Charting Presence in Virtual Environments and Its Effects on Performance

Michael P. Snow

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

Doctor of Philosophy
in
Industrial and Systems Engineering

Robert C. Williges, Chair
Paul T. Kemmerling, Jr.
Walter W. Wierwille
H. Rex Hartson
John M. Reising

December 5, 1996
Blacksburg, Virginia

Keywords: Virtual Environment, Virtual Reality, Presence, Magnitude Estimation

Charting Presence in Virtual Environments and Its Effects on Performance

by Michael P. Snow

Industrial and Systems Engineering

Dr. Robert C. Williges (Chair)

(ABSTRACT)

Virtual reality (VR) involves an attempt to create an illusion that the user of the VR system is actually present in a synthetic (usually computer-generated) environment. Little is known about how various system parameters affect the illusion of presence in a virtual environment (VE). In particular, there seem to be very little quantitative data on which to base VR system design decisions. Also, while presence (or immersion) in VEs is a primary goal of VR, not much is known about how this variable affects task performance. The goal of this research was to provide a ratio-scale measure of perceived presence in a VE, to explore the effects of a number of environmental parameters on this measure and construct empirical models of these effects, and to relate perceived presence to user performance.

This was done by manipulating eleven independent variables in a series of three experiments. The independent variables manipulated were scene update rate, visual display resolution, field of view, sound, textures, head-tracking, stereopsis, virtual personal risk, number of possible interactions, presence of a second user, and environmental detail. Participants performed a set of five tasks in the VE and rated perceived presence at the end of each set using the technique of free-modulus magnitude estimation. The amount of time spent in the VE was also recorded.

The results indicate that the VR system parameters manipulated and analyzed in this research did affect participants' subjective feeling of presence in the VE. Field of view, sound, and head-tracking showed the largest effects. Other significant effects found were those of visual display resolution, texture-mapping, stereopsis, and the presence of a second user. Free-modulus magnitude estimation worked well as a measure of perceived presence. A positive relationship was found between perceived presence and task performance, but this relationship was relatively weak. Second-order empirical models were constructed that predicted perceived presence with moderate success and, with less success, task performance.

DEDICATION

for Dad

ACKNOWLEDGMENTS

I would like to thank the members of my committee for their support throughout the process of completing this dissertation. Their advice and encouragement were critical in bringing to a successful conclusion a project that required significantly more time and effort than originally anticipated. I would especially like to thank Drs. Hartson and Williges for launching the virtual reality program at Virginia Tech and providing the funding for the equipment that was used in this research. My sincere thanks go to the people associated with the U.S. Air Force's Palace Knight program, which funded the last two and a half years of my graduate work. I would like to express my thanks to John Kelso for his invaluable assistance and patience in dealing with the myriad hardware and software difficulties we encountered throughout this effort. Finally, I owe a special debt of gratitude to my wife Nancy, who has been my personal infrastructure and without whom this research and, indeed, my graduate career would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Virtual reality, or VR as it is widely known, has one defining characteristic that differentiates it from every other form of human-computer interface in existence today: the goal of creating the illusion that a user is present in a virtual environment (VE) (Ellis, 1993). While other interfaces may use 2D (e.g., GUIs) or 3D (e.g., CAD) representations and metaphors, there is no attempt to immerse the user in these representations or metaphors. It is this illusion of presence (a.k.a., immersion) that makes VR what it is.

The concept of virtual reality is commonly credited to Ivan Sutherland (1965) who proposed a human-computer interface that, "could literally be the Wonderland into which Alice walked." Telerobotics (or “telefactor”, as it was then known) systems for operations in space that included, at least conceptually, most of the technology commonly associated with modern implementations of VR were proposed as early as 1967 (Bradley, 1967). Since the origin of this concept, the terminology associated with this field has greatly expanded, together with public and scientific interest in its potential. The terms "virtual reality", "synthetic environment", and "virtual environment" are often used and are essentially synonymous; however, "virtual environment" seems to have become the standard within the scientific community because of the somewhat oxymoronic quality of "virtual reality". The increased interest has also driven technological advance and diversity of VR applications. VR has been applied in entertainment, education and training, medicine, scientific and statistical visualization, networking, telerobotics and teleoperation, and business applications (Ellis, 1991; Travis, Watson, and Atyeo, 1994). VR equipment and technology range from relatively inexpensive LCD shutter glasses to relatively expensive sensing and force-reflecting exoskeletons (Shaw, 1993). Three general categories of VR have emerged, each associated with various combinations of VR technology: desktop VR, augmented reality, and immersive VR.

