subject.

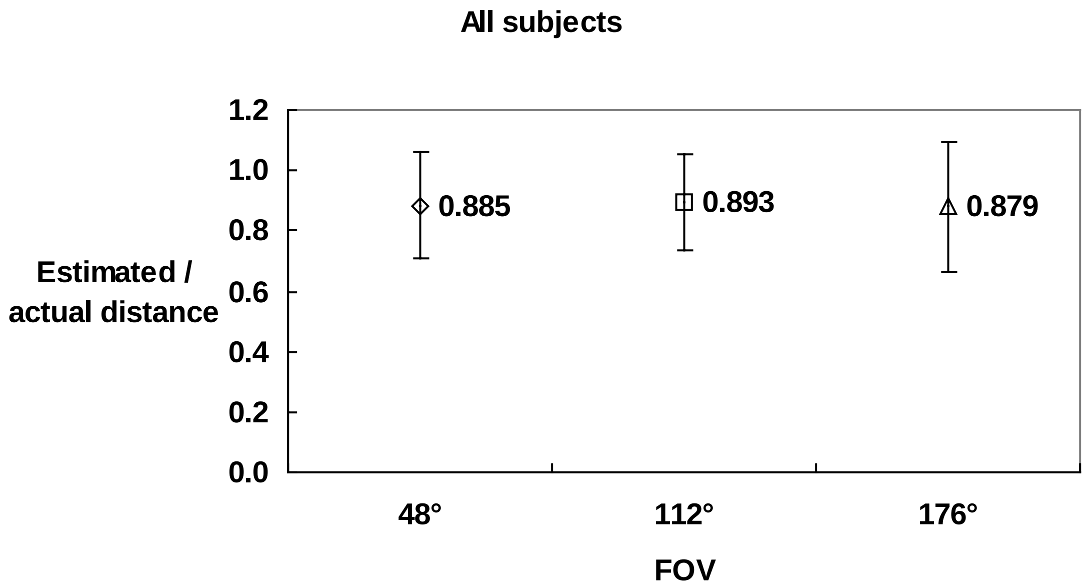

Figure 5-1. Average distance estimation accuracy across FOV.

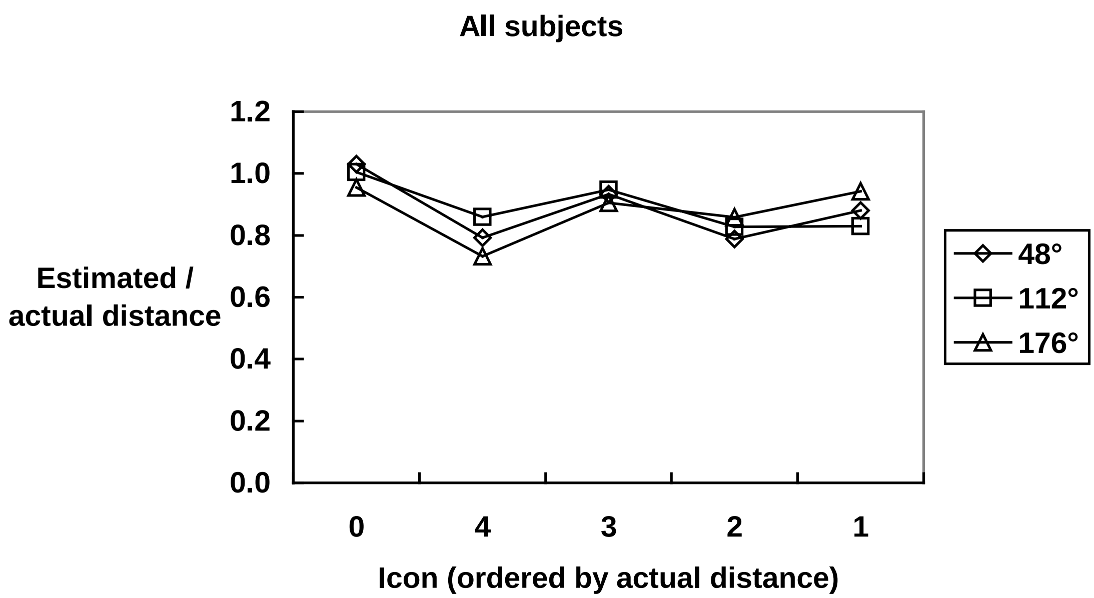

Figure 5-2. Average distance estimation accuracy across FOV and icon number.

5.1.2 Spatial memory

Data. Subjects performed the spatial memory task in two steps (for details on the methods see Section 3.2.4). The recall task for spatial memory was to place, on a 2D layout, five colored circles that were seen previously in the virtual environment. The subject performed the recall task once for each FOV. There are five circles and so there are five measures of accuracy per FOV per subject, giving up to 75 measurements in total. Data was lost for subject D at 48° and 176° and subject E at 112° because of equipment problems, leaving 60 measurements in total.

The measure used in the analysis for spatial memory accuracy is the error in placement for each of the five circles. For a given circle the error is the distance between the estimated and actual positions. The actual position is given by (x, y) coordinates in the virtual environment (world coordinates). The estimated position is indicated by the subject on the screen, and those screen coordinates were transformed to world coordinates (the scale and offsets are determined by the red and blue start/end circles).

The position of each circle was included in the analysis as an independent variable. There were five candidate positions, so circle position has the value 0, 1, 2, 3, or 4. The five candidate positions are different from the white circle positions used for the distance estimation task.

Analysis. An ANOVA was performed with error as the dependent variable and subject, FOV, and circle position as independent variables. All two-way interaction terms were included in the model – subject-by-FOV, subject-by-position, and FOV-by-position. The overall model fit was not significant (R^2 = .634, $F(30, 20) = 0.89$, $p = 0.633$). None of the independent terms were significant predictors.

Summary. No significant main effect was seen for FOV, so the hypothesis that spatial memory would be worse for restricted FOV was not supported by the data. Figure 5-3 shows average spatial memory error across FOV. There is no apparent trend, nor were the means significantly different. Circle position did not significantly affect error, nor was there any apparent trend with distance from the starting point (red circle). It appears that recall performance was worst for circle position number 2 (this is the topmost circle in the center of Figure 3-7). This may be because that circle was the closest to a wall. Appendix A.8 contains data summaries for each subject.

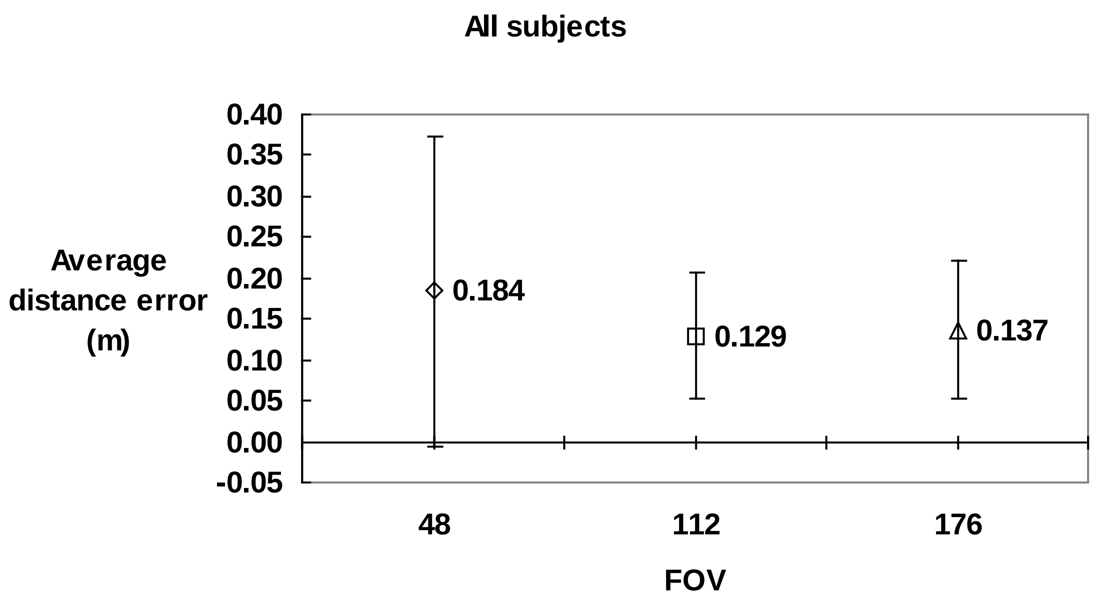

Figure 5-3. Average accuracy of spatial memory test across FOV.

5.1.3 Presence

Data. The modified Witmer and Singer Presence questionnaire (1998) was administered once per FOV. The questionnaire is reproduced in Appendix B.5. It was administered on paper at the end of the experiment session after the subject had completed all three VR exposures. For one subject at one FOV (subject D, 48°), presence data were not collected because the session ended early due to equipment problems. Thus there were 14 measures collected. When scored, the questionnaire gives a measure of total presence and measures on three subscales: involved/control, natural, and interface quality.

Analysis. Each presence score (total, involved/control, natural, and interface quality) was entered as a dependent variable in an ANOVA with subject and FOV as the only independent variables. The subject-by-FOV interaction was not included in the model because, with only 14 observations, there were not enough degrees of freedom to allow for it.

For total presence, the model was significant and accounted for 83.1% of the variance (R^2 = .831, F(6, 7) = 5.76, p = 0.018). Subject was a significant predictor (F = 7.49, p = .0114), but FOV was not. Appendix A.9 includes plots showing presence scores for each subject. Figure 5-4, below, shows averages of total presence for all subjects. The apparent trend is that scores increase with wider FOV, but this difference was not significant.

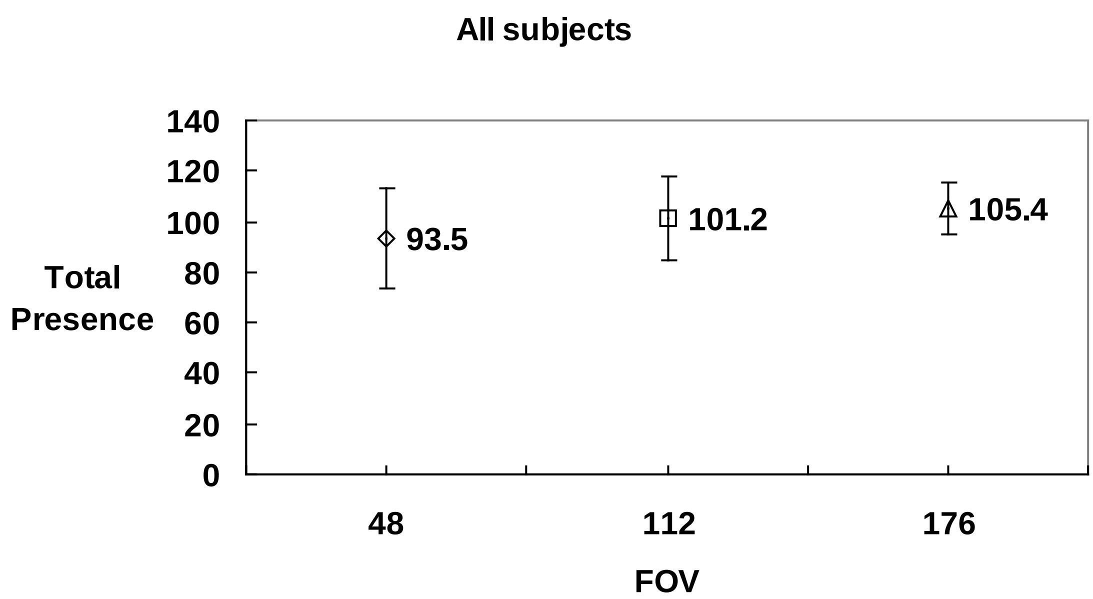

Figure 5-4. Average total presence across FOV.

For the "presence: involved/control" subscore, the results were similar: the model was significant and accounted for 83.1% of the variance (R^2 = .831, F(6, 7) = 5.77, p = 0.018). Subject was a significant predictor (F = 6.64, p = .015), but FOV was not a significant predictor. Figure 5-5 shows averages of presence: involved/control for all subjects.

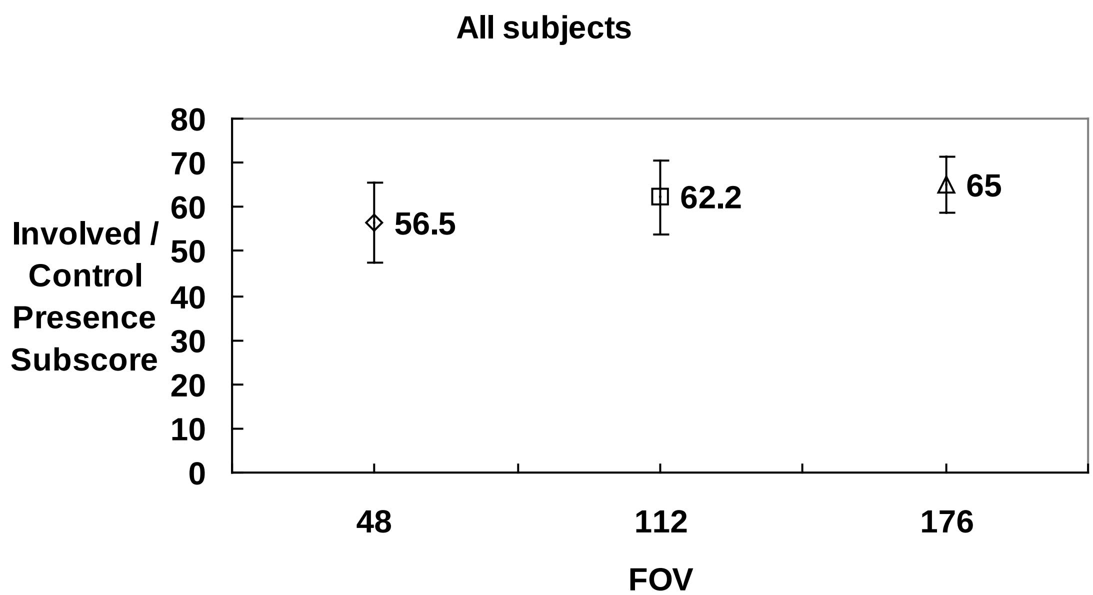

Figure 5-5. Average Involved/Control Presence Subscore across FOV.

For the "presence: natural" subscore, the results were again similar: the model was significant and accounted for 80.6% of the variance (R^2 = .806, $F(6, 7) = 4.86$, p = 0.0286). Subject was a significant predictor ($F = 5.24$, $p = .028$), but FOV was not a significant predictor. Figure 5-6 shows averages of presence: natural for all subjects. Again, scores appear to increase with wider FOV, but this difference was not significant.

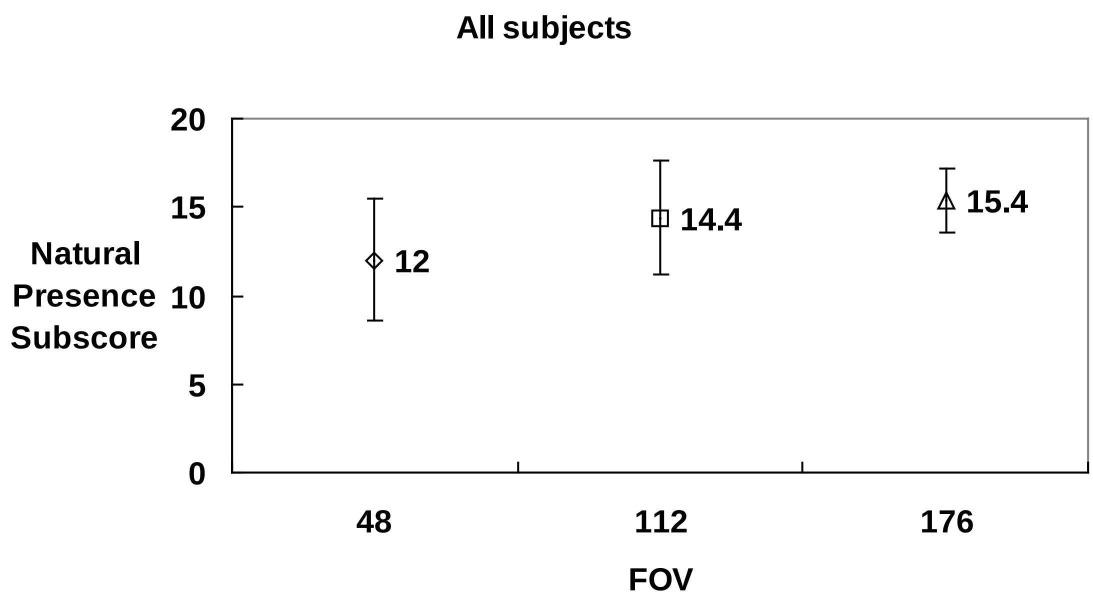

Figure 5-6. Average Natural Presence subscore across FOV.

For the “presence: interface quality” subscore, the model was significant and accounted for 80.5% of the variance (R^2 = .805, $F(6, 7)$ = 4.82, p = 0.0291). Subject was a significant predictor (F = 7.16, p = .012), but FOV was not a significant predictor. Figure 5-7 shows averages of presence: interface quality for all subjects. There does not appear to be any strong trend with FOV.

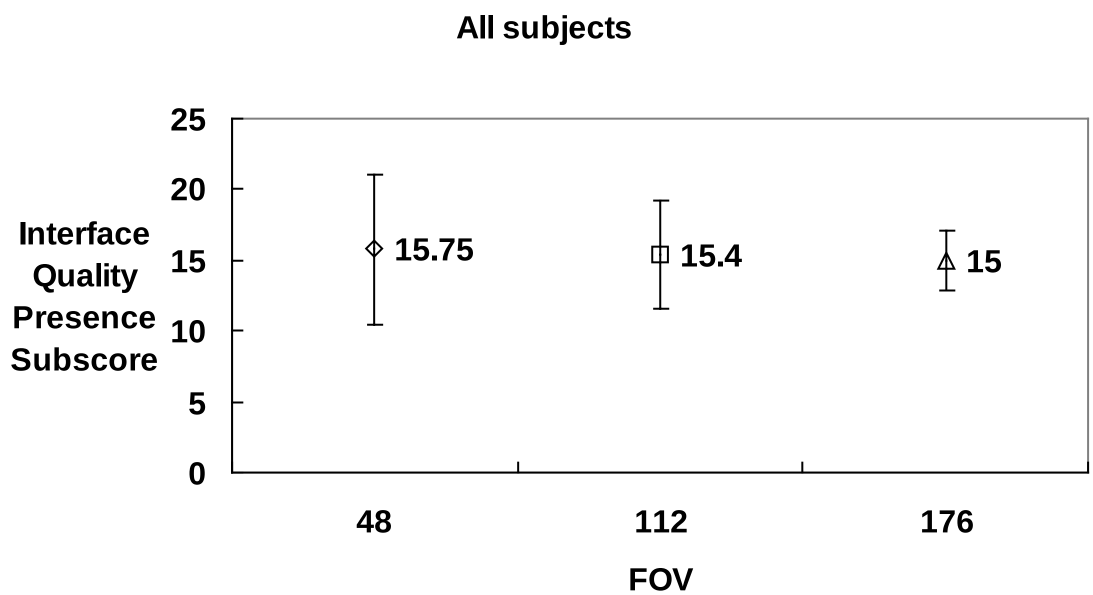

Figure 5-7. Average Interface Quality Presence subscore across FOV.

Summary. FOV was not a significant predictor of presence in the experiment. The hypothesis that presence is greater with wider FOV was not supported statistically. There are trends apparent, though, in total presence and the interactive/control and natural subscores. It appears that wider FOV may increase those measures of presence, but more samples are needed to verify this statistically (see Section 5.2 for a power analysis).

5.1.4 VR Sickness subscores

Data. As described in Section 4.4.3, sickness scores were obtained for analysis once after each VR exposure. The SSQ gives measures of total sickness, nausea, oculomotor discomfort, and disorientation. This section discusses the last three. There were 43 measures of each. Two measurements are missing because of equipment problems that caused the subject to end those sessions early.

Analysis. Analyses of variance were performed for each of the three SSQ subscores, with score as the dependent variable and subject, FOV, and exposure number as independent variables. All two way interaction terms were included in the first model: subject-by-FOV, subject-by-exposure, and exposure-by-FOV. After tests on the first model, nonsignificant terms were removed to find better models.

The first test for nausea gave overall model statistics of $R^2 = .845$, $F(28,14) = 2.73$, $p = .025$. The following independent terms were significant predictors: subject ($F = 11.03$, $p = .0003$), and exposure number ($F = 5.06$, $p = .0222$). Removing the interaction terms, which were not significant, and repeating the analysis gave a model for nausea with $R^2 = .667$, $F(8, 34) = 8.52$, $p < .0001$. The following terms were significant in that model: subject ($F = 12.46$, $p < .0001$), FOV ($F = 3.43$, $p = .0441$), and exposure number ($F = 5.71$, $p = .0073$).

The first test for oculomotor discomfort gave overall model statistics of $R^2 = .961$, $F(28,14) = 12.34$, $p < .0001$. The following independent terms were significant predictors: subject ($F = 75.0$, $p < .0001$), and subject-by-exposure-number ($F = 2.73$, $p = .048$). The nonsignificant terms were removed from this model and the analysis was repeated. In this final model for oculomotor discomfort, $R^2 = .898$, $F(14, 28) = 17.55$, $p < .0001$, and the two independent terms were significant: subject ($F = 57.07$, $p < .0001$), and subject-by-exposure-number ($F = 1.74$, $p = .0121$).