Desktop VR is commonly associated with CAD/CAM applications and usually involves the use of a liquid-crystal shutter (either mounted in glasses on the user's head or on the display itself) to achieve stereopsis, the presentation of a unique image to each eye. Stereopsis is used to create retinal disparity — a difference between the images each eye sees that is arguably a key component of depth perception at near distances (see Patterson and Martin, 1992). Essentially, this liquid-crystal shutter does the same thing as red and blue glasses did for 3D moviegoers back in the 1950s (albeit usually much better). When combined with head-coupling technology, desktop VR has also been called fishtank VR (Arthur and Booth, 1993).

Augmented reality refers to systems that superimpose VEs (or elements of VEs) upon a user's real environment. These systems use see-through displays and may or may not use head-tracking depending on the application. An example of one of these systems would be one in which a surgeon about to operate on a patient could see virtual internal organs and bones superimposed on the patient's real body with the virtual images based on actual data from conventional radiology techniques. Unfortunately, the speed and accuracy of currently available tracking technology are not adequate for most tasks to which this concept would be applied, including the example just given (Azuma, 1993). Also, there are significant technical obstacles to be overcome in matching vergence and accommodative cues between virtual and real stimuli.

Immersive VR is VR as it was originally conceived and VR as it is popularly portrayed in mass media. It involves the replacement of a user's real environment with a virtual environment. This VE can be synthetic or real. The latter implies telerobotics and teleoperations and, in this context, the term telepresence — the illusion of presence at a real, remote location — is often used. More often, the VE is generated by a computer. The hardware used to achieve immersive VR usually includes a powerful image generator, a head-mounted display (HMD), and a head-tracking system. A wide variety of other hardware components may be added to assist the user in navigating the VE or achieving immersion. These include 3D positioning devices, 3D sound spatializing equipment, gloves to sense the position of a user's hand and/or fingers, and exoskeletons or other force-reflecting control devices to simulate the force feedback resulting from normal interaction with solid objects. Almost invariably, these components are cumbersome and lack the speed and resolution that users experience in everyday interaction with the real world. However, technological advances in many of these areas — both evolutionary and revolutionary — seem to be occurring at rapid rate (Azuma, 1993; Kollin, 1993; Post, Sarma, Heinze, Ellis, Larson, Franklin, Trimmer, and Rodgers, 1994).

Problem Statement

Given that presence is the primary goal of VR and its defining characteristic as a human-computer interface, the current lack of understanding and empirical observation concerning this phenomenon and its relationship to user performance are remarkable. Held and Durlach (1992) write that, "An important obstacle at present to scientific use of the telepresence concept is the lack of a well-defined means for measuring telepresence... What is now needed is a systematic research effort designed to gain an understanding of the sensorimotor and cognitive factors that determine the sense of presence." Similarly, Barfield and Weghorst (1993) state that a series of studies using psychophysical techniques needs to be performed to evaluate the subjective level of presence reported by the VE participant as a function of a number of VE parameters. Sheridan (1992) cites the need for research on the connection between presence and performance. Kalawsky (1993, p. 81) writes:

> "Unfortunately, there is very little scientific literature that deals with this very important subject. This is further evidenced by the almost complete lack of objective measure... Ideally, we would like a set of repeatable objective measures for presence that indicates the degree of presence created by a particular system."

Travis, Watson, and Atyeo (1994) advocate a shift in focus from user immersion to user performance as set forth in usability specifications (e.g., Hix and Hartson, 1993). However, as these authors admit, exactly what the criteria should be for user performance in VEs is an unresolved issue. There seems to be a large gap in the VR literature concerning a practical measure of presence, the effects of VR system parameters on presence, and the relationship between presence and performance.

The purpose of this dissertation is to at least partially fill this gap. Specifically, the goals of this research are to:

1) Provide data on a variety of VR system parameters with regard to their effects on subjective user presence and performance in a VE.

2) Test the psychophysical technique of magnitude estimation as a measure of presence.

3) Provide quantitative data to be used in making VR system design and purchase decisions — specifically, to build empirical models of the variables investigated. Ideally, a VR system designer interested in maximizing presence could use the results of this research as a basis for decisions concerning design tradeoffs (e.g., field of view vs. resolution).

3) Provide data with which to support and/or contradict current theories of presence.

4) Relate users' subjective feeling of presence in a VE to their performance in that environment.

Research Hypotheses

The research hypotheses being tested by this research are:

1) That the VR system parameters manipulated and analyzed in this research affect users' subjective feeling of presence in a VR system. The degree to which each of the parameters affects presence is used to differentiate between currently existing theories, frameworks, and hypotheses concerning presence.

2) That free-modulus magnitude estimation can be used to measure subjective presence in a VE. Furthermore, that this measure can be used to generalize and compare results across users, across experiments, across tasks, and across differently-configured VR systems.

3) That a positive relationship exists between a user's subjective feeling of presence in a VE and the user's performance of tasks within that environment.

4) That empirical models of the effects of the parameters manipulated can be constructed based on the results of this research that will predict both presence and performance in a VE based upon the parameters of the VR system in question.

BACKGROUND LITERATURE

This research consists of a series of studies involving a relatively large number of independent variables, psychophysical scaling of the primary dependent variable (presence), and multiple user tasks and associated dependent measures. The following literature review is divided into three parts: presence and telepresence, independent variables, and dependent variables. An effort has been made to conduct an exhaustive literature review concerning theories and measures of presence. With regard to the independent and dependent variables, many of these variables involve topics (e.g., stereopsis, visual search) that are areas of human factors specialization in their own right and only a representative sample of the literature directly relevant to the research conducted is reviewed here.

Presence and Telepresence

Simulation and Simulator Design

While presence per se has not usually been the focus of research in the field of simulation and simulator design, a number of studies have investigated the issue of simulator fidelity: how it relates to operator performance, to effectiveness of simulation and simulators, and how these are measured. This issue is closely related to the current topic of presence in VEs.

Like researchers in the field of VR, researchers in the field of simulation have often wrestled with the question, “How much realism is enough?” In particular, the need for realism in visual and motion cues seems to be much debated, with a dichotomy sometimes found between user performance and user preference or acceptance (e.g., Reid and Nahon, 1987; Rolfe and Staples, 1986). The answer to the question of how much realism is necessary seems dependent upon the user, the user’s task, and the purpose of the simulator. There are some who advocate the highest simulator fidelity possible (for performance appraisal especially) and the Federal Aviation Administration has set standards for simulator fidelity in simulators used to license or assess flight crew personnel (Ray, 1995). Examples have been cited in the fields of flight and anesthesiology simulation of instances in which less-than-perfect simulator fidelity has later lead to critical lapses in performance in real-world conditions (Ray, 1996). It has also been reported that, in maintenance training, user performance declines as physical fidelity between simulated and real environments decreases (Allen, Hays, and Buffardi, 1986). However, others have concluded that physical correspondence of simulation to reality is overemphasized for many purposes (especially training). In support of this view, studies are cited in which low-fidelity simulators are sometimes indistinguishable from actual equipment in terms of transfer of training (Jones, Hennessy, and Deutsch, 1985). Many researchers acknowledge the need for a better understanding of the relationship between simulator fidelity and user performance and have called for more research in this area (e.g., Jones et al., 1985).

The number of methods used to assess users’ subjective impressions of simulation fidelity has varied widely. Rarely, though, have users been directly asked about the realism of a simulation. More often, they have been asked to assess “adequacy”, “acceptance”, “sufficiency” or individual characteristics of the simulation. The methods used have ranged from hash marks on continuous lines (Kawahara, Kaoru, Watanabe, and Funabiki, 1996) to relatively untested rating scales (e.g., Reid and Nahon, 1987) to rigorously tested and validated rating scales (e.g., the Cooper-Harper Pilot Rating Scale and the Simulator Sickness Severity Index — see Frank, 1986 and Bailey and Knotts, 1987 for examples). There do not seem to be any published reports of studies directly investigating the connection between a user’s performance in a simulator and a user’s subjective impression of simulation realism.