The first test for disorientation gave overall model statistics of $R^2 = .940$, $F(28,14) = 7.81$, $p < .0001$. The following independent terms were significant predictors: subject ($F = 28.19$, $p < .0001$), exposure number ($F = 21.58$, $p < .0001$), and subject-by-exposure-number ($F = 6.18$, $p = .0016$). The nonsignificant terms were removed from this model and the analysis was repeated. In this final model for disorientation, $R^2 = .883$, $F(14, 28) = 15.12$, $p < .0001$, and the independent terms were significant: subject ($F = 29.05$, $p < .0001$), exposure number ($F = 22.23$, $p < .0001$), and subject-by-exposure-number ($F = 6.37$, $p = .0001$).

The following chart shows the averages of nausea, oculomotor discomfort, and disorientation for all subjects. While FOV was significant only for predicting nausea, there is a (non-significant) trend in the disorientation scores as well. Nausea decreases on average with increasing FOV, but disorientation increases on average with increasing FOV (but that difference was not statistically significant).

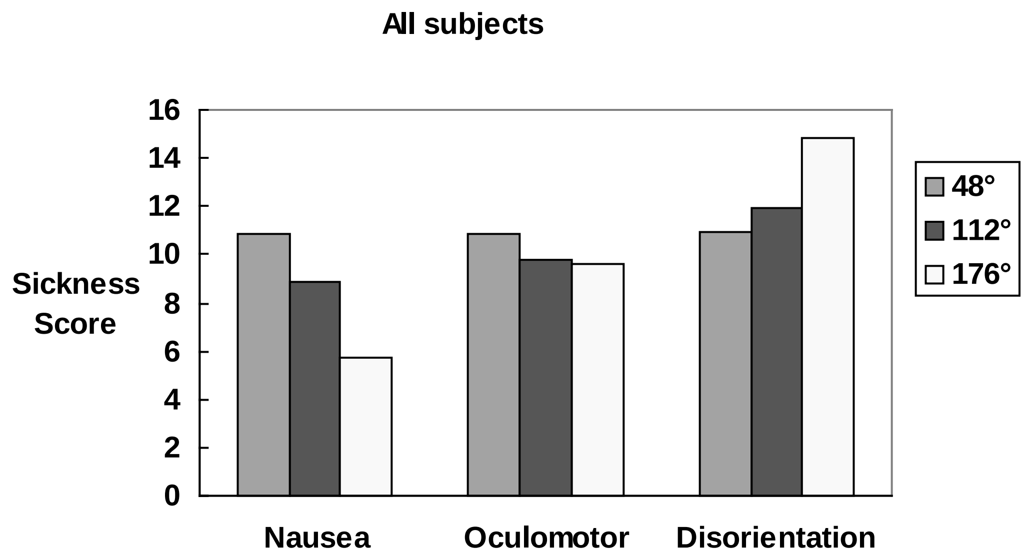

Figure 5-8. Average sickness subscores across FOV.

To understand the opposing trends I looked at the original questionnaire data and the subscore computations. The sixteen symptoms of sickness contribute to the subscores as shown in Table 5-1. The subject responds on a scale of 0 through 3 for each symptom, where 0 means “none” and 3 means “severe”. These scores are then added in different combinations to obtain the sickness subscores, as indicated in the table (for more details see Appendix B.4).

Table 5-1. Symptoms for Simulator Sickness Questionnaire.

Symptom	Weight for symptoms		
(scored 0, 1, 2, or 3)	Nausea	Oculomotor	Disorientation
1. General discomfort	1	1	
2. Fatigue		1	
3. Headache		1	
4. Eye strain		1	
5. Difficulty focusing		1	1

6. Increased salivation	1		
7. Sweating	1		
8. Nausea	1		1
9. Difficulty concentrating	1	1	
10. Fullness of head			1
11. Blurred vision		1	1
12. Dizzy (eyes open)			1
13. Dizzy (eyes closed)			1
14. Vertigo			1
15. Stomach awareness	1		
16. Burping	1		

Figure 5-9 below shows the totals obtained for each of the sixteen symptoms for each FOV. Each entry is the sum of responses (each on a 0-3 scale) over all 5 subjects (thus a minimum of 0, maximum of 15).

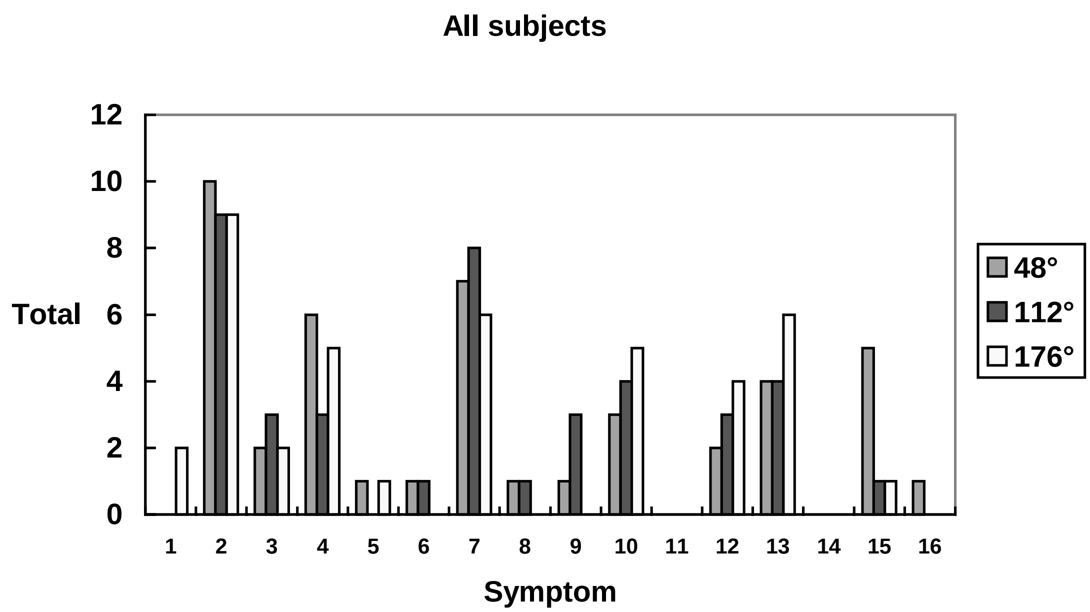

Figure 5-9. Total scores on all sickness symptoms across FOV.

The next chart, Figure 5-10, is similar but shows only the symptoms that contribute to nausea. Figure 5-11 shows only the symptoms that contribute to disorientation.

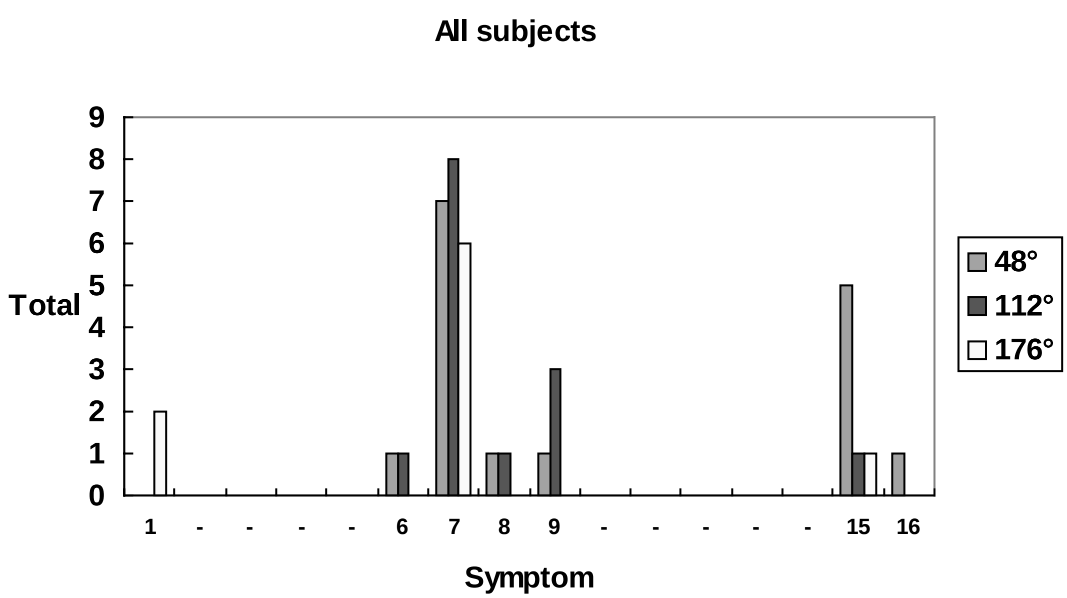

Figure 5-10. Scores on sickness symptoms that contribute to nausea subscore.

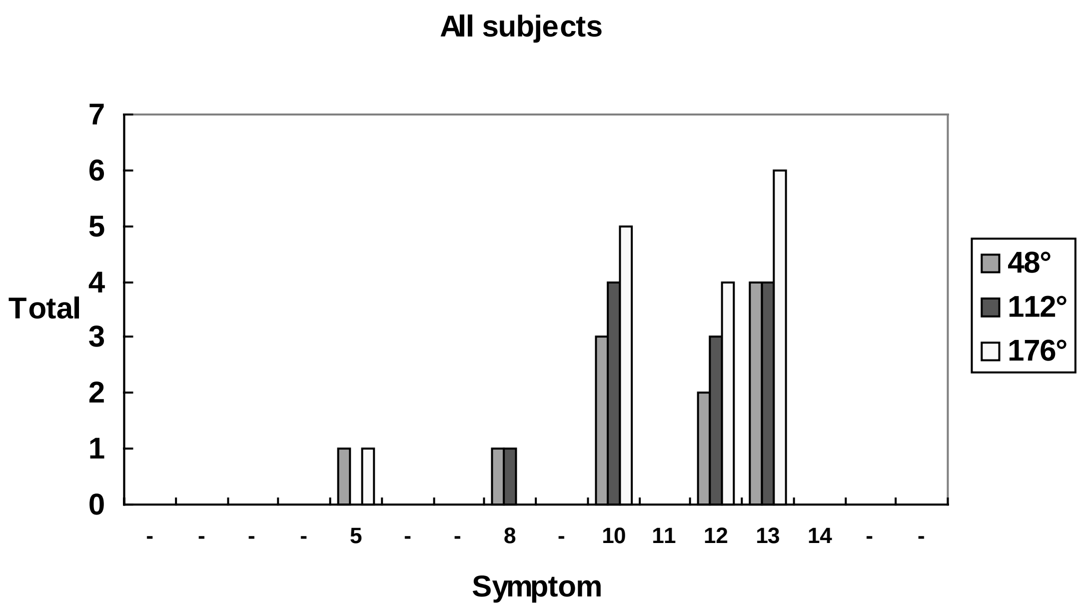

Figure 5-11. Scores on sickness symptoms that contribute to disorientation subscore.

From Figure 5-10 it appears that the trend in nausea scores is due most to the responses on symptom 15 – stomach awareness. From Figure 5-11 it appears that the trend in disorientation scores occurred because of responses on three symptoms: 10 – fullness of head, 12 – dizzy (eyes open), and 13 – dizzy (eyes closed).

Summary. There is a trend in the data showing that disorientation may increase with FOV and nausea may decrease with FOV, though the difference for disorientation is not statistically significant. Further work with more samples is needed to determine whether the effects of FOV on these subscores is real or not.

5.1.5 Head-movement velocity

Head movement was recorded throughout the experiment sessions. I used head movement data collected during postural instability trials to compute a measure of postural instability (average translational head-movement velocity), as described in Section 4.4.4. Head movement data were also kept for all other experiment trials.

Figure 5-12 shows mean translational head-movement velocity during all walking trials for all subjects. Velocity appears to increase with FOV. This relationship is approximately the inverse of that found in Section 4.4.2 – that walking time decreases with FOV. It has been suggested that restricting FOV might increase head movement (beyond this expected increase) (Dolezal, 1982b), but the walking data do not verify or refute this hypothesis. That trend might be seen with a more complete analysis that looks at rotational velocity as well.

Mean translational head-movement velocity during search tasks was also computed for one subject, but no trend was apparent in the data.

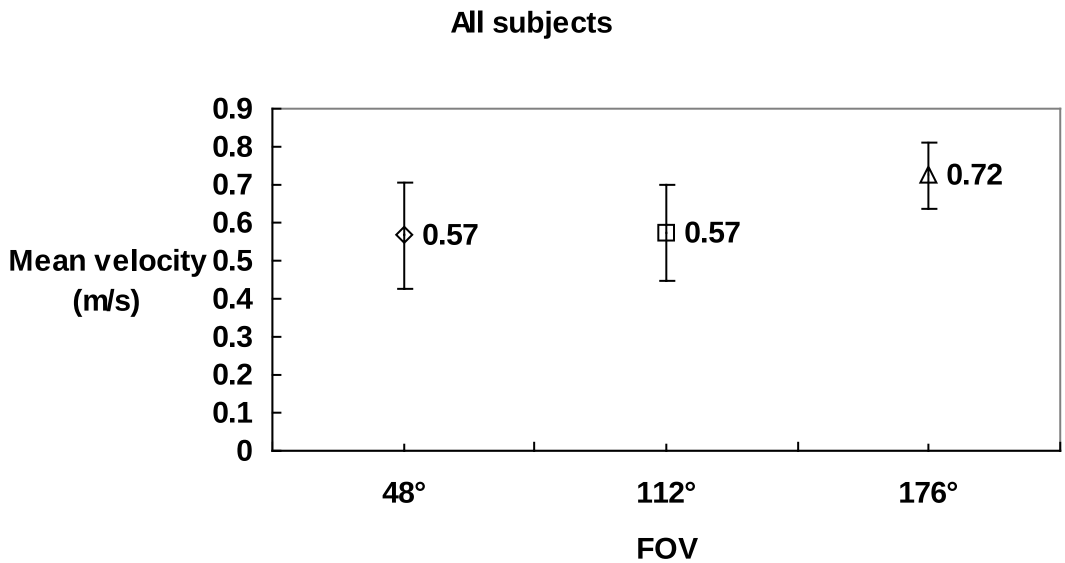

Figure 5-12. Head-movement velocity across FOV for walking task.

5.1.6 Room size estimation

As part of the presence questionnaire (Appendix B.5), an additional question (#33) asked "How would you describe the size of the *virtual* room that you saw inside the headset compared to the size of the *real* room that you saw when you took off the headset? The virtual room was: Smaller → Larger (1-7)." The question was included to check whether any trend in size estimation with FOV might occur, as has been found in previous real-world studies (Hagen et al., 1978; Dolezal, 1982b).

All answers fell between 3 and 5 (4 meant "same size"). The data are presented in Appendix A.10 for each subject. Almost all variation in responses was due to subject differences. No trend was apparent with FOV – most subjects answered with the same number each time.

5.2 Estimates of statistical power

This section contains a power analysis for variables and statistical tests used in Chapters 4 and 5.

In Chapters 4 and 5 I referred to the p-level or α that I chose, .05, or .0125 after Bonferroni correction. The value α is the probability of making a Type I error (a false positive) concluding that there is a significant difference between two groups when in truth there is no difference. Conversely, $(1 - \alpha)$ is the probability of correctly detecting that there is no difference. Selecting a low value for α means that we are likely to make the correct conclusion when there is no real difference in the data. Alpha is conventionally set at .05 or less (Lipsey, 1990).

The value β is analogous to α and determines our likelihood to make a correct conclusion where there *is* a real difference in the data. The value β is the probability of making a Type II error (false negative) – concluding that there is no significant difference between two groups when in truth there is a difference. Conversely, $(1 - \beta)$ is the probability of correctly detecting a difference. The value $(1 - \beta)$ is called the statistical power. The higher this value the more likely we are to make a correct conclusion about a real difference in the data. Typical recommended levels for β are .80 or higher (Cohen, 1988; Muller and Benignus, 1992).

The statistical power is determined by four things: sample size, effect size, level of α, and the particular statistical test under consideration. We can choose a required value for β and determine values for the other parameters. In the discussion below, all power values are from tables and charts by Lipsey (1990). Other sources of power tables are available (Kraemer and Thiemann, 1987; Cohen, 1988), as well as statistical power software.

I performed a simple power analysis for the effect of FOV on selected dependent variables. I looked only at the power needed for detecting differences between two levels of FOV – 48° and 176°, rather than between combinations of the three levels. This simplification is suggested by Lipsey (1990, p. 70).

For each variable I did the following. First I computed the effect size (ES) present in the data I collected. The effect size is the difference between means divided by the standard deviation. Given ES and the number of samples, I estimated the power available in the current analysis for $\alpha = .05$ for a two-tailed t-test (Lipsey (1990, p. 91)). I used $\alpha = .05$ rather than .0125 because the effect of that difference on β is very small.

Next I recomputed N for a power of .80, for the same ES and α. The results are shown in the Table 5-2. The revised values of N here are the number of samples needed per FOV to show significance between 48° and 176° for the difference seen in the current study.

Table 5-2. Summary of power analysis.

	As tested			**Revised**		
Variable	ES	N	Power	ES	N	power
Search time	0.8	300	~1.0	0.8	27	0.80
Walking time	1.2	100	~1.0	1.2	13	0.80
SSQ total	0.15	15	0.08	0.15	> 500	0.80
SSQ Nausea	0.6	15	0.35	0.6	45	0.80
SSQ Disorientation	0.3	15	0.10	0.3	175	0.80
Presence	0.8	5	.20	0.8	27	0.80

The search and walking tasks had more power than needed; both those tasks could be done with fewer subjects or repetitions. The presence task could be replicated with about 6 times as many subjects, and this would give enough data to find whether the differences I saw were real or not. For sickness measures, it would be impractical to replicate the study with the numbers needed to show significance for the effect sizes I saw. Longer VR exposures might increase the effect sizes, making such an experiment feasible.

5.3 Comparisons with V8 HMD and restricted-real performance

In Sections 3.3 and 3.5.2 I described the secondary viewing conditions that were tested. Subjects used the V8 HMD with 48° FOV in one session and performed all tasks. Subjects wore the restricted-real display with 48° FOV for another session and performed the walking task. The following sections compare the results found with the V8 and restricted-real displays with the results found with the Kaiser HMD. Three performance measures are discussed: search, walking, and VR sickness.

5.3.1 Search

Data. Subjects performed the search task 80 times in the V8 display, the same number of times as for each of the three FOV conditions in the Kaiser HMD. The data were processed as in Section 4.4.1 (including the check for outliers, which yielded none). There were no missing values, so there were 300 measurements for the V8 display condition.

Figure 5-13 shows search performance with the V8 HMD compared to performance with the Kaiser HMD. On average, search time for the V8 at 48° was about the same as for the Kaiser at 176°. All four display conditions differed significantly, as described below.

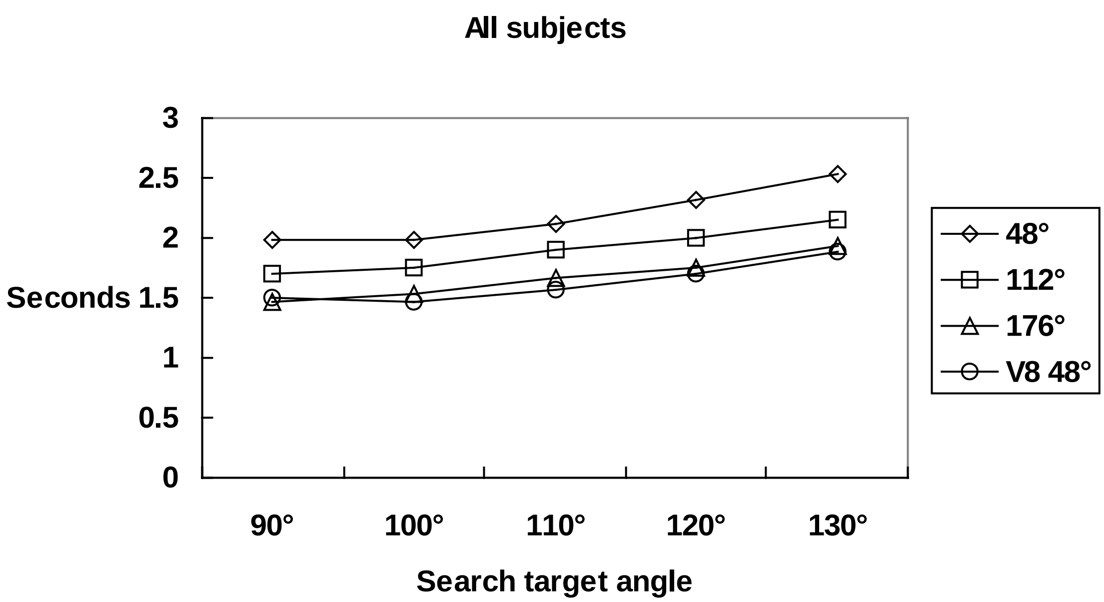

Figure 5-13. Search task performance across FOV in Kaiser and V8.

Analysis. A repeated measures ANOVA was performed, and a first model accounted for 89.2% of the variance. The dependent variable in this model was display type, and had four values (V8 48°, Kaiser 48°, Kaiser 112°, and Kaiser 176°). The independent variables were subject, FOV, and target angle, and all possible two-way interaction terms were included in the model (subject-by-FOV, subject-by-angle, and FOV-by-angle). The overall model statistics are $R^2 = 0.892$, $F(51, 1141) = 184.38$, $p < .0001$. The following independent terms were significant: subject ($F = 1390.92$, $p < .0001$), FOV ($F = 446.3$, $p < .0001$), target angle ($F = 194.79$, $p < .0001$), subject-by-FOV ($F = 131.76$, $p < .0001$), and subject-by-target-angle ($F = 7.21$, $p < .0001$).