In summary, although a great deal more research has been done in the field of simulator design than in the field of virtual environment design (and certainly more is known in the former area than the latter), the value of realism in simulation remains somewhat undetermined and seems to be dependent upon many variables. The approach most often taken by simulator designers seems to be that currently taken most often by virtual environment system designers: namely, to build as much realism into the system as time and money allow.

Theories

A number of theories, taxonomies, and hypotheses have been proposed concerning factors that influence presence in VEs and how it may be measured.

Held and Durlach (1992) consider the following items crucial to high telepresence:

- Wide range of sensorimotor interaction — when the head or eyes move, the visual scene must move correspondingly.
- Movement of viewed effectors — movement of the hand must correspond to the sight of the robot effector moving.
- High correlation between kinesthetically sensed movement and movement of the robot effectors — this correlation is hypothesized to be reduced by time delays, internally generated noise, and noninvertable distortions between the actions of the operator and the sensed actions of the slave robot. They state that, "How these variables interact, combine, and trade in limiting telepresence and teleoperator performance is a crucial topic for research."
- Similarity between visual appearance of the operator and the slave robot.

These authors further suggest, "Sensory factors that must certainly contribute to telepresence include high resolution and large field of view." Finally, they propose that telepresence can increase with operator familiarization with the teleoperation system through adaptation and learning of system characteristics. The telepresence system transforms the operator, his task, and his sensory environment, and the operator is required to model these transformations to interact with the VE. Presence is therefore dependent upon the success with which the operator models these transformations and this success is hypothesized to increase with user familiarity with the telepresence system.

Loomis (1992a, 1992b) looks at presence in terms of perception of self and nonself. He ties these perceptions to afference received in response to efference. The link between afference and efference can be nearly transparent (e.g., redirecting one's gaze and seeing a corresponding change in visual scene) or it can be almost entirely mediated (e.g., exploring the surface of a hidden object with a hand-held probe). Individuals often attribute sensations to externalized objects (in his words, "distal attribution"), even though there is something mediating the link between efference and afference. One example he cites is a Tactile Vision Substitution System in which users used a video camera to scan high-contrast objects. The signal from the video camera fed a matrix of vibrating stimulators placed on the observer's back or abdomen. Initially, users reported feeling only changing vibration patterns on the torso; however, extensive practice led some to experience the objects in front of the camera as external objects. He even cites an anecdotal case in which the camera lens was abruptly zoomed and an observer ducked in response to what appeared to be a looming object. The implication is that users can learn to be "present" or increase their sense of presence through learning and experience with unfamiliar characteristics of the linkage between efference and afference. It is hypothesized that presence is determined by the success with which a user models the linkage between efference and afference. It therefore follows that the harder this linkage is to model, the less presence will be felt or the longer it will take to build a sense of presence. Loomis equates "distal attribution" (perception of objects as external) with the sensation of presence in a VE.

Carr and England (1993) stress that realism cannot be considered as an objective, independent quality of the environment, but is a product of the user's interaction with the environment. Presence is only partially dependent on hardware and environmental parameters. They cite examples of movie-goers crying or exhibiting startle reactions despite relatively low sensory fidelity. They concentrate on theories of perception and compare two: direct perception and constructivist. According to these authors, direct perception theory holds that all experience with the environment results from direct perception of the qualities of the environment. It is suggested that there is an innate ability, not requiring interpretation or calculation, to see constants in sensory patterns and that these constants are key to correct perception of the environment. The implication of this theory in the current context is that presence will increase as the VE provides more and better cues (e.g., texture, optical flow, stereopsis, and other depth cues) that allow better modeling and perception of constants by the user. It also implies that artifacts that interfere with perception of these constants (e.g., spatial or temporal aliasing) will interfere with presence.