As in Section 4.4.1, I successively dropped terms from the model in a step-down analysis. The resulting final model contained only main effects and no interactions, the same as the model found in 4.4.1. The statistics for that model (with display type as the dependent variable and subject, FOV, and target angle as the independent terms) are $R^2 = .729$, $F(11, 1181) = 288.09$, $p < .0001$. The independent terms were significant: subject ($F = 573.81$, $p < .0001$), FOV ($F = 184.12$, $p < .0001$), and target angle ($F = 80.36$, $p < .0001$).

Post-hoc Bonferroni t-tests showed that the difference was significant between all pairs of display types. That is, performance with the V8 display was significantly better than performance with the Kaiser display at 48°, 112°, or 176°. For target angle, the differences were significant between all pairs except between 90° and 100°.

Summary. Subjects' performance improved with the V8 display, either because of practice effects (subjects did the V8 session after all three of the Kaiser sessions had been completed) or because of the V8's better visual quality, lower weight, lower moment of inertia, and lower latency compared to the Kaiser HMD. The data from the current study are not sufficient for telling which factors led to the improvement.

5.3.2 Walking

Data. Subjects performed the walking task once in the V8 display, giving 20 trials plus 10 practice. In the restricted-real display they performed the walking task twice, but the first administration was dropped because it took that long for the experimenters to get practiced with handling the cables to the display. Recall that in the restricted-real condition the real styrofoam walls were present and the subject wore the mock HMD with HiBall tracker attached. Because of the walls the experimenters had to follow the subject closely along their path and hold cables over the walls. The restricted-real condition gave 20 trials plus 10 practice. The data were processed for outliers as described in Section 4.4.2 (four outliers were removed).

Figure 5-14 shows walking task performance with the V8 and with the real-world restricted FOV display. The analysis described below found that walking performance with the V8 was significantly better than the Kaiser HMD at 48° or 112°, and that walking performance with the restricted-real display was significantly better than all viewing conditions using the Kaiser or the V8 HMDs.

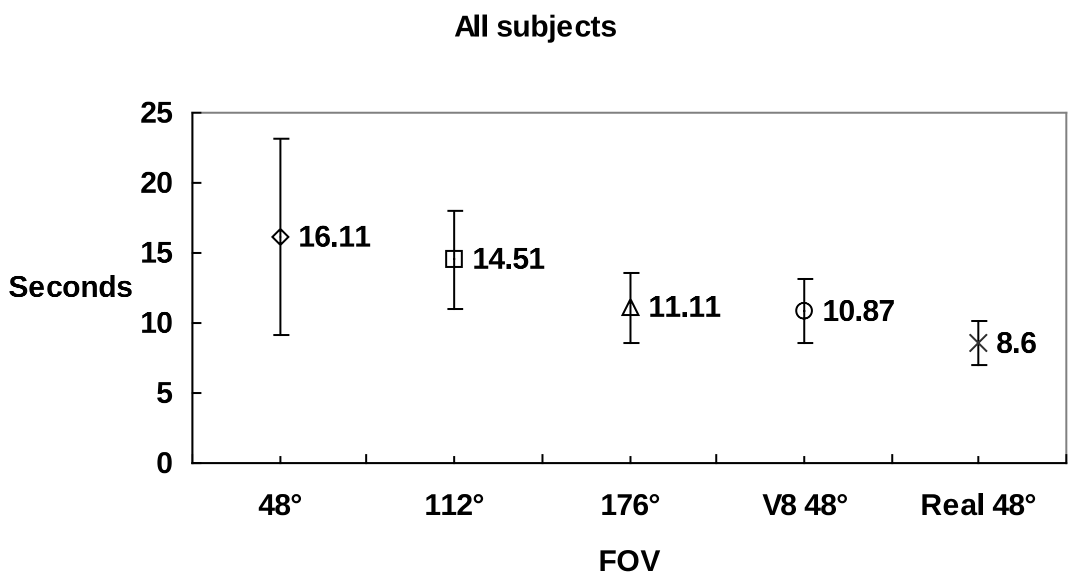

Figure 5-14. Walking task performance for Kaiser, V8, and restricted-real conditions.

Analysis. A repeated measures ANOVA was performed with display type as the dependent variable. It had five values (Kaiser 48°, Kaiser 112°, Kaiser 176°, V8 48°, and Restricted-real 48°). This analysis corresponds closely to the analysis in Section 4.4.2. A model with subject, FOV, and the subject-by-FOV as independent terms accounted for 89.5% of the variance. The model statistics are: R^2 = .895, F(23, 445) = 165.01, p < .0001). All independent terms were significant: subject (F = 365.98, p < .0001), FOV (F = 354.06, p < .0001), and subject-by-FOV (F = 61.0, p < .0001). Section 4.4.2 explains the reasons for the subject-by-FOV interaction.

Post-hoc Bonferroni t-tests showed that performance differed significantly between all pairs of the five display types, except between V8 48° and Kaiser 176°.

Summary. Walking performance with the V8 display was better than with the Kaiser HMD at 48° or 112°; reasons for this may include practice or improvements in display quality, as discussed in the previous section. Walking performance with restricted-real display was better than in any of the V8 or Kaiser HMD conditions. This improvement may again have resulted from practice (the restricted-real session was performed after the V8 session) or from better display quality. The restricted-real display had better visual quality, lower weight, and lower latency than the V8 HMD.

5.3.3 VR Sickness

The simulator sickness questionnaire was administered after VR exposures in the V8 HMD and restricted-real display. Figure 5-15 summarizes the total severity scores from the SSQ. The values seen for the V8 HMD are comparable to those found for the Kaiser display, and sickness again appears to increase with successive exposures, though this difference is not significant (only 15 measures were available, 3 for each subject). The total severity reported for the restricted-real condition is after the second exposure. The first exposure data was dropped as practice, as described in Section 5.3.2. Sickness with the restricted-real display appears to be lower on average than for the Kaiser or V8 displays, though this difference was not significant. Some subjects commented after the restricted-real session that they felt dizzy, indicating that some sickness was indeed present.

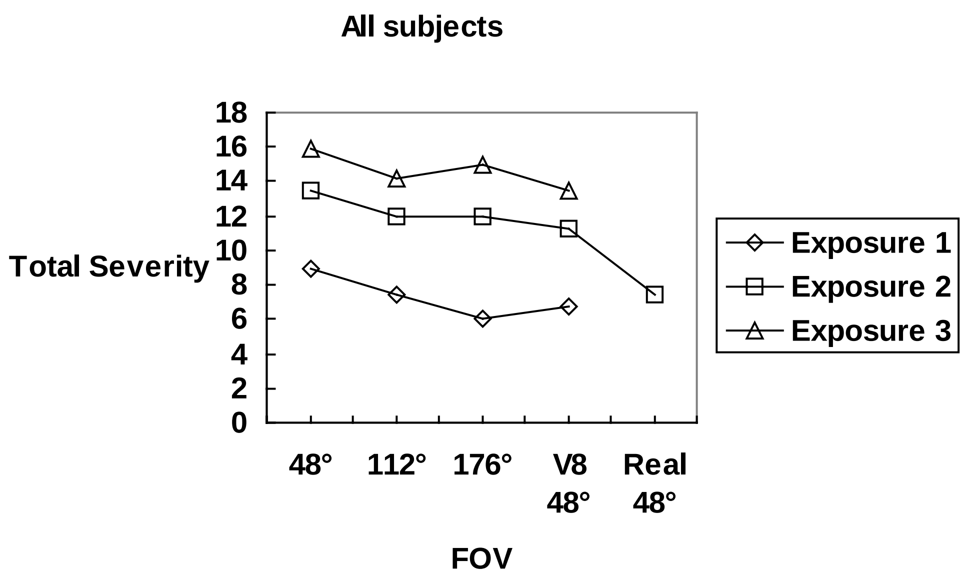

Figure 5-15. Total Severity of sickness showing V8 and restricted-real conditions.

5.4 Subject comments

Appendix A.11 contains the written comments given by subjects at the end of the presence questionnaire. They were asked: “Please enter below any additional comments you wish to make about the virtual environments that you experienced today.” The following are typical comments.

- On performing better with wider FOV: “was able to perform tasks much easier” (subject A at 176°), “additional screens helped” (B at 112°), “a lot more real” (B at 176°), “much harder to find pyramid with limited vision … less able to survey field/room” (E at 48°), “much easier staying w/in the maze the 2nd day (E at 176°),

- One subject noted feeling more disoriented with wider FOV: “I felt more dizzy” (subject A at 176°),

- Subject E thought “limited vision helped me in the maze a bit… because of less white space surrounding” and did in fact perform better at 48°, whereas other subjects did not.

- Several comments were made noting the better visual quality of the V8 display compared to the Kaiser display.

Chapter 6

Future Work

Restricting FOV degraded performance on searching and walking tasks at the three levels of horizontal FOV I tested. This chapter describes some remaining unanswered or untested research questions about FOV in HMDs. In Section 6.1 I consider additional aspects of the independent variable, FOV. In Section 6.2 I consider better tasks and better measures of the dependent variables.

6.1 Other Aspects of FOV

6.1.1 Vertical FOV

The present study varied horizontal FOV and neglected vertical FOV, but restricting the vertical FOV will also degrade performance, though perhaps less. All HMDs restrict the vertical FOV to much less than our normal 135°. The vertical FOV of the V8 HMD, for example, is 36°.

Vertical FOV (VFOV) differs in nature from horizontal FOV. Our visual world is basically symmetric from side to side, but it is not symmetric from top to bottom. We use our lower and upper visual fields differently. In particular, our lower visual field is where we usually see our body, or our virtual body in VR, and see our feet hitting the ground. For this reason we should consider lower VFOV and upper VFOV as separate variables, at least with regard to some tasks. The lower VFOV, below the horizon, is probably most interesting for travel tasks. For headcentric search tasks or memory tasks, the distinction between lower and upper VFOV is probably not as important. The upper and lower VFOV are usually equal in size in most HMDs (the VFOV is centered on the pupil).

Slater et al. (1995) have shown that providing a virtual body in VR is important for presence, and that the more users are aware of their virtual body the higher their sense of presence. It follows that the larger the lower VFOV, the more presence the user will experience, because they see their virtual body more. This hypothesis could be tested by conducting a study with independent variables for lower-VFOV size and virtual body (present, absent, or type of representation) and with presence as the dependent variable.

To test the effects of lower VFOV on travel performance, the same walking task implemented here could be used. Virtual body (presence, absence, and whether the feet move) might also have an effect on walking performance.

Of the other main tasks I implemented – distance memory, spatial memory, and search – only the search task is inherently horizontal and inappropriate for use in a study of vertical FOV.

6.1.2 Interaction of FOV with other display qualities

Subjects performed just about as well with a narrow-FOV V8 display or "restricted-real" display as with the wide-FOV Kaiser display. Practice was probably part of the reason for this (they used the Kaiser first). Whether or not practice was responsible could be determined with a study that balanced the order of display conditions. Some part of the improvement was probably caused by the quality of the V8 and restricted-real displays. They were better than the Kaiser HMD in display resolution, color, contrast, brightness, latency, tiling artifacts, weight, moment of inertia, fan noise, fit and comfort. There was also a difference in subject's impressions of the displays and this may have introduced bias – subjects said they preferred using the V8. I hypothesized that the effects of FOV are largely independent of these other display qualities; addressing this question requires new studies with better quality wide-FOV displays.

6.1.3 FOV in Augmented HMDs

If a practical wide-FOV HMD is to be realized, it is likely to require a design that divides the FOV into regions with different resolutions. Such designs increase the FOV without seriously decreasing the frame rate. The techniques include moving area-of-interest (Yoshida et al., 1995), fixed area-of-interest (Watson et al., 1997), area-of-interest in one eye (Kaiser Electro-Optics, 2000), and add-on peripheral displays (Monterey Technologies, 1999). The last approach offers perhaps the largest increase in FOV because it avoids the problem of designing wide-FOV optics for a single image plane. Rendering for the add-on peripheral displays might be offloaded to a PC worn by the user. If that PC has direct access to head tracker data it could generate low-latency horizon cues for the peripheral display that would benefit travel performance. For all of these designs, studies are required to determine the optimum FOV sizes for the different regions and to assess the tolerance for artifacts at region boundaries.

6.2 Extensions to Dependent Variables

6.2.1 Presence

Restricting FOV probably reduces presence. In this study I did not establish statistically significant confirmation of that hypothesis, but I did see a consistent, non-significant trend. Mean total presence, measured by the Witmer and Singer questionnaire, decreased with restricted FOV. This could be verified with a larger sample size, approximately 6 times the number of measurements taken here. A follow-up study should start with a more presence-inducing virtual environment to induce a larger effect size, as was done by Usoh et al. (1999).

Long presence questionnaires are not the best way to measure presence in a repeated measures study. The Witmer and Singer questionnaire that I used, and most other presence questionnaires, are usually given to each subject only once. I gave the questionnaire to each subject after each session (i.e., five times in total, once for each display condition). Giving a long questionnaire more than once introduces the possibility of bias; subjects may answer differently on second and third administrations because they don't reread the questions or because they become tired of the questionnaire. To reduce the chances of this, I familiarized the subjects with the questionnaire before their first session so that all administrations would be repeats. A better way to measure presence in repeated measures studies might be with short checklists (similar to the format of the Simulator Sickness Questionnaire), rating scores (Dinh et al., 1999), breaks in presence (Slater, 1999), or physiological measures.

6.2.2 VR Sickness and postural instability

As with presence, a prerequisite for measuring variations in sickness requires first that we use a sickness-inducing virtual environment. Current VR systems as used in this study are not creating serious levels of sickness for typical short exposures (5 to 10 minutes). Given these low sickness levels, it seems unlikely that FOV will have appreciable effects on sickness during typical exposures. FOV effects on sickness might be reproduced in an artificial setting like those used in previous studies of vection (Stern et al., 1990), but those settings are not likely to arise in practice in VR. So and Lo (1999) made a similar assessment, that viewing a stationary virtual environment for 15 minutes or less is not likely to induce nausea.

Using head tracking to measure head movement for postural stability analysis is undoubtedly better than videotaping and manually analyzing the tape frame-by-frame. But finding good measures of postural instability in VR requires again that we first find an instability-inducing environment. The environment for the present study was probably not sufficient for that purpose; even with improved analysis (using, for example, frequencies or acceleration data), any instability present was probably smaller than could be detected. There was no indication, either from the translational analysis or from direct observation of subjects, that subjects were unusually unstable in their footing after VR exposure.

6.2.3 Distance estimation and spatial memory

The distance estimation and spatial memory tasks implemented here failed to produce any answers to the hypotheses about restricted FOV causing misperceived distance or size. If these tasks were to be used again they should be administered in separate sessions (distance memory separate from spatial memory) with new, simpler stimuli. The variability in the stimuli could be reduced by, for example, placing the circles along one dimension and using a more symmetric virtual environment (the styrofoam wall layout was not symmetric).

A simpler and cleaner experiment to assess errors in judgments of space might take the “room size estimation” question (question #33 on the presence questionnaire I used) and vary the virtual room size in a randomized design. This would be closer to methods used by Alfano and Michel (1990) in their real-world study. The dependent variable here would be the subject’s estimate of (virtual size/real size), and the independent variable would be the actual value of that ratio. A second interesting independent variable would be the degree of subject motion: fixed view, turning only (as in Alfano’s study), or full motion (walking). It’s possible that more motion would result in less perceptual error. An alternate method would be to allow the subject to adjust the virtual room size with a dial until they felt it matched the size of the real room.

.

Appendix A
Data summaries

A.1 Subjects

The questions corresponding to the fifth and sixth columns in Table A-1, asked by the experimenter during screening sessions, were:

- "Have you used a virtual reality system before?" (never, 3 times or less, or more)
- "How often do you use video games?" (monthly, weekly, or daily)

Table A-1. Subject descriptions.

Subject	Gender	Age	Hand	VR Use	Games
A	M	21	R	Never	Weekly
B	M	21	R	Never	Daily
C	F	22	R	Never	< monthly
D	F	22	R	3 or less	< monthly
E	F	30	R	Never	< monthly

A.2 Search data

The plots show averages over two blocks of 20 trials for each FOV (they do not include 20 practice trials).

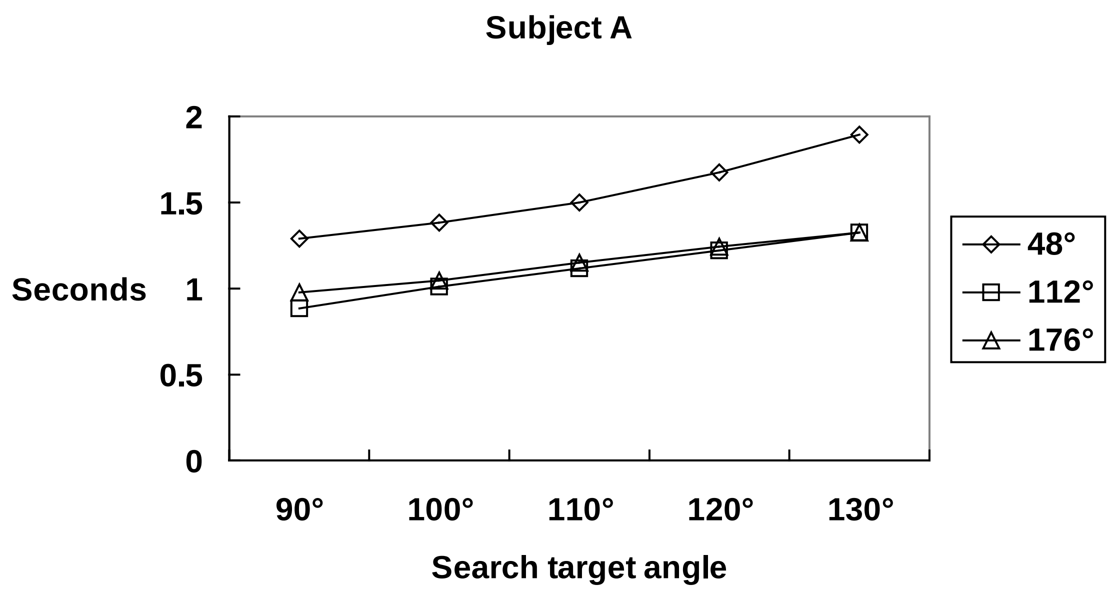

Figure A-1. Search task performance for subject A.

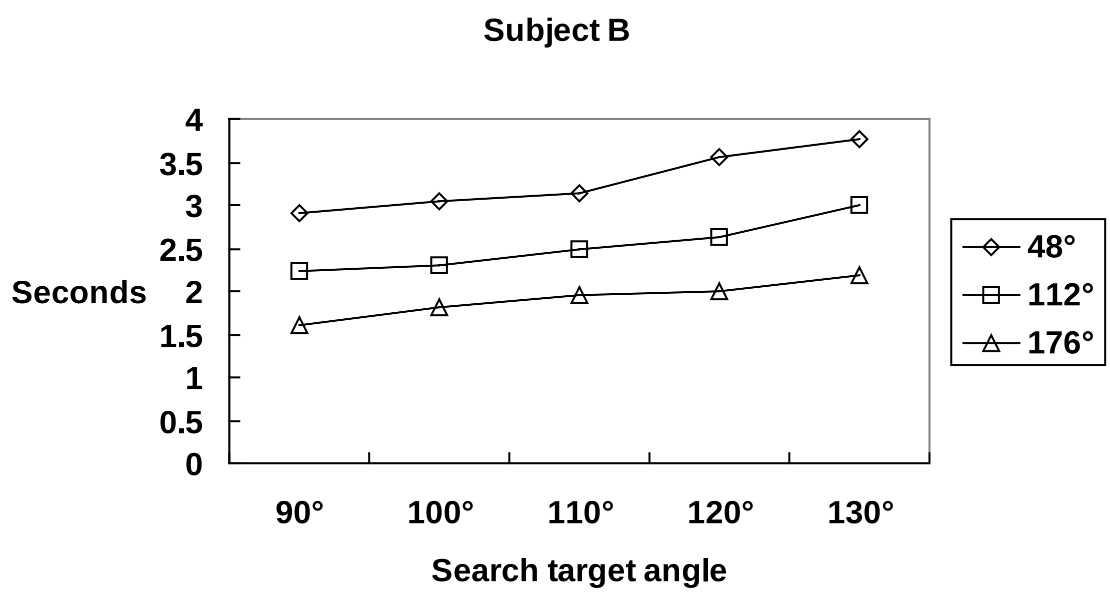

Figure A-2. Search task performance for subject B.

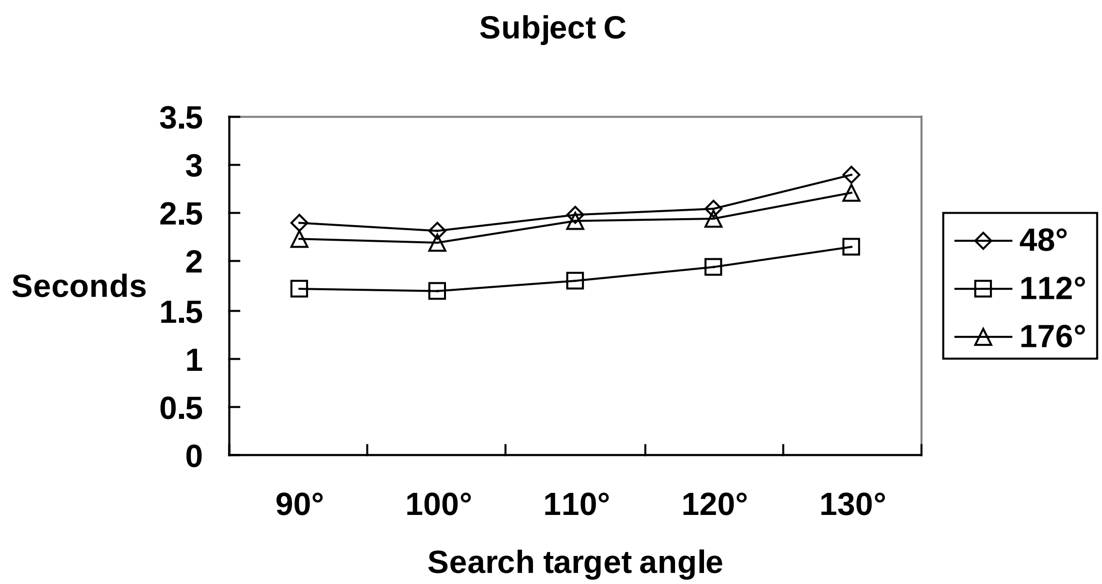

Figure A-3. Search task performance for subject C.

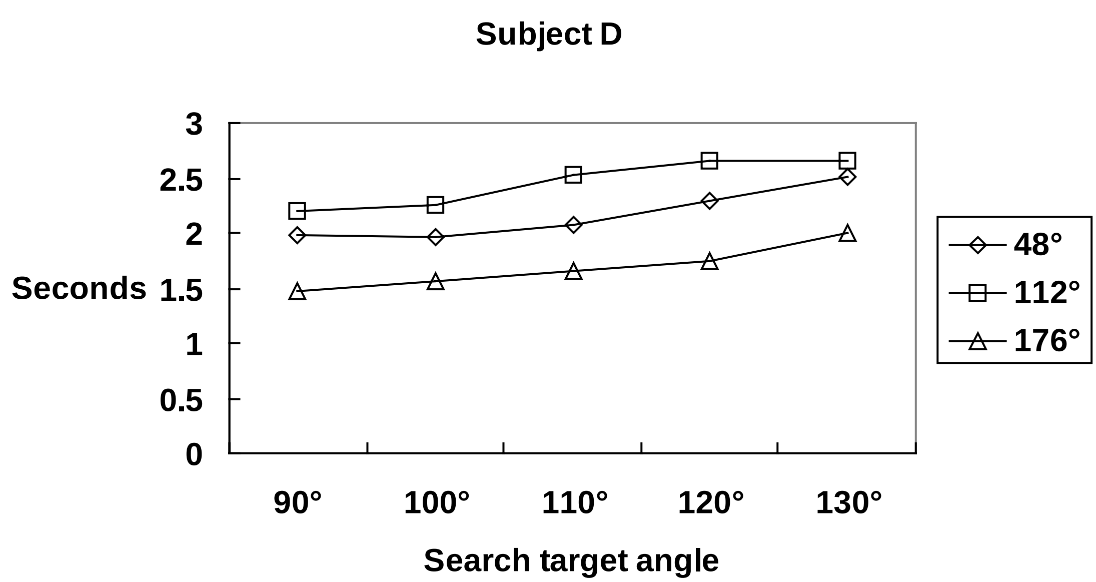

Figure A-4. Search task performance for subject D.

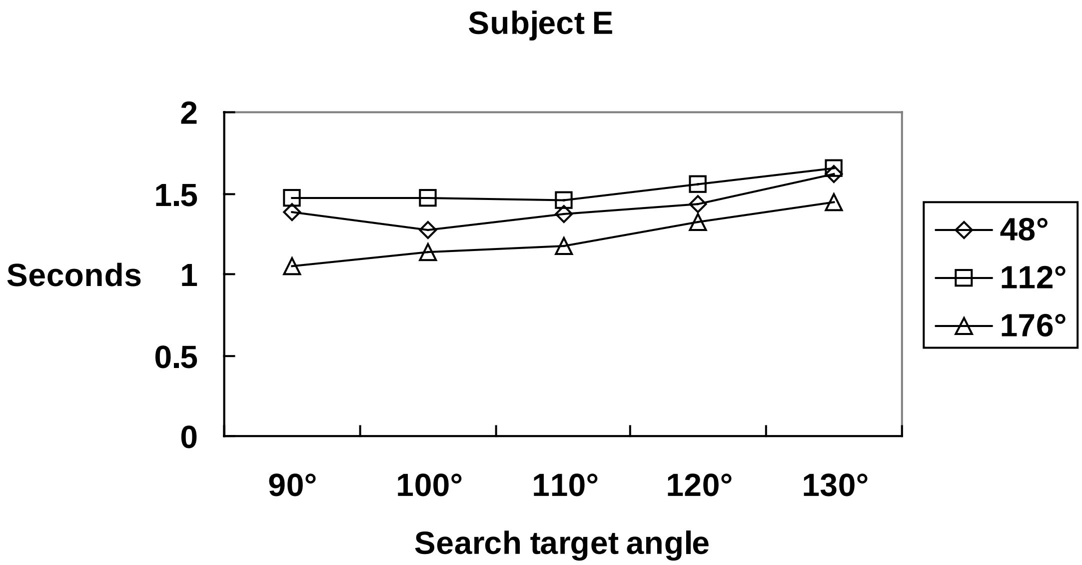

Figure A-5. Search task performance for subject E.

A.3 SAS Code for search time analysis

The following code was used to perform the repeated measures ANOVA for search time. Code for the other main analyses is similar.

Table A-2. Sample SAS code for search time analysis.

```
title ’model 1, with interactions’;
proc anova;
  class subj fov angle;
  model dt = subj fov angle subj*fov subj*angle fov*angle;
  means subj fov angle / bon;
run;

title ’model 2, drop fov*angle’;
proc anova;
  class subj fov angle;
  model dt = subj fov angle subj*fov subj*angle;
  means subj fov angle / bon;
run;

title ’model 3, drop subject*angle’;
proc anova;
  class subj fov angle;
  model dt = subj fov angle subj*fov;
  means subj fov angle / bon;
run;
```

```
title 'model 4, drop subject*fov (no interactions)';
proc anova;
  class subj fov angle;
  model dt = subj fov angle;
  means subj fov angle / bon;
run;
```

A.4 Walking data

The following plots show data points for two blocks of 10 trials for each FOV. They do not include the earlier block of 10 practice trials for each FOV, and do not include outliers (data outside 3 standard deviations from the median were removed from the original raw data).

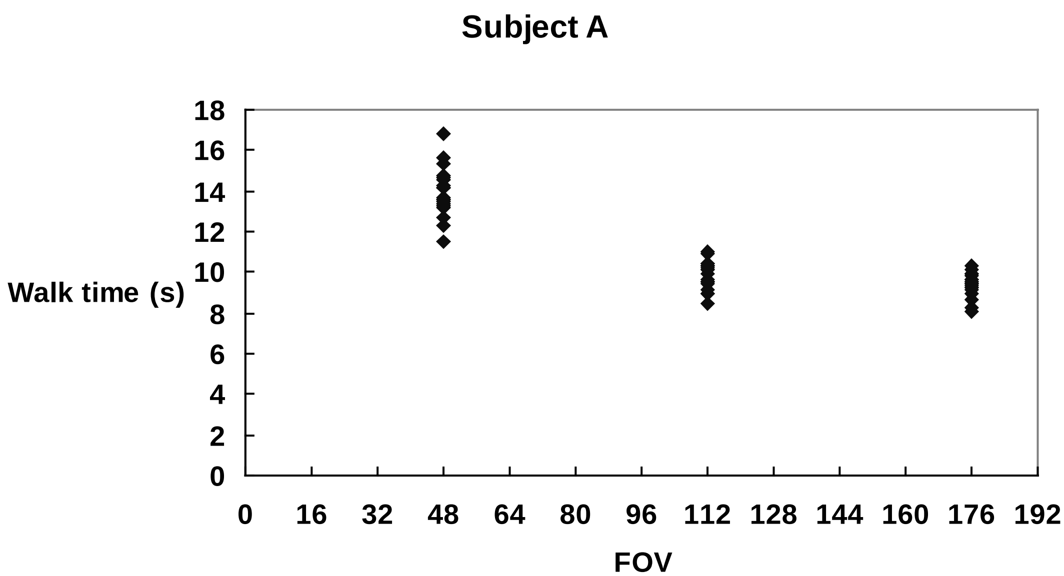

Figure A-6. Walking task performance for subject A.

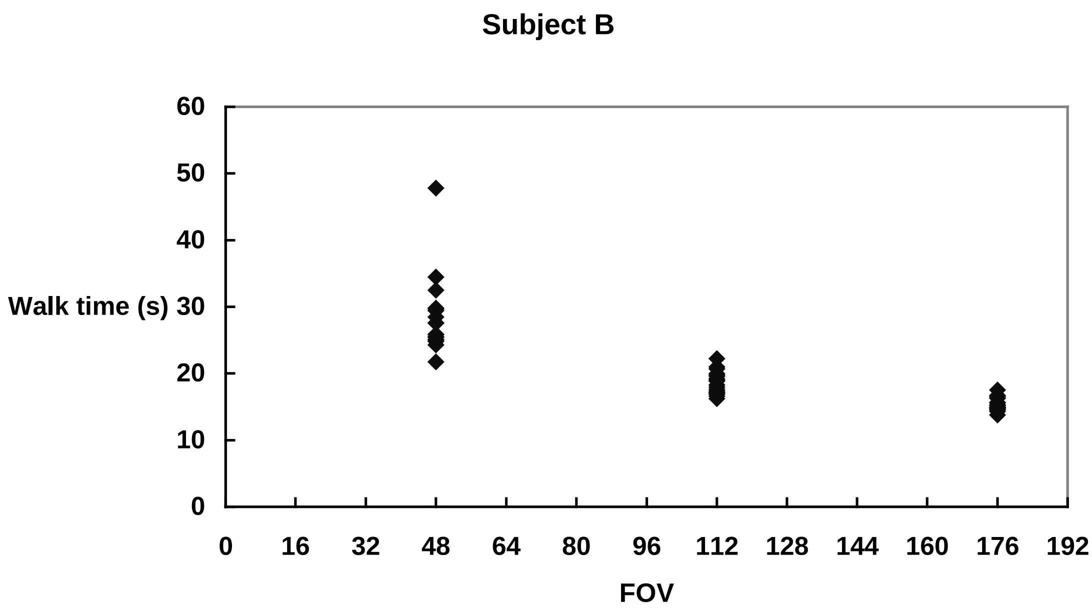

Figure A-7. Walking task performance for subject B.

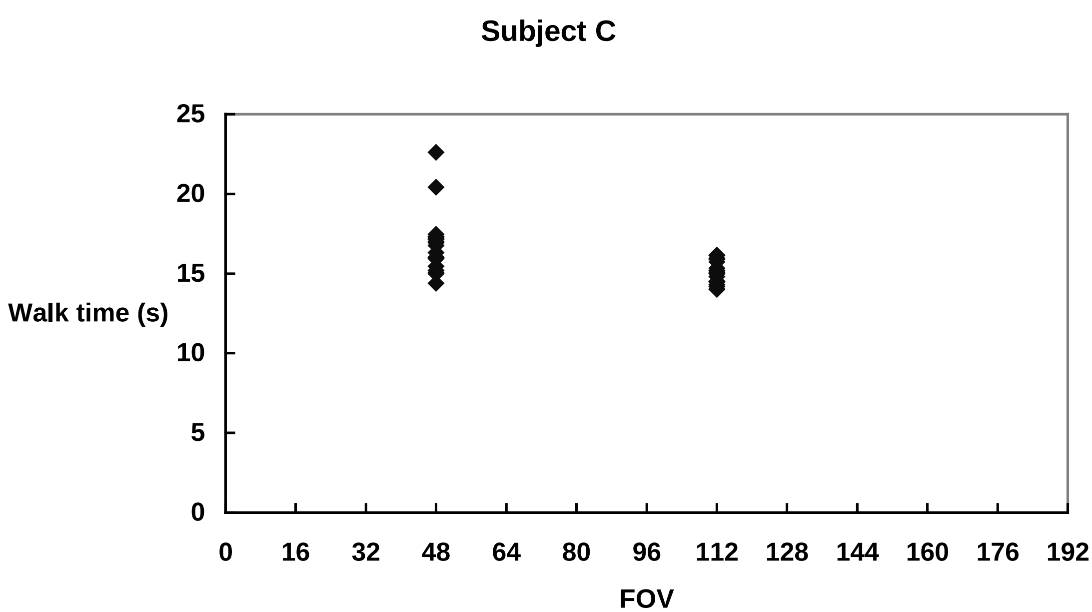

Figure A-8. Walking task performance for subject C.

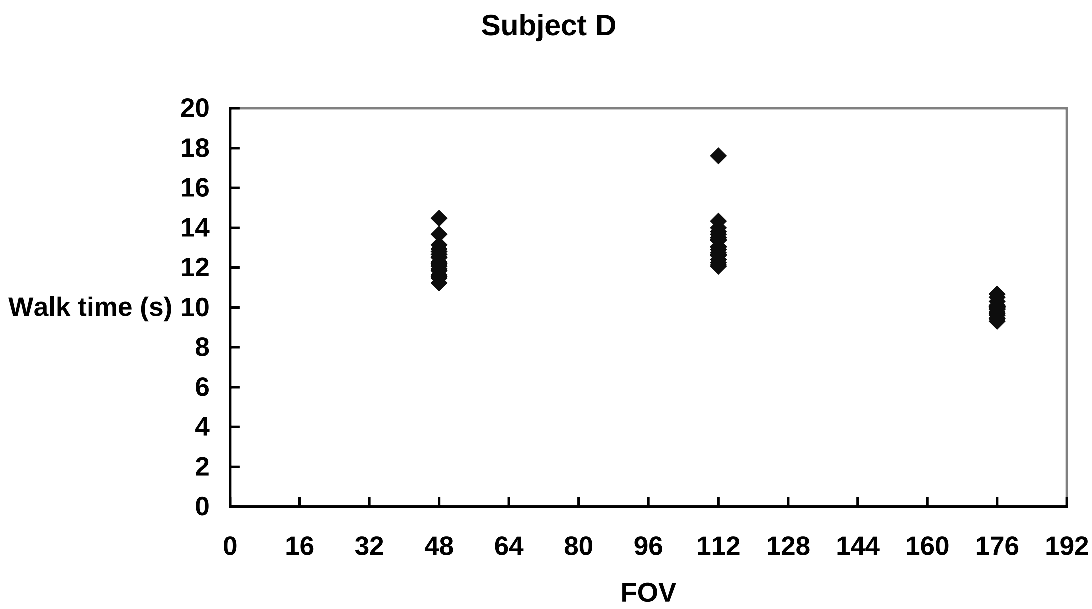

Figure A-9. Walking task performance for subject D.

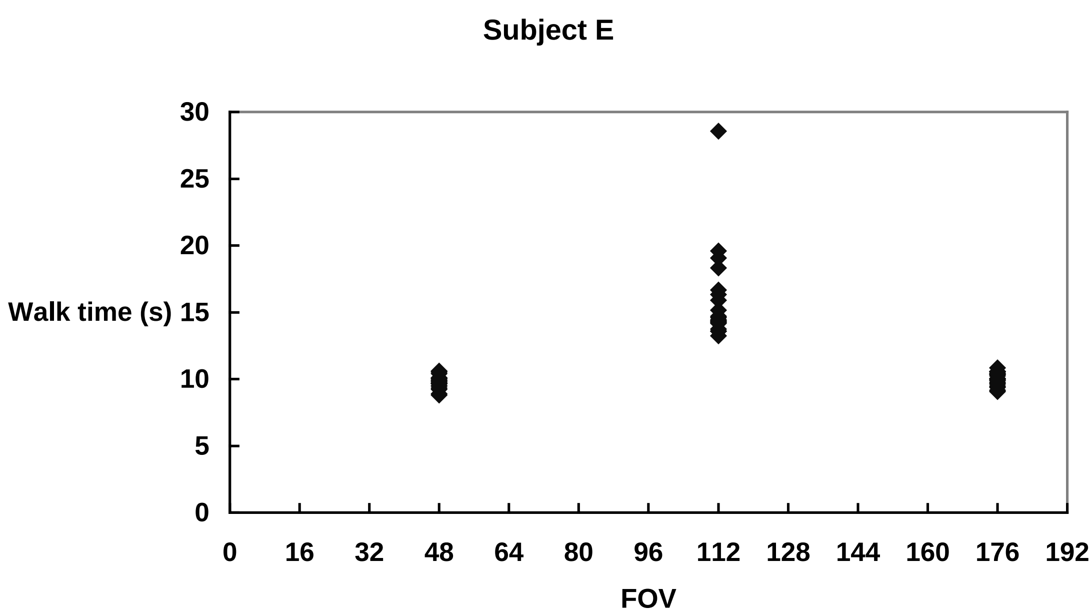

Figure A-10. Walking task performance for subject E.

A.5 VR Sickness data

The plots below show the Total Severity scores from the SSQ. Each subject's test yielded up to 9 measures, 3 at each FOV, once after each VR exposure (Test 1 after the search exposure, Test 2 after the walking exposure, and Test 3 after the memory exposure).

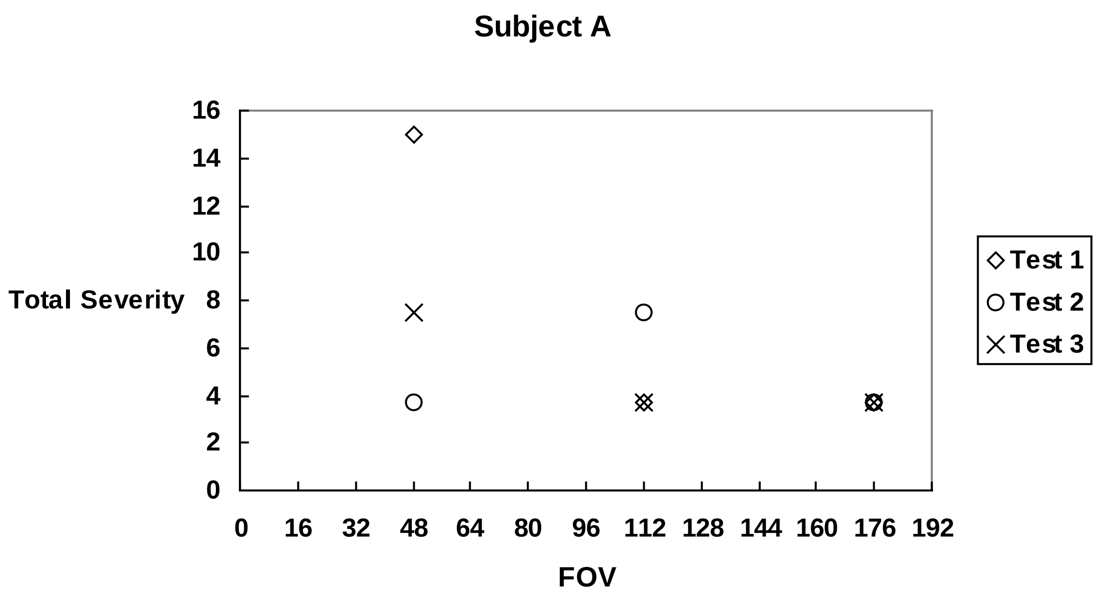

Figure A-11. Sickness scores for subject A.

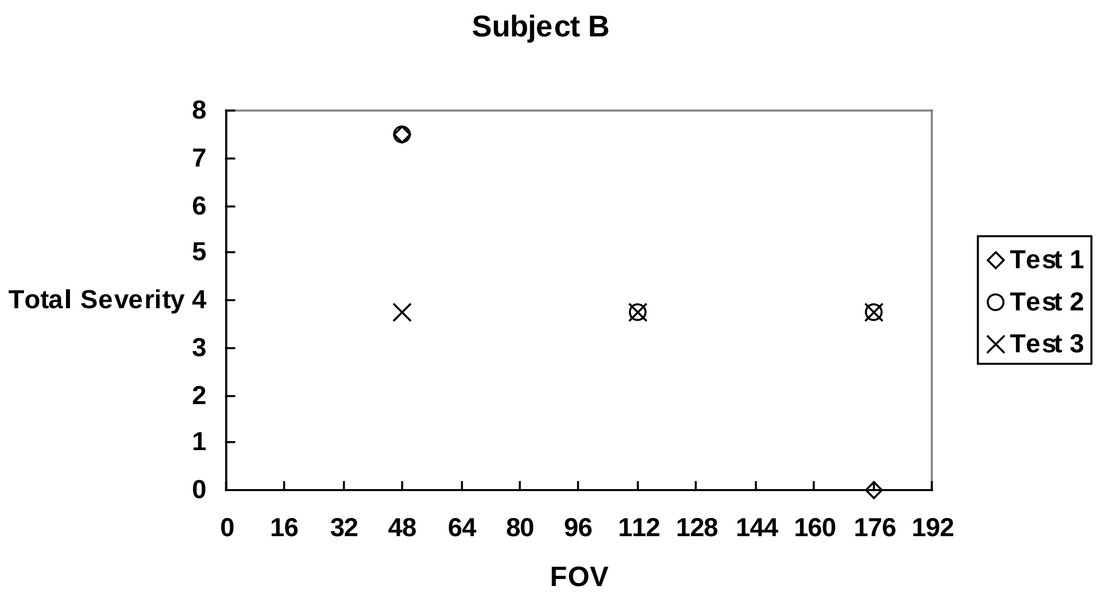

Figure A-12. Sickness scores for subject B.