Constructivist theories describe perception in terms of exploration of the environment. People form hypotheses and inferences about the environment based on sensory input. Interpretation of sensory patterns is therefore learned — not innate — and can be modified through subsequent learning. As evidence in support of this interpretation, Carr and England (1993) cite the ability of people to adjust to viewing the world through prisms and learning to perceive the altered sensory pattern as reality (Harris, 1965). The implication of constructivist theories is that presence will increase as the user learns and models the altered sensory pattern provided by the VE hardware, that it will increase with users' opportunities to explore the environment (e.g., with time spent in the environment and with greater number and variety of possible interactions), and that it will decrease with inconstancy in the VE (e.g., if head-tracking suddenly changed from 2 DOF to 3 DOF or if a bouncing ball uncharacteristically shattered as it hit the virtual floor). It also follows that the extent of duplication of the real environment by the VE will be positively related to presence: presence will increase as the amount of perceptual learning required by the transition to the VE decreases.

Zeltzer (1992) suggests a taxonomy of graphic simulation systems with three orthogonal axes: autonomy of objects/persona in the environment, interaction between the user and the environment, and user presence in the environment. He suggests that presence is independent of autonomy of virtual objects and user interaction with the environment and dependent entirely upon the number and fidelity of sensory input and output channels. He highlights the connection between presence and the user's task and contends that the most important aspects of sensory input and output with respect to presence will be the ones most important to the task performed.

Steuer (1992) discusses presence as dependent upon the following two environmental variables:

- Vividness: The ability of the apparatus to produce a sensorially rich mediated environment. This is broken down into sensory breadth (the number of senses stimulated) and sensory depth (the resolution of the sensory channels stimulated).

- Interactivity: The extent to which users are able to modify the form and content of the mediated environment in real time. He breaks this down into speed (the rate at which input can be assimilated into the mediated environment), range (the number of possible interactions available to the user), and mapping (the extent to which interaction with the mediated environment is natural and predictable).

Both vividness and interactivity are stimulus-driven and, as such, dependent upon the implementing technology. He calls for research into the interaction between sensory breadth and sensory depth in determining presence and proposes that the contribution of sensory breadth to presence is likely to be highly task-dependent. For example, sensory breadth would be less likely to contribute to presence for a primarily visual task than it would for a task more dependent upon multisensory information. Finally, he states that the presence of other users in the VE should increase a user's sense of presence due to the normal experience of interacting with others in the real world and implies that, because of this, the absence of others could aid in making a virtual experience seem unreal.

Schloerb (1995) describes a theory of presence in which presence is either objective or subjective. He defines objective presence as the case in which a teleoperator is objectively present (causally interacts in some manner) with a remote environment. The degree of objective presence is defined as the probability that a specified task will be completed successfully in the remote environment. He suggests that many different kinds of objective presence are possible, corresponding to many different tasks. Subjective presence results from a teleoperator (or VE participant) feeling present in a remote environment or VE. The degree of subjective presence is defined as the probability that a user will perceive that he is physically present in a VE. Subjective presence is operationally defined in terms of signal detection theory using the probabilities of a user responding that he is physically present in a VE when he either is or is not. This definition requires a very close match (to cause uncertainty) between the virtual and real environments. This requirement causes the use of this measure to be severely limited, making it impractical for use with the vast majority of VEs and VR systems in use today (and in the foreseeable future). Unless a VR system approaches reality, a truthful participant will always respond that he or she is physically present in the real environment. Also, the method does not allow testing of VEs or interactions within VEs that cannot be matched in the real environment. Many VE applications include components that cannot be duplicated in a real environment. For example, this method could only be applied to VEs in which the participant's mode of locomotion is one that is feasible to duplicate in a real environment. Finally, there is the problem (admitted in his paper) of what do between trials — put the patient to sleep to prevent sensation of movement or lack thereof? While theoretically interesting, this paper and the method described seem extremely limited in practical use.