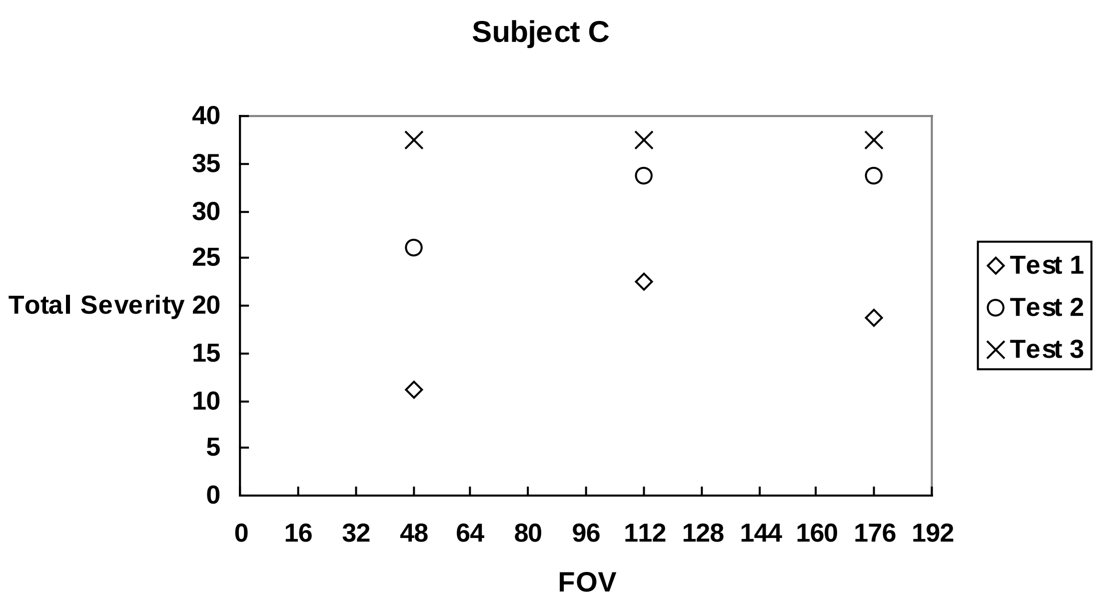

Figure A-13. Sickness scores for subject C.

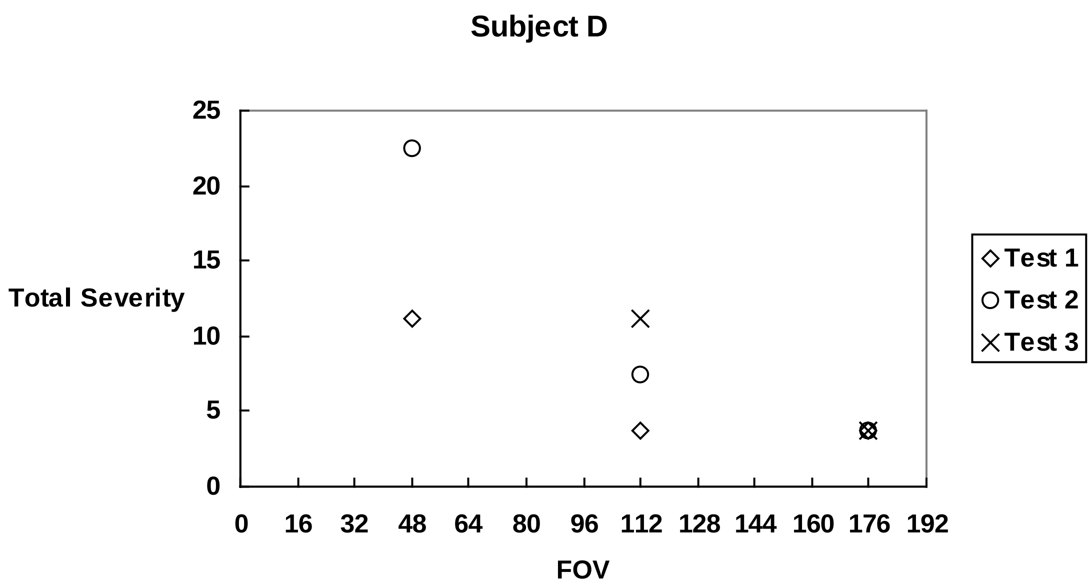

Figure A-14. Sickness scores for subject D.

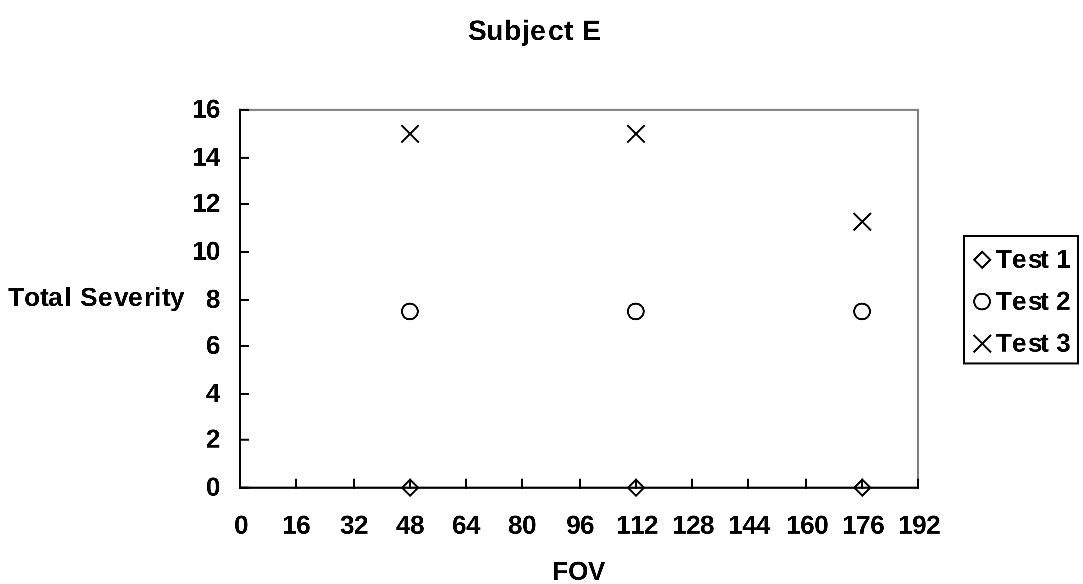

Figure A-15. Sickness scores for subject E.

A.6 Postural instability data

The postural instability measure was average head movement velocity over 20 seconds. The test was performed before and after each of the three VR exposures, giving 2 measures for 3 sessions, 6 measures in total.

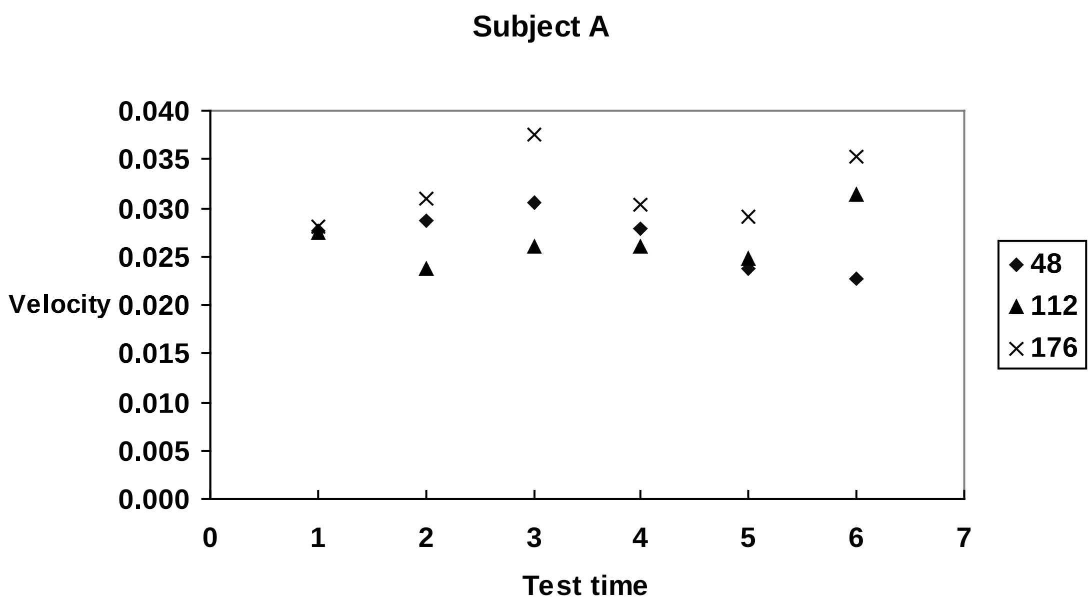

Figure A-16. Postural instability measures for subject A.

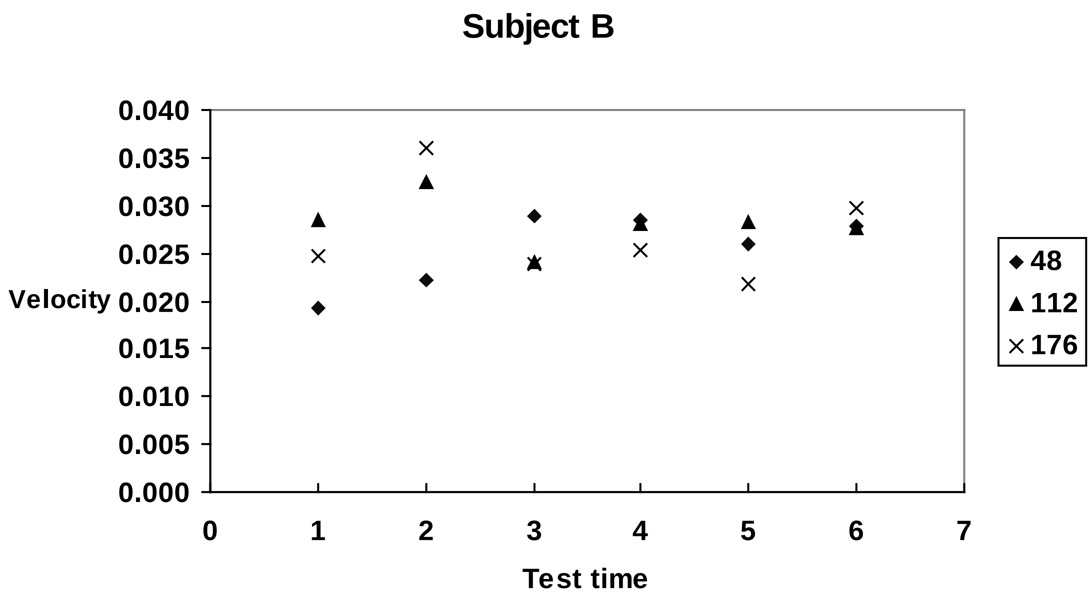

Figure A-17. Postural instability measures for subject B.

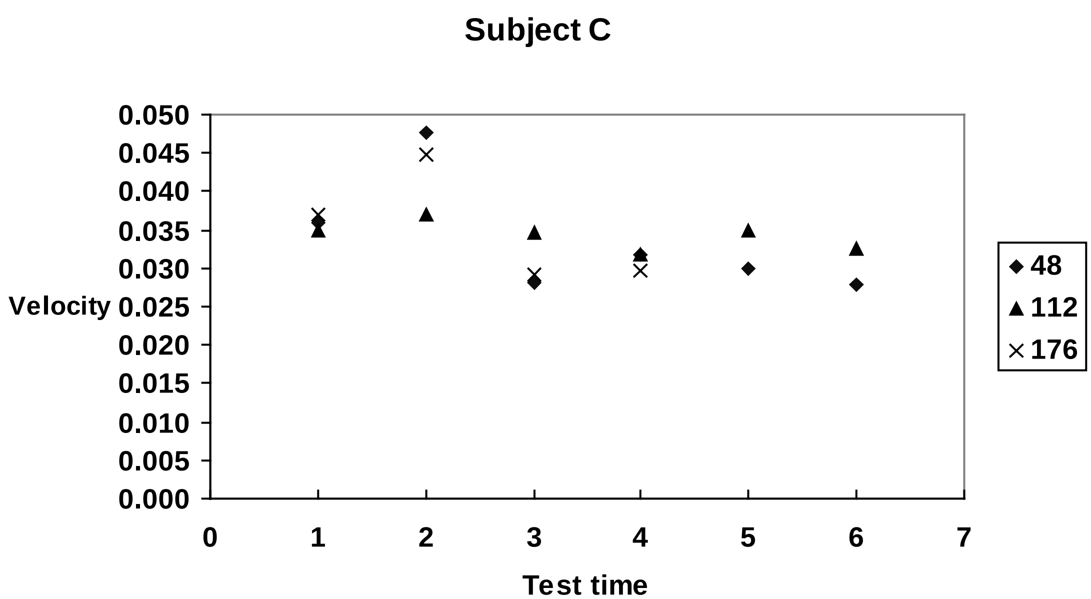

Figure A-18. Postural instability measures for subject C.

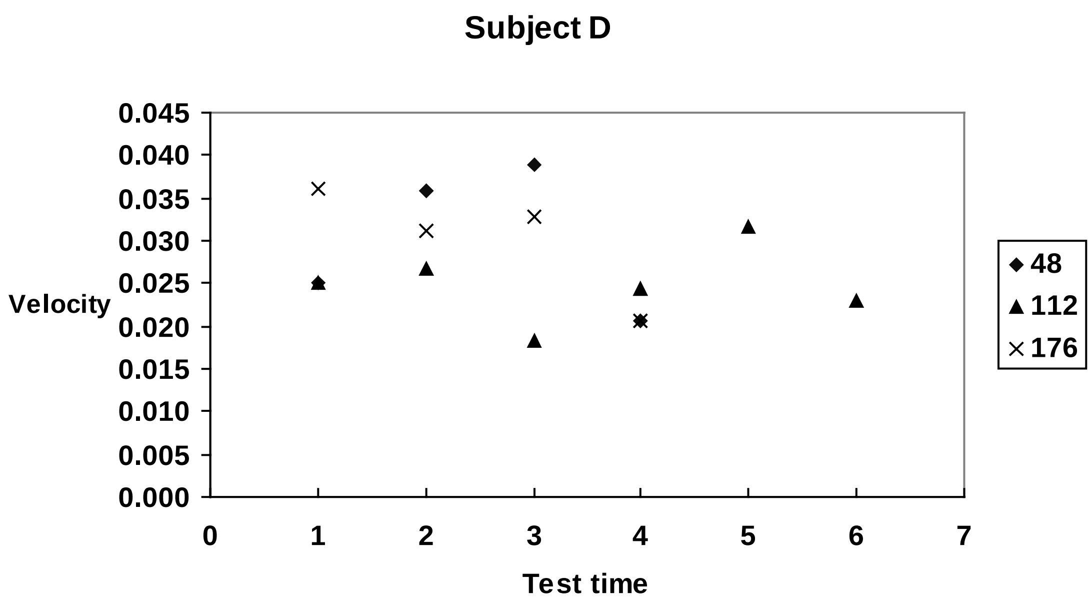

Figure A-19. Postural instability measures for subject D.

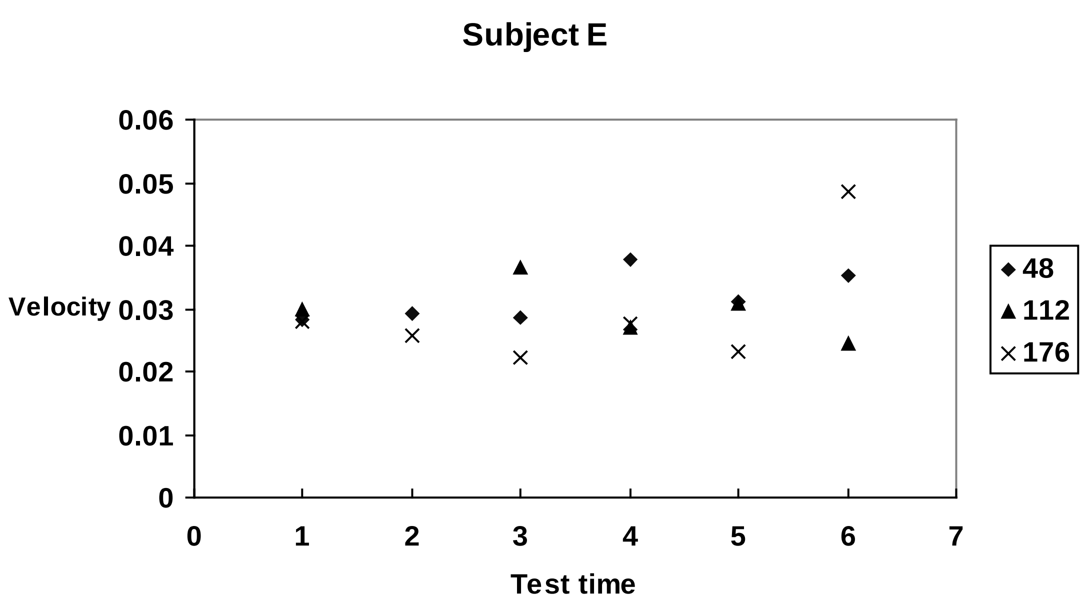

Figure A-20. Postural instability measures for subject E.

A.7 Distance estimation data

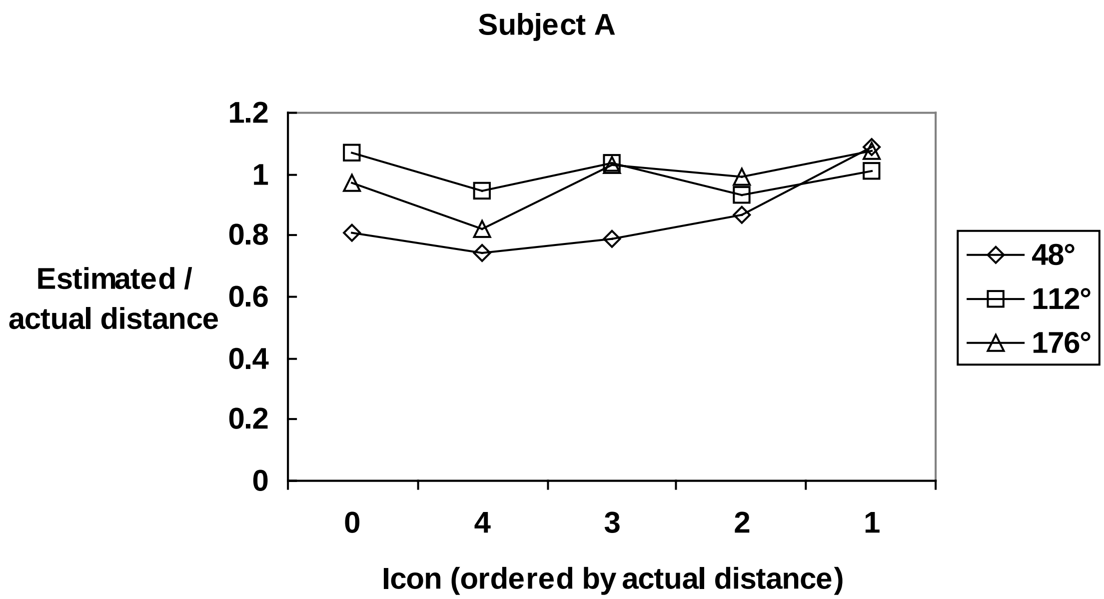

Figure A-21. Distance estimation accuracy for subject A.

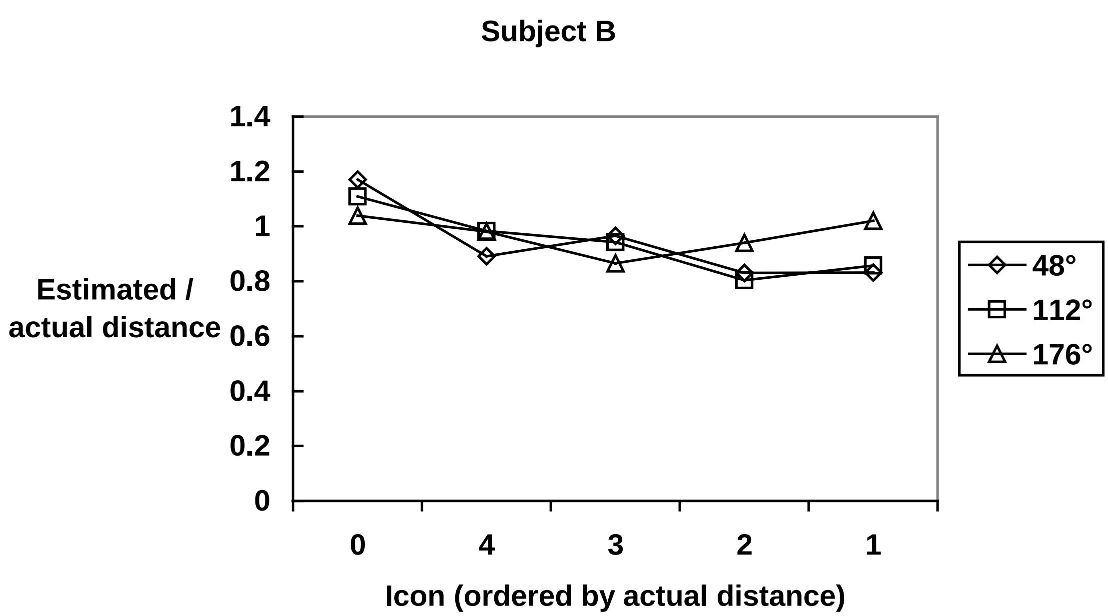

Figure A-22. Distance estimation accuracy for subject B.

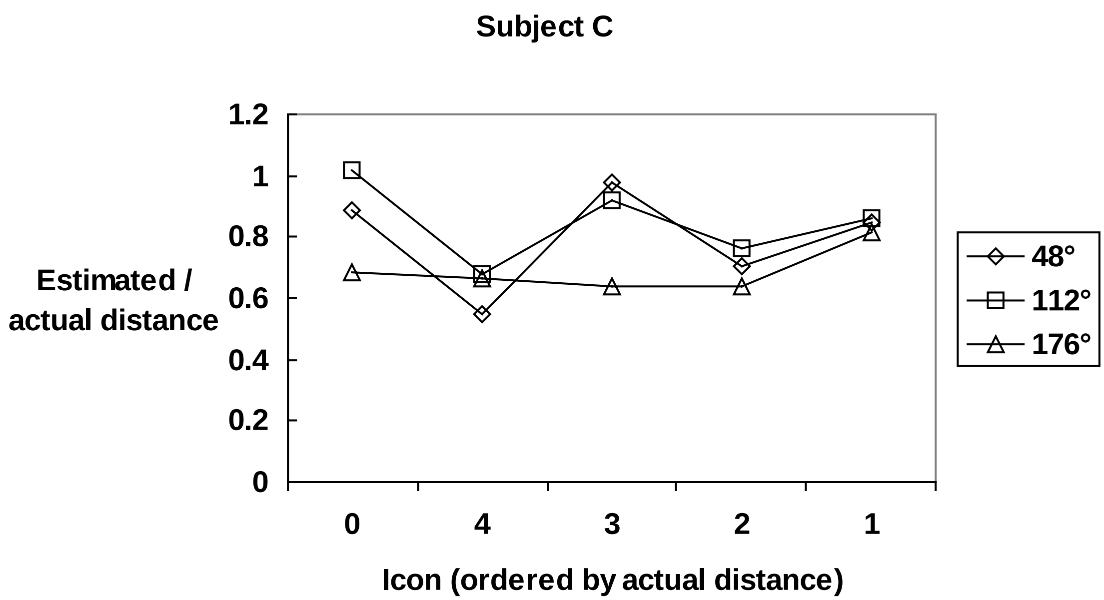

Figure A-23. Distance estimation accuracy for subject C.

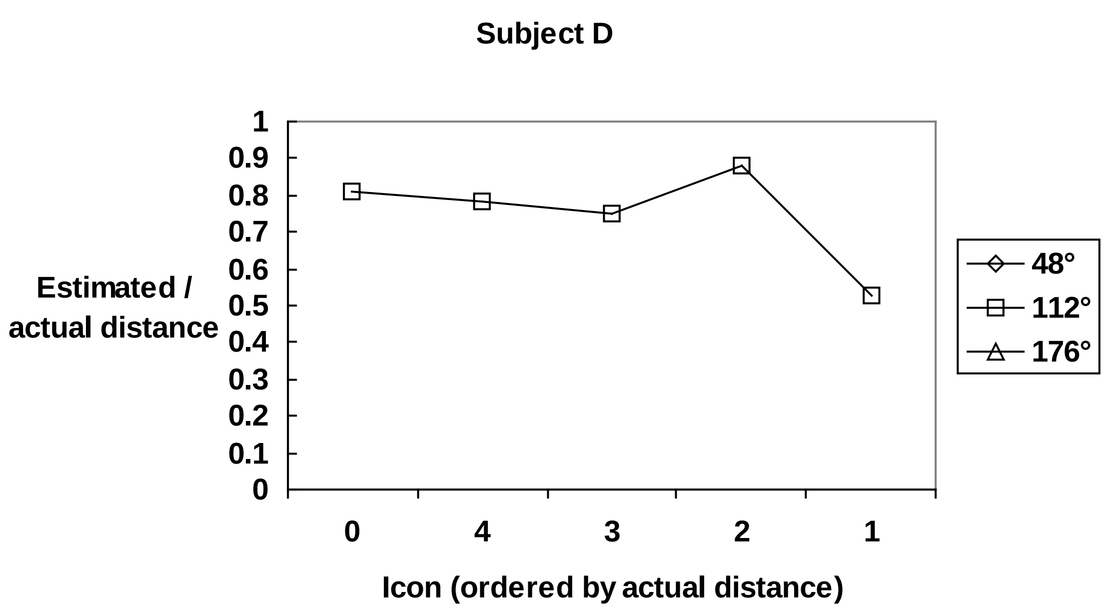

Figure A-24. Distance estimation accuracy for subject D.

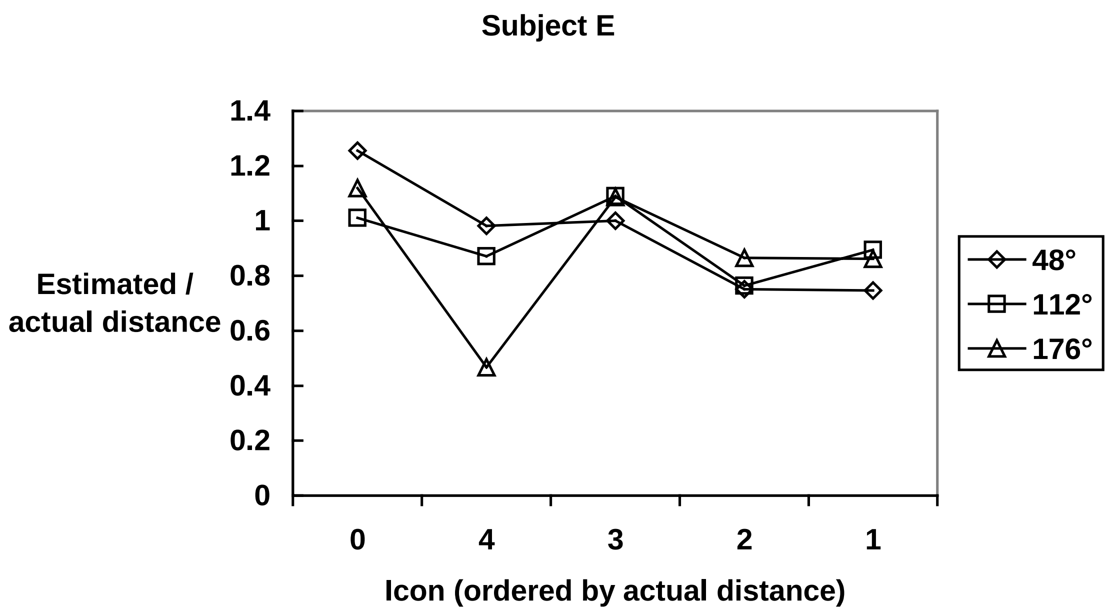

Figure A-25. Distance estimation accuracy for subject E.

A.8 Spatial memory data

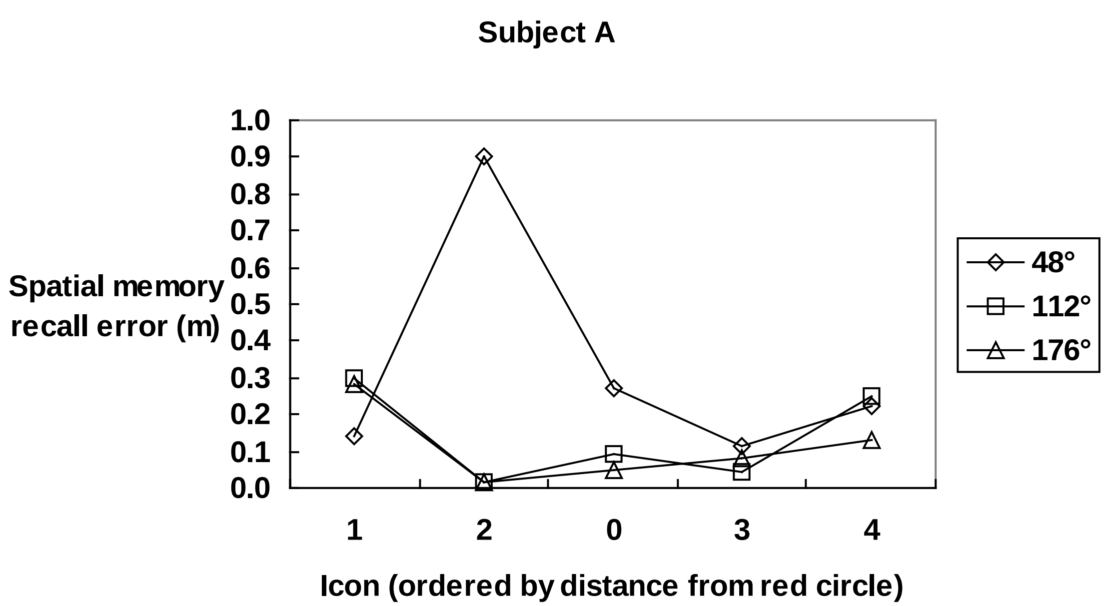

Figure A-26. Spatial memory error for subject A.

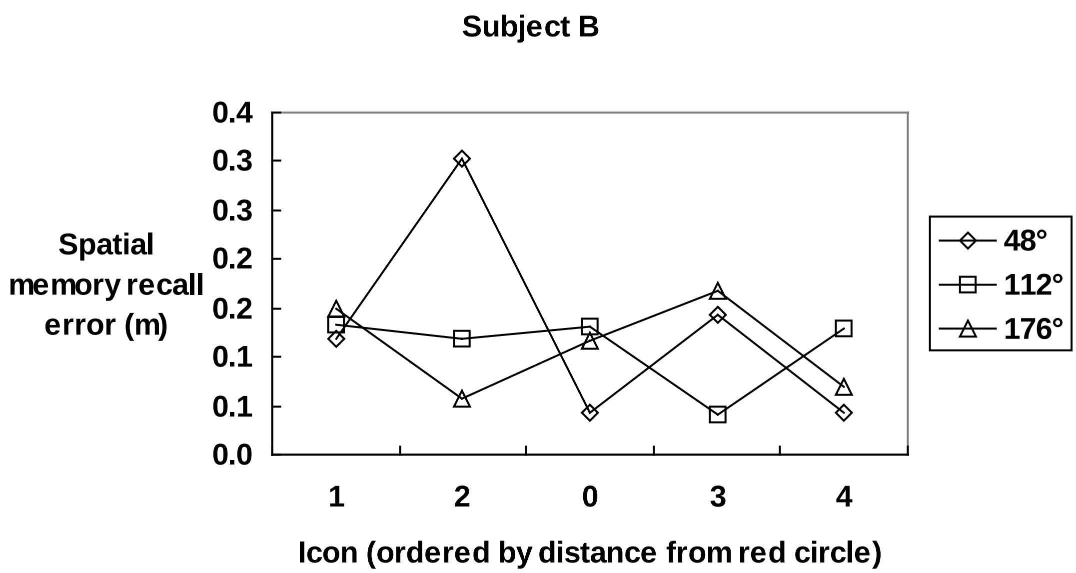

Figure A-27. Spatial memory error for subject B.

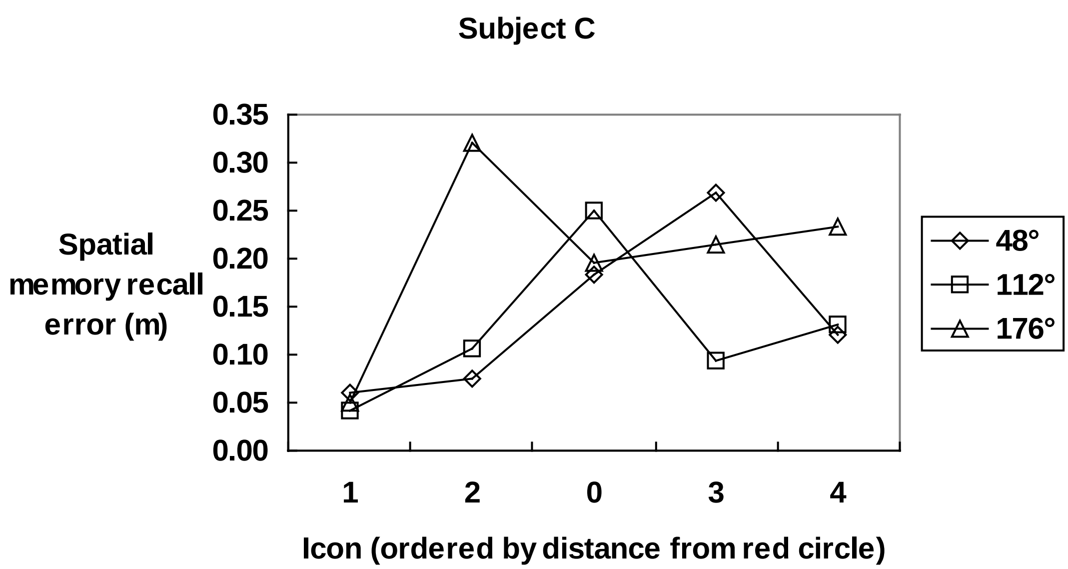

Figure A-28. Spatial memory error for subject C.

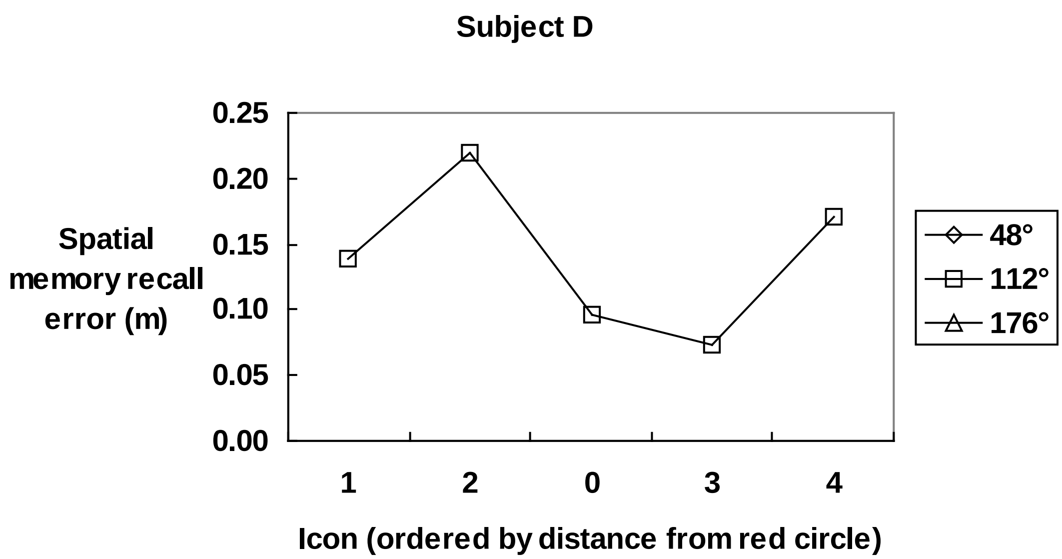

Figure A-29. Spatial memory error for subject D.

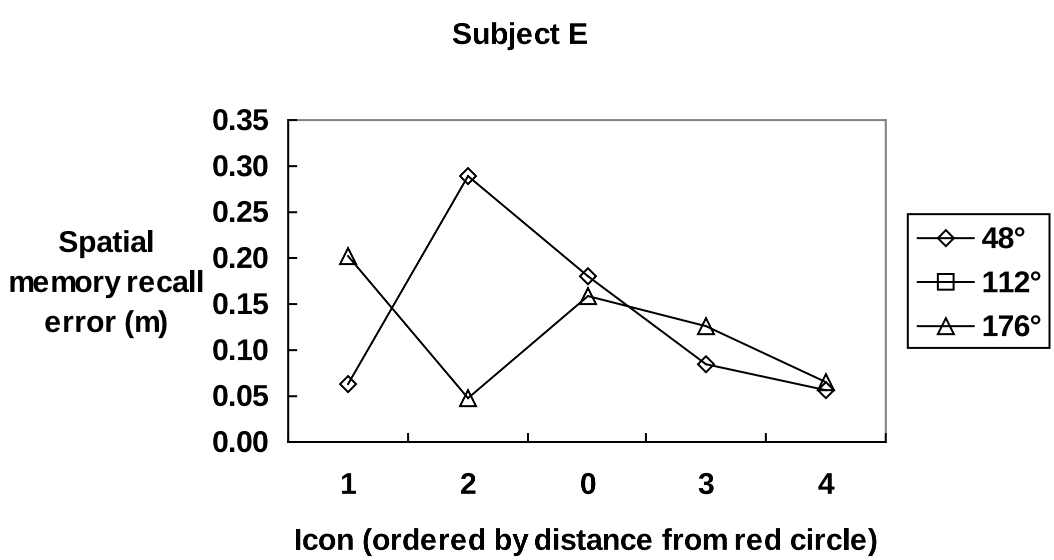

Figure A-30. Spatial memory error for subject E.

A.9 Presence data

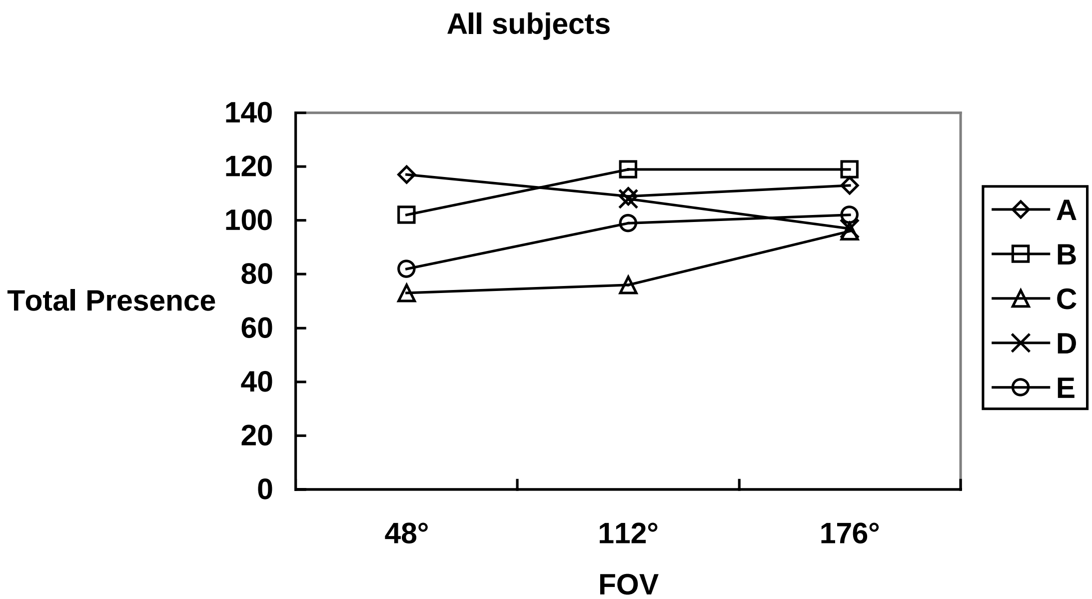

Figure A-31. Presence scores by FOV and subject.

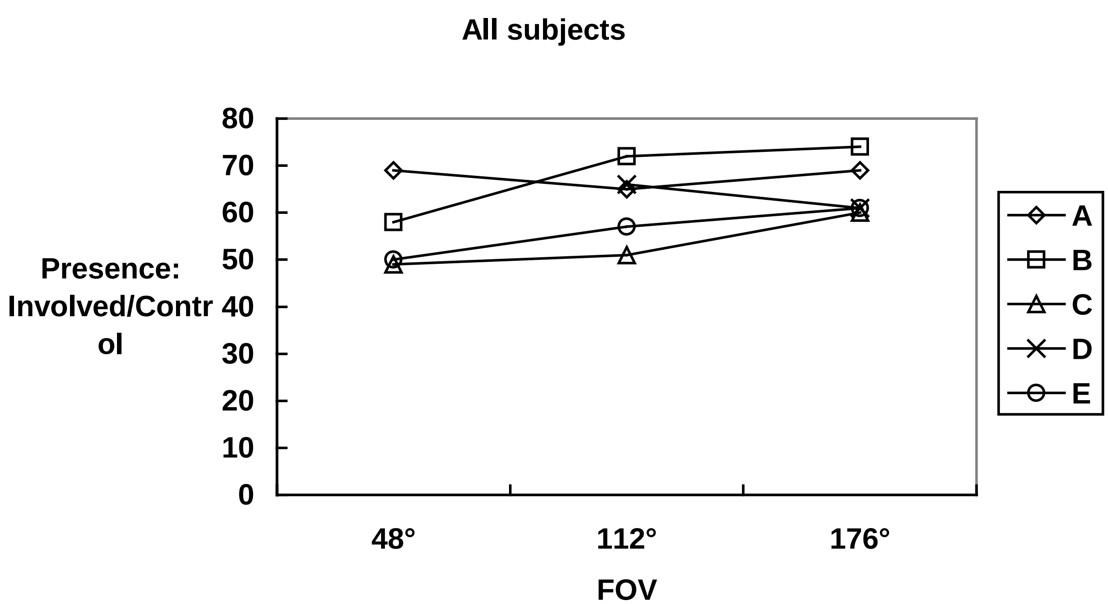

Figure A-32. Involved/Control Presence scores by FOV and subject.