Citing a need for research in this area, Sheridan (1992) states that presence may range from helping performance to being a distraction. He suggests as measures of presence subjective ratings similar to those used in mental workload (i.e., multifactor rating procedures), a participant's response to unexpected or threatening stimuli, and socially conditioned responses to virtual social encounters (e.g., saying "Gesundheit" in response to a virtual actor's sneeze). He proposes three principal determinants of presence: 1) extent of sensory information, 2) control of sensors in relation to the environment (e.g., head-tracking), and, 3) ability to modify the environment. It is suggested that efferent and afferent distortions affect presence, training efficiency, and task performance, but that the extent of these distortions is largely unknown and unresearched with the exception of time delay. He advocates training efficiency and task performance as performance measures to be associated with presence in future research. In his latest paper on this topic, Sheridan (1996) reiterates these ideas and extends his proposed measures of presence to include one that would equate presence with the probability of a user discriminating between real and virtual environments. He cites Schloerb (1995), but, rather than improving a virtual environment until it becomes indistinguishable from a real environment, he advocates degrading the real environment until it becomes indistinguishable from the virtual. While this seems a more realistic and feasible approach than that proposed by Schloerb, it is still difficult to imagine a system that could add noise to a real environment — especially in dimensions such as sensor control and ability to modify the environment — that would not be prohibitively expensive and complex, especially given currently available technology.

In a response to Sheridan (1996), Ellis (1996) points out that, for Sheridan’s measures and three-determinant concept of presence to be useful, the dimensions must be quantifiable (preferably, in his words, into “isopresence conditions”). He also questions the entire premise that presence should be a goal of VE designers, contending that the usefulness of presence in a VE application should be known before it is made a goal of VE system design. In support of this last contention, he cites examples of interfaces and tasks in which user performance was improved by altering the interface in ways actually designed to *decrease* presence (e.g., changing the user’s frame of reference from egocentric to exocentric).

Data

There are surprisingly few studies that have attempted to collect data concerning the experience of presence. Those that exist have largely been of an exploratory nature. Although many authors have stated the need for studies of factors influencing presence and/or a replicable measure of this phenomenon, few have yet been published.

Heeter (1992) separates presence into personal presence (extent of afferent/efferent fidelity combined with suspension of disbelief), social presence (extent to which virtual actors react to the user's presence in the VE), and environmental presence (extent to which the environment reacts to the user's presence within it). She quotes anecdotal evidence for responsiveness to movement as more important than display resolution in determining personal presence. She also cites anecdotal evidence for head-tracking and representation of a virtual body as important to personal presence. She suggests, like many authors, that experience is also a factor determining personal presence. Her suggestion of "social presence" implies that acknowledgment by others (real or simulated) of the user's presence within the virtual world strongly contributes to presence. She even speculates that an environment that is more responsive to the user than the real world (e.g., a room that verbally greets the user as it is entered) could evoke greater presence than an environment that attempts to strictly model the real world.

In exploring these hypotheses, she conducted questionnaires of BattleTech and ENTER 3D participants. BattleTech is a location-based entertainment application of VR in which the user enters a simulation of the cab of a combat robot and interacts with the VE and other participants through the displays and controls of the simulator. ENTER 3D is an application in which users stand in front of a blue screen and see themselves superimposed upon the VE on a television monitor several feet away. The virtual world then interacts with the image of the user on the monitor. One common, strong finding from both questionnaires was participants' preference for participation with others in the virtual world. ENTER 3D participants expressed a strong desire for more interactions with the virtual world and more complex interactions. It should be noted that both of these applications are quite different from immersive VR applications involving an HMD and head-tracking, and the results of these questionnaires may not be generalizable to immersive VR applications.

Barfield and Weghorst (1993) propose a conceptual framework for presence in which attentional resources mediate the influence of several factors on presence. They emphasize that attention must be allocated to an aspect of the VE for it to affect presence. Factors affecting presence in their framework include display fidelity, environmental stability, sensory bandwidth, interactive fidelity, person variables, task variables, and context variables. Potential indicators of presence suggested include subjective assessment, physiometric indicators (e.g., heart rate variability), virtual world task performance, natural world task performance, frame of reference conflict resolution (i.e., which way a conflict is resolved when information from the real and virtual worlds conflicts), and context reorientation time or degree of disorientation in transitioning between real and virtual environments. Among the display fidelity examples they cite are spatial resolution, contrast resolution, field of view, optical distortion, stereopsis, and lighting and shading models. They suggest dual-task measures and workload-type measures as presence measures based on the assumption that greater presence is associated with allocation of more attention to virtual stimuli. Unfortunately, there is the danger that the outside distractions associated with these measures could decrease or even eliminate the very phenomenon they attempt to measure.