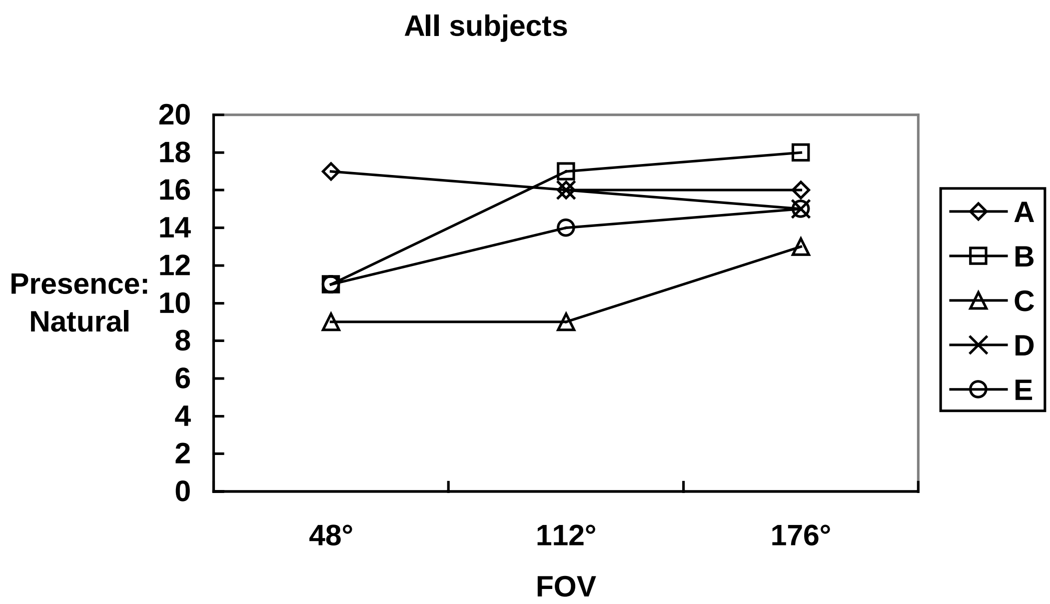

Figure A-33. Natural Presence scores by FOV and subject.

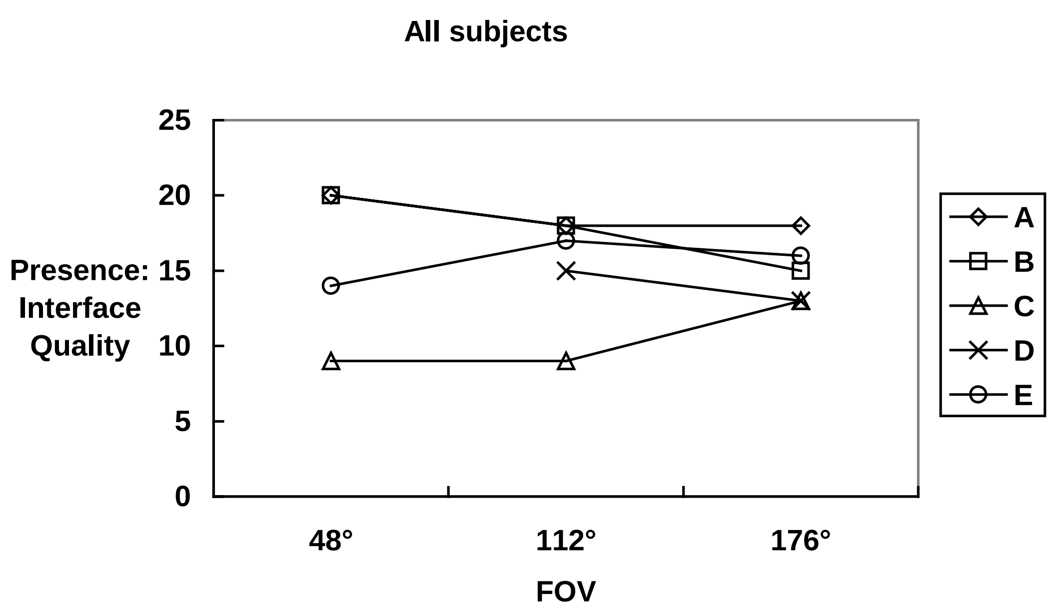

Figure A-34. Interface Quality Presence scores by FOV and subject.

A.10 Room size estimation data

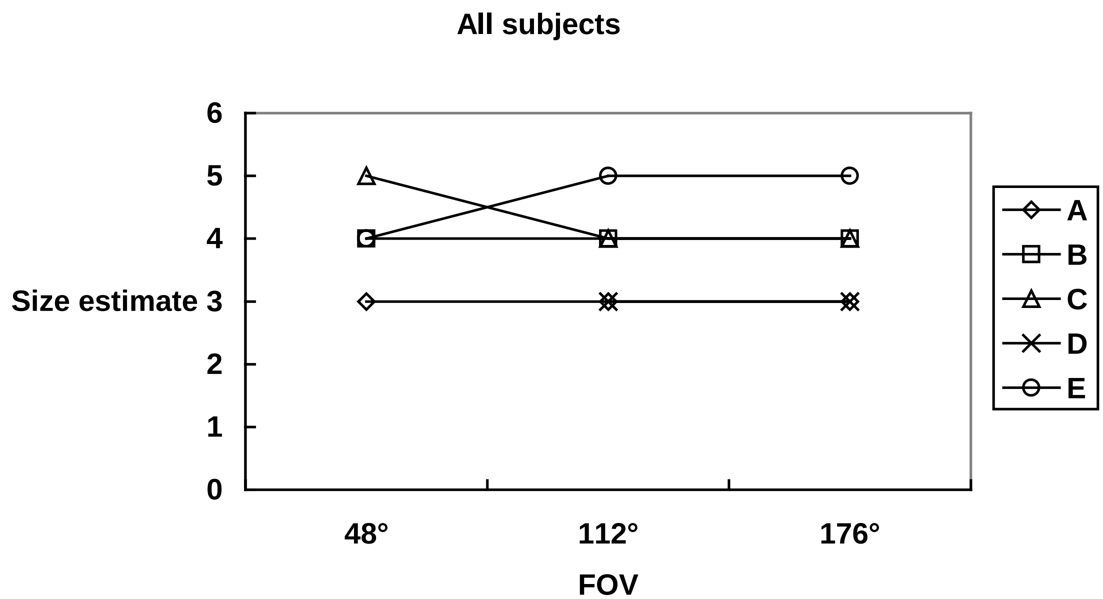

Figure A-35. Room size estimate responses by FOV and subject.

A.11 Subject comments

The following comments were given in response to question 34 of the presence questionnaire. The number beside the FOV indicated the order of experiment sessions (see Table 4-1).

Table A-3. Subject comments.

Subject	FOV	Comment
A	48° (1)	It was frustrating to see red when I felt as though I was not touching the blocks. I never got a sense for how large my virtual body was.
A	112° (3)	(none)
A	176° (2)	I was able to perform tasks much easier with all the cameras on, however, I felt more dizzy as well. When I took the headset off I was fine again.
B	48° (1)	The block walls were very easy to ignore, for the most part, because all I got

		of them was a visual sense. When I avoid things, it's because I have a solid knowledge of how my body moves and I can sort of "feel" other objects without actually touching them. I couldn't do that in the virtual environment, and my virtual body was of different dimensions (and moved differently) than my own body, so I was constantly bumping into walls. Also, the white circle test was tough because I memorize locations via landmarks. 'Twas still fun, though.
B	112° (2)	The additional screens helped in task 2, but had no real effect on my performance on tasks 1 and 3. I still felt somewhat like I was occasionally bumping into walls that I wouldn't in real life. Task 3 was still surprisingly difficult, but fun.
B	176° (3)	It seemed a lot more real today than it did before, mostly thanks to the additional screens on the edge of my vision, giving me a wider area to see. That made me better able to control myself in the maze, and my eyes having the ability to track really made the first task go by quicker, as I could target the start button & crystal early and click them as soon as the box went over them. A *much* more realistic experience.
C	48° (2)	(none)
C	112° (3)	(none)
C	176° (1)	(none)
D	48° (2)	(none)
D	112° (1)	(none)
D	176° (3)	(none)

E	48° (3)	Helmet seemed louder today; probably b/c I put it on when it was off/silent, then when turned on it buzzed… much harder to find the pyramid w/ limited vision → I had to move slower in order to find it/not pass it limited vision helped me in the maze a bit, I think b/c of less white space surrounding (but going thru last door to red was harder – lost sight of where the opening was once I started to step thru limited vision made the dot exercise much harder too – less able to survey the field/room
E	112° (1)	Depth was easier for me to see with the colored dots in the environment
E	176° (2)	Still easier to see the depth w/ the color dots – I wonder if the walls were a diff color I would hit them less walking. Much easier staying w/in the maze the 2nd day.

Appendix B
Documents

B.1 Consent form

Effects of Field of View in Virtual Reality -- Informed Consent

Introduction and purpose of the study:

We are inviting you to participate in a study of behavior in virtual reality (VR) systems. The purpose of this research is to measure how people perform or behave differently when using different types of virtual reality headsets. We hope to learn things that will improve our understanding of human behavior in virtual environments, and that will allow us to improve the technology.

The principal investigator is Kevin Arthur (UNC Chapel Hill Department of Computer Science, 217 Sitterson Hall, 962-1729, e-mail: arthur@cs.unc.edu). The faculty advisor is Dr. Frederick P. Brooks, Jr. (UNC Chapel Hill Department of Computer Science, 216 Sitterson Hall, 962-1931, e-mail: brooks@cs.unc.edu).

What will happen during the study:

We will ask you to come to the laboratory for five sessions, each on a different day, with each session lasting approximately one to two hours. During the sessions you will perform a number of simple tasks while wearing a VR headset. You will also be given questionnaires asking about your perceptions and feelings during and after the VR experience. Approximately 6 people will take part in this study.

We will use computers to record your head and hand motion during the VR experience, and will also make video and audio recordings of the session.

Protecting your privacy:

We will make every effort to protect your privacy. We will not use your name in any of the data recording or in any research reports. We will use a code number rather than your name. No images from the videotapes in which you are personally recognizable will be used in any presentation of the results.

Risks and discomforts:

While using virtual reality systems, some people experience slight symptoms of disorientation, nausea, or dizziness. These can be similar to “motion sickness” or to feelings experienced in wide-screen movies and theme park rides. We do not expect these effects to be strong or to last after you leave the laboratory. If at any time during the study you feel uncomfortable and wish to stop the experiment you are free to do so.

Your rights:

You have the right to decide whether or not to participate in this study, and to withdraw from the study at any time without penalty. We will pay you $8 per hour you spend participating in the study.

Institutional Review Board approval:

The Academic Affairs Institutional Review Board (AA-IRB) of the University of North Carolina at Chapel Hill has approved this study. If you have any concerns about your rights in this study you may contact the Chair of the AA-IRB, David A. Eckerman, at CB#4100, 201 Bynum Hall, UNC-CH, Chapel Hill, NC 27599-4100, (919) 962-7761, or e-mail: aa-irb@unc.edu.

Summary:

I understand that this is a research study to measure behavior in virtual reality.

I understand that if I agree to be in this study:

- I will visit the laboratory five times for sessions lasting one to two hours.
- I will wear a virtual reality headset to perform tasks, and my movements and behavior will be recorded by computer and on videotape, and I will respond to questionnaires between and after the sessions.
- I may experience slight feelings of disorientation, nausea, or dizziness during or shortly after the VR experiences.

I certify that I am at least 18 years of age.

I have had a chance to ask any questions I have about this study and those questions have been answered for me.

I have read the information in this consent form, and I agree to be in the study. I understand that I will get a copy of this consent form after I sign it.

____ _______________

Signature of Participant Date

B.2 Experiment information sheet

Virtual Reality Study – General Information

Introduction to the study

Thank you for participating in this study. This document is intended to familiarize you with the things you'll be doing during the experiment. It's important that you read this before your first session. Please feel free to ask questions of the experimenters in person or by e-mail or phone.

The experiment will take place in five sessions, each on a different day (about two days apart). Once you've started these sessions, please don't discuss any aspects of the experiment with other people who may also be subjects, until after you have completed the experiment.

Payment

We will pay you after you have completed the study (or after your last session if you don't finish it). Payment is $8 per hour.

Equipment and tasks

We will demonstrate to you the VR headset and the general experiment scenario. When you wear the headset, you'll see a computer-generated scene (the "virtual environment"). We'll give you verbal instructions, telling you to perform different tasks while wearing the headset. For example, your instructions might be like the following:

"Your task is to walk from the red circle, where you're now standing, to the blue circle across the room, and then click the button. Stay clear of any obstacles in your way and don't step outside the red lines on the floor. Walk at a pace that's comfortable and is not too fast that you run into the obstacles. When you're ready to begin this task, click the button and begin."

The experimenter will demonstrate each of the tasks, and you will also have time to practice. We will be measuring your movement through the virtual environment, and your performance rates and times. Please listen to the instructions carefully and ask questions if they are not clear.

Health information and the Simulator Sickness Questionnaire

At several times during each session we'll have you enter responses (on a computer) to a Simulator Sickness Questionnaire (SSQ). This is a checklist of 16 symptoms that people sometimes experience after using flight simulators or virtual reality systems. Please review the attached copy before your first session, and if any of the terms seem unfamiliar please ask questions.

The other "Participant health questionnaire" is to check that you are in good health to start with. We'll have you fill this in each day you come to the lab. We'll also have you fill in the

SSQ once each day before you start, again to check that you're feeling well enough. It's important that you are in your usual state of good health when you do the experiment, so please keep this in mind.

The Presence Questionnaire

Once after each session we'll have you fill out (on paper) a questionnaire about your sense of "presence" in the virtual environment that you experienced. Presence is a term to describe the subjective sense of being in a particular environment. For example: ordinarily we feel present in the physical real environment, but we might feel present in an engrossing movie or a theme park ride (we feel like we're "really there" inside the movie instead of the real world).

Please read the attached copy of the questionnaire once before your first session, and ask questions if anything is unclear.

Location and scheduling

The experiments will take place in the HMD Lab (room 265) in Sitterson Hall. We have a tight schedule with several subjects, so please arrive promptly at your scheduled times. The spacing between sessions is important due to the nature of the study, and therefore we cannot reschedule sessions if you miss one. If you are late or absent for any session we will have to drop you from the study.

B.3 Pre-session health questionnaire

This questionnaire was administered using an HTML form. Below is a reproduction of the information on that form.

Participant Health Information

Subject ID:

1. Are you in your usual state of good fitness (health)?

   Yes
   No (please explain)

2. In the past 24 hours which, if any, of the following substances have you used? (Check all that apply)

   None
   Sedatives or Tranquilizers
   Decongestants
   Anti-histamines
   Alcohol (three drinks or more)
   Other (please explain)

B.4 Simulator Sickness Questionnaire

This questionnaire was administered using an HTML form. Below is a reproduction of the information on that form.

Simulator Sickness Questionnaire

Subject ID:

For each of the following conditions, please indicate how you are feeling *right now*, on the scale of "none" through "severe".

1. General discomfort	None	Slight	Moderate	Severe
2. Fatigue	None	Slight	Moderate	Severe
3. Headache	None	Slight	Moderate	Severe
4. Eye Strain	None	Slight	Moderate	Severe
5. Difficulty Focusing	None	Slight	Moderate	Severe
6. Increased Salivation	None	Slight	Moderate	Severe
7. Sweating	None	Slight	Moderate	Severe
8. Nausea	None	Slight	Moderate	Severe
9. Difficulty Concentrating	None	Slight	Moderate	Severe
10. Fullness of Head	None	Slight	Moderate	Severe
11. Blurred Vision	None	Slight	Moderate	Severe
12. Dizzy (Eyes Open)	None	Slight	Moderate	Severe
13. Dizzy (Eyes Closed)	None	Slight	Moderate	Severe
14. Vertigo	None	Slight	Moderate	Severe
15. Stomach Awareness	None	Slight	Moderate	Severe
16. Burping	None	Slight	Moderate	Severe

For each question, a score of 0 (none), 1 (slight), 2 (moderate), or 3 (severe) is assigned. The scores are then combined as follows (Kennedy et al., 1993). (See also Table 5-1.)

COLUMN 1 = SUM (1, 6, 7, 8, 9, 15, 16)
COLUMN 2 = SUM (1, 2, 3, 4, 5, 9, 11)
COLUMN 3 = SUM (5, 8, 10, 11, 12, 13, 14)

NAUSEA = COLUMN 1 $\times$ 9.54
OCULOMOTOR DISCOMFORT = COLUMN 2 $\times$ 7.58
DISORIENTATION = COLUMN 3 $\times$ 13.92

TOTAL SEVERITY = (COLUMN 1 + COLUMN 2 + COLUMN 3) $\times$ 3.74

Definitions for Simulator Sickness Questionnaire

The following information was available through a link from the SSQ form.

Explanation of Conditions

General Discomfort	
Fatigue	weariness or exhaustion of the body
Headache	
Eye Strain	weariness or soreness of the eyes
Difficulty Focusing	
Increased Salivation	
Sweating	
Nausea	stomach distress
Difficulty Concentrating	
Fullness of Head	
Blurred Vision	
Dizzy (with your eyes open)	
Dizzy (with your eyes closed)	
Vertigo	surroundings seem to swirl
Stomach Awareness	a feeling just short of nausea
Burping	

B.5 Presence questionnaire

The presence questionnaire I used was modified from that of Witmer and Singer (1998). I obtained a copy of that questionnaire (version 3.0 dated November 1994) from the authors on February 17, 1999. The questionnaire is scored as follows (the following text is from the original questionnaire):

> "Simply score the boxes for each question from left to right beginning with one and increasing in value to the box the subject has marked, and the number of that box becomes the score. Some of the questions have reversed response anchors, and are scored so the left-most box receives a seven and the rest decrease in value. The subscale scores are the sum of the scores for each subscale item. There is no weighting of items or subscales. The questionnaire total and subscales are comprised as follows:
>
> PRESENCE QUESTIONNAIRE
>
> > <u>Total</u>: Items 1, 2, 3, 4, 6, 7, 8, 9, 10, 14, 15, 16, 18, 19+, 20, 21, 22+, 23+, 24.
> > <u>PQ-Involved/Control</u>: Items 1, 2, 4, 7, 9, 10, 14, 18, 19+, 20, & 21.
> > <u>PQ-Natural</u>: Items 3, 6, & 8.
> > <u>PQ-Interface Quality</u>: Items 22+, 23+, & 24.
> > <u>PQ-Auditory*</u>: Items 5, 11, 12.
> > <u>PQ-Haptic*</u>: Items 13 & 17.
> > <u>PQ-Resolution*</u>: Items 15 & 16.
>
> The last three subscales listed for the PQ are marked with an asterisk (*) because they have yet to be used in analyses, but are being retained on a theoretical basis. Since there have been no haptic or auditory interfaces, nor any differences in resolution to judge, those items have been scored as zero. Items marked with a plus (+) have to be reverse scored (see above) in order to contribute to the subscale and overall totals."

Because our system did not use audio or haptics , I removed questions that contributed to those subscores and to no others. The original questionnaire also contains experimental questions numbered 25 through 32, which the authors have not yet used for computing any scores. I removed those questions as well. I added a question on room estimation (33), and a question for general comments (34). The full questionnaire that I used is as follows.

PRESENCE QUESTIONNAIRE

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

WITH REGARD TO THE VIRTUAL ENVIRONMENTS
THAT YOU EXPERIENCED TODAY

1. How much were you able to control events?

|________|________|________|________|________|________|________|
NOT AT ALL SOMEWHAT COMPLETELY

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

|________|________|________|________|________|________|________|
NOT RESPONSIVE MODERATELY RESPONSIVE COMPLETELY RESPONSIVE

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

|________|________|________|________|________|________|________|
EXTREMELY ARTIFICIAL BORDERLINE COMPLETELY NATURAL

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

|________|________|________|________|________|________|________|
NOT AT ALL SOMEWHAT COMPLETELY

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

|________|________|________|________|________|________|________|
EXTREMELY ARTIFICIAL BORDERLINE COMPLETELY NATURAL

7. How compelling was your sense of objects moving through space?

|________|________|________|________|________|________|________|

NOT AT ALL	MODERATELY COMPELLING	VERY COMPELLING

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

|________|________|________|________|________|________|________|

NOT CONSISTENT	MODERATELY CONSISTENT	VERY CONSISTENT

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

|________|________|________|________|________|________|________|

NOT AT ALL	SOMEWHAT	COMPLETELY

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

|________|________|________|________|________|________|________|

NOT AT ALL	SOMEWHAT	COMPLETELY

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

|________|________|________|________|________|________|________|

NOT COMPELLING	MODERATELY COMPELLING	VERY COMPELLING

15. How closely were you able to examine objects?

|________|________|________|________|________|________|________|

NOT AT ALL	PRETTY CLOSELY	VERY CLOSELY

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

|________|________|________|________|________|________|________|

NOT AT ALL	SOMEWHAT	EXTENSIVELY

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

|________|________|________|________|________|________|________|

NOT INVOLVED	MILDLY INVOLVED	COMPLETELY ENGROSSED

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

|________|________|________|________|________|________|________|

NO DELAYS	MODERATE DELAYS	LONG DELAYS

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

|________|________|________|________|________|________|________|

NOT AT ALL	SLOWLY	LESS THAN ONE MINUTE

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

|________|________|________|________|________|________|________|

NOT PROFICIENT	REASONABLY PROFICIENT	VERY PROFICIENT

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

|________|________|________|________|________|________|________|

NOT AT ALL	INTERFERED SOMEWHAT	PREVENTED TASK PERFORMANCE

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

|________|________|________|________|________|________|________|

NOT AT ALL	INTERFERED SOMEWHAT	INTERFERED GREATLY

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

|________|________|________|________|________|________|________|

NOT AT ALL	SOMEWHAT	COMPLETELY

33. How would you describe the size of the *virtual* room that you saw inside the headset compared to the size of the *real* room that you saw when you took off the headset? The virtual room was:

|________|________|________|________|________|________|________|
SMALLER   SAME SIZE   LARGER

34. Please enter below any additional comments you wish to make about the virtual environments that you experienced today:

REFERENCES

Alfano, Patricia L. and George F. Michel. (1990). Restricting the field of view: perceptual and performance effects. *Perceptual and Motor Skills, 70*(1), 35-45.

Barfield, Woodrow, Claudia Hendrix, Ove Bjorneseth, Kurt A. Kaczmarek, and Wouter Lotens. (1995). Comparison of Human Sensory Capabilities with Technical Specifications of Virtual Environment Equipment. *Presence, 4*(4), 329-356.

Barfield, W. and S. Weghorst. (1993). The sense of presence within virtual environments: a conceptual framework. In G. Salvendy & M. Smith (Eds.), *Human-Computer Interaction: Software and Hardware Interfaces* (pp. 699-704). Amsterdam: Elsevier.

Billinghurst, Mark and Suzanne Weghorst. (1995, March 11-15). The use of sketch maps to measure cognitive maps of virtual environments. In Proceedings of Virtual Reality Annual International Symposium 95, Research Triangle Park, NC, IEEE Computer Society Press, 40-47.

Birch, Jennifer. (1993). *Diagnosis of Defective Colour Vision*. Oxford; New York: Oxford University Press.

Bittner, Alvah C., Jr., Brian F. Gore, and Becky L. Hooey. (1997). Meaningful assessments of simulator performance and sickness: Can't have one without the other? In Proceedings of 41st Annual Meeting of the Human Factors and Ergonomics Society, 1089-1092.

Brooks, Frederick P., Jr. (1986, October). Walkthrough -- A dynamic graphics system for simulating virtual buildings. In Proceedings of Workshop on Interactive 3D Graphics, Chapel Hill, NC, *Computer Graphics*, 9-21.

Brooks, Frederick P., Jr. (1999). What's Real About Virtual Reality? *IEEE Computer Graphics and Applications, 19*(6), 16-27.

Cannon, Mark W., Jr. (1986). Recent advances in understanding peripheral vision. In Proceedings of 30th Annual Meeting of the Human Factors Society, 601-604.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). New York: Academic Press.

Cohen, Jacob and Patricia Cohen. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.

Craig, Ian R., Andrew A. Marshall, and Christopher S. Jordan. (1997). An evaluation of a methodology to develop future helmet-mounted display symbology. In Proceedings of Head Mounted Displays II, *SPIE Proceedings volume 3058*, 105-114.

Cunningham, James A., W. Todd Nelson, Lawrence J. Hettinger, Michael W. Haas, and Chris A. Russell. (1996). Perceptual & Human Factors Issues in Virtual Environments (Course Notes): Assessing Target Detection Performance Using Head Position Data: A Qualitative and Quantitative Analysis. In Proceedings of IMAGE '96.

Davids, Keith and Rolf Stratford. (1989). Peripheral vision and simple catching: the screen paradigm revisited. *Journal of Sports Sciences, 7*, 139-152.

Davis, Elizabeth Thorpe. (1997). Visual requirements in HMDs: What can we see and what do we need to see? In James E. Melzer & Kirk Moffitt (Eds.), *Head-Mounted Displays: Designing for the User* (pp. 207-251). New York: McGraw-Hill.

Davson, Hugh. (1990). *Physiology of the Eye*. (5th ed.). New York: Pergamon Press.

de Vries, Sjoerd C. and Pieter Padmos. (1997). Steering a simulated Unmanned Aerial Vehicle using a head-slaved camera and HMD. In Proceedings of Head Mounted Displays II, *SPIE Proceedings 3058*, 24-33.

Dinh, Huong Q., Neff Walker, and Larry F. Hodges. (1999). Evaluating the Importance of Multi-sensory Input on Memory and the Sense of Presence in Virtual Environments. In Proceedings of IEEE Virtual Reality '99, 222-228.

Dolezal, Hubert. (1982a). *Living in a world transformed: perceptual and performatory adaptation to visual distortion*. New York: Academic Press.

Dolezal, Hubert. (1982b). The Restricted Field of View: A Neglected Factor in Adaptation Research, *Living in a world transformed: Perceptual and performatory adaptation to visual distortion* (pp. 57-79). New York: Academic Press.

Draper, Mark. (1998). *The Adaptive Effects of Virtual Interfaces: Vestibulo-Ocular Reflex and Simulator Sickness.* PhD dissertation, University of Washington Human Interface Technology Lab.

Edwards, Kenneth L., John W. Buckle, Mark J. Doherty, Lionel J. Lee, Adam C. Pratty, and John F. White. (1997). An operationally-applicable objective method for the analysis and evaluation of the flights of helicopter mission task elements during field-of-view trials. In Proceedings of Head Mounted Displays II, *SPIE Proceedings 3058*, 235-251.

Ellis, S.R. and U.J. Bucher. (1994). Distance perception of stereoscopically presented virtual objects optically superimposed on physical objects in a head-mounted see-through display. In Proceedings of 38th Annual Meeting of the Human Factors and Ergonomics Society, 1300-1305.

Ellis, Stephen R., Mary K. Kaiser, and Arthur J. Grunwald. (1993). *Pictorial communication in virtual and real environments*. (2nd ed.). London: Taylor and Francis.

Feldman, J.A. (1985). Four frames suffice: A provisional model of vision and space. *The Behavioural and Brain Sciences, 8*, 265-289.

Fischer, Robert E. (1997). Fundamentals of HMD Optics. In James E. Melzer & Kirk Moffitt (Eds.), *Head-Mounted Displays: Designing For the User* (pp. 83-116). New York: McGraw-Hill.

Gauthier, G.M., G. Obrecht, C. Pedrono, J.L. Vercher, and L. Stark. (1987). Adaptive optimization of eye-head coordination with degraded peripheral vision. In J.K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 201-210). Amsterdam: Elsevier Science Publishers.

Geiselman, Eric E. and Brian H. Tsou. (1996). Helmet-display resident target locator line symbology: an evaluation of vector length depiction. In Proceedings of Head Mounted Displays, *SPIE Proceedings 2735*, 233-244.

Gerth, Jeff. (1997). Design Issues in Human Performance-Based Test and Evaluation of HMDs. In James E. Melzer & Kirk Moffitt (Eds.), *Head-Mounted Displays: Designing For the User* (pp. 313-335). New York: McGraw-Hill.

Gibson, James J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Hagen, Margaret A., Rebecca K. Jones, and Edward S. Reed. (1978). On a neglected variable in theories of pictorial perception: Truncation of the visual field. *Perception & Psychophysics, 23*(4), 326-330.

Henry, Daniel and Tom Furness. (1993). Spatial perception in virtual environments: evaluating an architectural application. In Proceedings of Virtual Reality Annual International Symposium, IEEE Computer Society Press, 33-40.

Hettinger, Lawrence J. and Robert S. Kennedy. (1996). Perceptual & Human Factors Issues in Virtual Environments (Course Notes). In Proceedings of IMAGE '96.

Hoffman, Hunter G., Keith G. Hullfish, and Stacey J. Houston. (1995, March 11-15). Virtual-Reality Monitoring. In Proceedings of Virtual Reality Annual International Symposium 95, Research Triangle Park, NC, IEEE Computer Society Press, 48-54.

Hood, Donald C. and Marcia A. Finkelstein. (1986). Sensitivity to light, *Handbook of Perception and Human Performance* (Vol. I, Chapter 5,). New York: Wiley.

Howard, Ian P. (1993). Spatial vision within egocentric and exocentric frames of reference. In Stephen R. Ellis, Mary K. Kaiser, & Arthur J. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (2nd ed., pp. 338-358). London: Taylor & Francis.

Howlett, Eric M. (1990). Wide angle orthostereo. In Proceedings of 1990 SPIE Conference on Stereoscopic Displays and Applications, Santa Clara, CA, *SPIE Proceedings 1256*, 210-223.

Jorgensen, Craig, Jeff Ogden, E. Kerry Willis, Mary Blessing, Kathryn Ann Caudell, Graham Patrick, and Thomas P. Caudell. (1997). Locomotion in a virtual environment: Performance measures and physiological responses. In Proceedings of 41st Annual Meeting of the Human Factors and Ergonomics Society, 1148-1151.

Kaiser Electro-Optics. (2000). HiDef Family of Head-Mounted Displays . Available at http://www.keo.com.

Kennedy, Robert S., Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. (1993). A simulator sickness questionnaire (SSQ): A new method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*(3), 203-220.

Kennedy, Robert S. and Michael G. Lilienthal. (1995, March 11-15). Implications of balance disturbances following exposure to virtual reality systems. In Proceedings of Virtual Reality Annual International Symposium 95, Research Triangle Park, NC, IEEE Computer Society Press, 35-39.

Kenyon, Robert V. and Edward W. Kneller. (1993). The effects of field of view size on the control of roll motion. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(1), 183-193.

Kolasinski, Eugenia M. (1995). *Simulator sickness in virtual environments* (Technical Report 1027). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kolasinski, Eugenia M. (1996). *Prediction of Simulator Sickness in a Virtual Environment.* PhD dissertation, University of Central Florida Psychology Department, Orlando.

Kraemer, H.C. and S. Thiemann. (1987). *How many subjects? Statistical power analysis in research.* Newbury Park, CA: Sage.

Lampton, Donald R., Bruce W. Knerr, Stephen L. Goldberg, James P. Bliss, J. Michael Moshell, and Brian S. Blau. (1994). The Virtual Environment Performance Assessment Battery (VEPAB): Development and Evaluation. *Presence, 3*(2), 145-157.

Latham, Roy. (1998). Head-Mounted Display Survey. *Real Time Graphics, 7*(2), 8-12.

Latham, Roy. (1999). Head-Mounted Display Survey. *Real Time Graphics, 8*(2), 6-13.

Lee, D.N. (1977). The functions of vision. In H.L. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information* . Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Leibowitz, H.W. (1986). Recent advances in our understanding of peripheral vision and some implications. In Proceedings of 30th Annual Meeting of the Human Factors Society, 605-607.

Leung, Y.K. and M.D. Apperley. (1994). A review and taxonomy of distortion-oriented presentation techniques. *ACM Transactions on Computer-Human Interaction, 1*(2), 126-160.

Lipsey, Mark W. (1990). *Design Sensitivity: Statistical Power for Experimental Research*. Newbury Park, CA: Sage.

Malcolm, Richard. (1984). Pilot disorientation and the use of a peripheral vision display. *Aviation, Space, and Environmental Medicine*(March 1984).

McCauley, Michael E. and Thomas J. Sharkey. (1992). Cybersickness: Perception of self-motion in virtual environments. *Presence, 1*(3), 311-318.

McKenna, Michael. (1992, March). Interactive viewpoint control and three-dimensional operations. In Proceedings of Symposium on Interactive 3D Graphics, Boston, MA, *Computer Graphics*, 53-56.

McLean, William E., Clarence E. Rash, Joseph McEntire, Malcolm G. Braithwaite, and John C. Mora. (1997). A Performance History of AN/PVS-5 and ANVIS Image Intensification Systems in U.S. Army Aviation. In Proceedings of Head Mounted Displays II, *SPIE Proceedings 3058*, 264-298.

Melzer, James E. and Kirk Moffitt. (1997). HMD Design--Putting the User First. In James E. Melzer & Kirk Moffitt (Eds.), *Head-Mounted Displays: Designing for the User* (pp. 1-16). New York: McGraw-Hill.

Monterey Technologies. (1999). Presentation of Aircraft State Information in an Ambient Head-Mounted Display . Available at http://www.montereytechnologies.com/hmd.htm.

Mon-Williams, M., J.P. Wann, and S. Rushton. (1993). Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display. *Ophthalmic Physiol. Opt., 13*(4), 387-391.

Muller, Keith E. and Vernon A. Benignus. (1992). Increasing scientific power with statistical power. *Neurotoxicology and Teratology, 14*, 211-219.

Nordwall, Bruce D. (1989). Pilots sense attitude with peripheral vision using new Garrett display. *Aviation week and space technology, 130*(19), 97-99.

Oman, Charles M. (1993). Sensory conflict in motion sickness: an Observer Theory approach. In Stephen R. Ellis, Mary K. Kaiser, & Arthur J. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (2nd ed., pp. 362-376). London: Taylor & Francis.

Pausch, Randy, Thomas Crea, and Matthew Conway. (1992). A literature survey for virtual environments: Military flight simulator visual systems and simulator sickness. *Presence, 1*(3), 344-363.

Pausch, Randy, Dennis Proffitt, and George Williams. (1997). Quantifying Immersion in Virtual Reality. In Proceedings of SIGGRAPH 97, Los Angeles, CA, *Computer Graphics Proceedings*, Annual Conference Series, ACM SIGGRAPH, 13-18.

Pausch, Randy, M. Anne Shackelford, and Dennis Proffitt. (1993). A user study comparing head-mounted and stationary displays. In Proceedings of IEEE 1993 Symposium on Research Frontiers in Virtual Reality, 41-45.

Perry, Chris E. and John R. Buhrman. (1997). HMD Head and Neck Biomechanics. In James E. Melzer & Kirk Moffitt (Eds.), *Head-Mounted Displays: Designing For the User* (pp. 147-173). New York: McGraw-Hill.

Peruch, Patrick and Daniel Mestre. (1999). Between Desktop and Head Immersion: Functional Visual Field During Vehicle Control and Navigation in Virtual Environments. *Presence, 8*(1), 54-64.

Piantanida, Thomas P., Duane Boman, James Larimer, Jennifer Gille, and Charles Reed. (1992). Studies of the field-of-view/resolution tradeoff in virtual reality. In Proceedings of Human Vision, Visual Processing and Digital Display III, *SPIE Proceedings 1666*, 448-456.

Raskar, Ramesh, Matt Cutts, Greg Welch, and Wolfgang Stuerzlinger. (1998, June). Efficient Image Generation for Multiprojector and Multisurface Displays. In Proceedings of Ninth EuroGraphics Rendering Workshop, Vienna, Austria, 139-144.

Riccio, G.E. and T.A. Stoffregen. (1991). An ecological theory of motion sickness and postural instability. *Ecological Psychology, 3*(3), 195-240.

Rinalducci, Edward J., Dan Mapes, Stephen G. Cinq-Mars, and Kent E. Higgins. (1996). Determining the Field of View in HMDs: A Psychophysical Method. *Presence, 5*(3).

Robinett, Warren and Richard Holloway. (1992). Implementation of flying, scaling, and grabbing in virtual worlds. In Proceedings of ACM Symposium on Interactive 3D graphics, 189-192.

Robinett, Warren and Richard Holloway. (1995). The Visual Display Transformation for Virtual Reality. *Presence, 4*(1), 1-23.

Robinett, Warren and Jannick P. Rolland. (1992). A computational model for the stereoscopic optics of a head-mounted display. *Presence, 1*(1), 45-62.

Rolland, J.P., W. Gibson, and D. Ariely. (1995). Towards quantifying depth and size perception in virtual environments. *Presence, 4*(1), 24-49.

Roscoe, S.N. (1993). The eyes prefer real images. In S.R. Ellis, M.K. Kaiser, & A.J. Grunwald (Eds.), *Pictorial Communication in Virtual and Real Environments* (2nd ed., pp. 577-586). London: Taylor and Francis.

Sampson, J. B. (1993). Cognitive performance of individuals using a head-mounted display while walking. In Proceedings of 37th Annual Meeting of the Human Factors Society.

Sivak, Barbara and Christine L. MacKenzie. (1992). The contributions of peripheral vision and central vision to prehension. In L. Proteau & D. Elliot (Eds.), *Vision and Motor Control* (pp. 233-259). Amsterdam: Elsevier Science Publishers.

Slater, Mel. (1999). Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence, 8*(5), 560-565.

Slater, Mel and Martin Usoh. (1993a). Representations systems, perceptual position and presence in virtual environments. *Presence, 2*(3), 221-233.

Slater, Mel and Martin Usoh. (1993b). Simulating peripheral vision in immersive virtual environments. *Computers & Graphics, 17*(6), 643-653.

Slater, M., M. Usoh, and A. Steed. (1995). Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality. *ACM Transactions on Computer-Human Interaction, 2*(3), 201-219.

So, Richard H. Y. and W. T. Lo. (1999). Cybersickness: An Experimental Study to Isolate The Effects of Rotational Scene Oscillations. In Proceedings of IEEE Virtual Reality '99, 237-241.

Stanney, Kay. (1995, March 11-15). Realizing the full potential of virtual reality: human factors issues that could stand in the way. In Proceedings of Virtual Reality Annual International Symposium 95, IEEE Computer Society Press, 28-34.

Stanney, Kay and Gavriel Salvendy. (1998). Aftereffects and Sense of Presence in Virtual Environments: Formulation of a Research and Development Agenda. *International Journal of Human-Computer Interaction, 10*(2), 135-187.

Stanney, Kay M. and P. Hash. (1998). Locus of user-initiated control in virtual environments: Influences on cybersickness. *Presence, 7*(5), 447-459.

Stanney, Kay M., Robert S. Kennedy, and Julie M. Drexler. (1997). Cybersickness is not simulator sickness. In Proceedings of 41st Annual Meeting of the Human Factors and Ergonomics Society, 1138-1142.

Stanney, Kay M., Ronald R. Mourant, and Robert S. Kennedy. (1998). Human Factors Issues in Virtual Environments: A Review of the Literature. *Presence, 7*(4), 327-351.

Stark, L.W., K. Ezumi, T. Nguyen, R. Paul, G. Tharp, and H.I. Yamashita. (1992). Visual search in virtual environments. In Proceedings of Human Vision, Visual Processing, and Digital Display III, *SPIE Proceedings 1666,* 577-589.

Stern, Robert M., Senqi Hu, Richard B. Anderson, Herschel W. Leibowitz, and Kenneth L. Koch. (1990). The effects of fixation and restricted visual field on vection-induced motion sickness. *Aviation, Space, and Environmental Medicine, 61*(8), 712-715.

Stevens, James. (1996). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sutherland, Ivan. (1965). The ultimate display. In Proceedings of Proceedings of the IFIP Congress, Federation of Information Processing Societies, 506-508.

Task, H.L. (1997). HMD Image Source, Optics, and the Visual Interface. In James E. Melzer & Kirk Moffitt (Eds.), *Head-Mounted Displays: Designing for the User* (pp. 55-82). New York: McGraw-Hill.

Toto, Lawrence and Harold Kundel. (1987). An eye motion dependent interactive display. In J.K. O'Regan & A. Levy-Schoen (Eds.), *Eye Movements: From Physiology to Cognition* (pp. 643-649). Amsterdam: Elsevier Science Publishers.

Triesman, M. (1977). Motion sickness: An evolutionary approach. *Science*, *197*, 493-495.

Usoh, Martin, Kevin Arthur, Mary C. Whitton, Rui Bastos, Anthony Steed, Mel Slater, and Jr. Frederick P. Brooks. (1999). Walking > Walking-in-Place > Flying, in Virtual Environments. In Proceedings of SIGGRAPH 99, Los Angeles, CA, *Computer Graphics Proceedings*, Annual Conference Series, ACM SIGGRAPH, 359-364.

VRPN. (1999). Virtual Reality Peripheral Network . Available at http://www.cs.unc.edu/Research/vrpn/.

Ward, Mark, Ronald Azuma, Robert Bennett, Stefan Gottschalk, and Henry Fuchs. (1992). A demonstrated optical tracker with scalable work area for head-mounted display systems. In Proceedings of 1992 Symposium on Interactive 3D Graphics, 43-52.

Ware, Colin and Ravin Balakrishnan. (1994). Reaching for objects in VR displays: lag and frame rate. *ACM Transactions on CHI, 1*, 331-356.

Ware, Colin and Stephen Osborne. (1990, March). Exploration and virtual camera control in virtual three dimensional environments. In Proceedings of 1990 Symposium on Interactive 3D Graphics, *Computer Graphics*, 175-183.

Watson, Benjamin, Neff Walker, Larry F. Hodges, and Martin Reddy. (1997). An Evaluation of Level of Detail Degradation in Head-Mounted Display Peripheries. *Presence, 6*(6), 630-637.

Welch, Gregory and Gary Bishop. (1997). SCAAT: Incremental tracking with incomplete information. In Proceedings of SIGGRAPH 97, Los Angeles, CA, *Computer Graphics Proceedings*, Annual Conference Series, ACM SIGGRAPH, 333-344.

Wells, Maxwell J. and Michael Venturino. (1990). Performance and head movements using a helmet-mounted display with different sized fields-of-view. *Optical Engineering, 29*(8), 870-877.

Wells, Maxwell J., Michael Venturino, and R.K. Osgood. (1989). Effect of field-of-view-size on performance at a simple simulated air-to-air mission. In Proceedings of Helmet Mounted Displays, *SPIE Proceedings 1116*, 126-137.

Werner, Elliot B. (1991). *Manual of visual fields*. New York: Churchill Livingstone.

Whitestone, Jennifer J. and Kathleen M. Robinette. (1997). Fitting to Maximize Performance of HMD Systems. In James E. Melzer & Kirk Moffitt (Eds.), *Head-Mounted Displays: Designing For the User* (pp. 175-206). New York: McGraw-Hill.

Witmer, Bob G. and Michael J. Singer. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence, 7*(3), 225-240.

Wood, Joanne M. and Rod Troutbeck. (1994). Effect of visual impairment on driving. *Human Factors, 36*(3), 476-487.

Yoshida, Akitoshi, Jannick P. Rolland, and John H. Reif. (1995, March 11-15). Design and applications of a high-resolution insert head-mounted display. In Proceedings of Virtual Reality Annual International Symposium 95, IEEE Computer Society Press, 84-93.