Barfield and Weghorst (1993) conducted two studies. In the first study, subjects were given interactive flythroughs of two virtual worlds. A subsequent questionnaire included three presence-relevant ratings (rating scale was 1-10): "Sense of being there", "Sense of inclusion in the virtual world", and, "Sense of presence in the virtual world". They found relatively low correlations (absolute value ranging between 0.20 and 0.50) between these measures and other items. The general pattern of their results indicates a negative relationship between presence and being lost, and positive relationships between presence and equipment comfort, ease of navigation, visual display color quality, clarity of the visual image, and overall enjoyment of the experience. Their second study investigated primarily social issues with mixed results. The one relatively clear finding of this study was a positive relationship between feeling like a part of the virtual world and the enjoyment subjects derived from designing and building the virtual world.

Waldern (1993) discusses VE strategy and gives anecdotal evidence concerning presence resulting from experience with entertainment applications of VR. He contends that the following increase presence:

- Sensory fidelity, with emphasis on the importance of updating the VE in real time and avoiding perceived lag.

- Presence of other users within the environment.

- Consistency of sensory modalities. He uses examples of, 1) walking into a virtual cave and hearing one's own voice and the voices of others begin to echo, and 2) hearing one's own voice and the voices of other players modified to match visual appearances within the virtual world (e.g., the voice of the person playing a tall, burly warrior should be modified so that everyone in the VE hears it with deep, bass tones added).

Slater and Usoh (1993a, 1993b, 1993c, 1993d) discuss presence in terms of exogenous (technology-dependent) and endogenous (participant-dependent) factors. They suggest the following factors as important to exogenous presence:

- High resolution information presented to the appropriate sensory organs, with the information received through channels to all sensory organs describing a consistent world.

- Information free from signals that indicate the existence of the input devices or display.

- A wide range of interactions based on movement of the subject's sensory organs, with the operator able to see the effect of moving his limbs in the VE.

- A high correlation between movements of the operator sensed directly and the actions of the representation of the person in the VE.

- An ability to change the VE.

- A similarity in visual appearance between the subject and his representation in the VE, including the subject's identification between his own body and that of the representation.

- Adaptation through learning over time, and consequent increase in subject familiarization with the relationship between motor actions and feedback through the input channels to the senses.

- That the feedback loop from operator to system to operator forms a consistent and lawful whole (afference lawfully related to efference and vice-versa).

- That the linkage between afference and efference be simple enough for the subject to model given time.

- Objects in the VE spontaneously respond to the subject.

- Participants' ability to move around in the VE.

- Presentation of a scenario or events that, if presented in everyday reality, would normally result in an assessment of their personal relationship to those events; in particular, the implications for their own state of mind or body. Facing the subject with a potential danger is one way of realizing this.

The four papers cited above resulted from one exploratory study. The independent variable manipulated in this study was representation in the VE of the user: by means of a virtual body or simply as a 3D arrow. Participants went through six virtual rooms in which they:

- Navigated through a room cluttered with furniture.

- Reacted to objects flying toward their face or body.

- Were turned virtually upside down.

- Built a structure from blocks.

- Walked a plank over a virtual precipice and dropped a rock onto a checkerboard-patterned floor below.

The dependent variables in this experiment consisted of subjects' responses on questionnaires. In one part of the questionnaire, subjects were asked to rank perceived presence on a six-point scale. Another part of the questionnaire asked two free-response questions concerning factors that increased or decreased presence. In addition, they tracked users' responses (physical and verbal) to dangerous virtual stimuli (i.e., objects flying at them, the virtual precipice), user reaction to an outside noise caused by the experimenter, and reaction to a socially-conditioned response (being asked the time by a virtual actor).

Their results are limited by the preliminary nature of the study: a number of things they expected to happen did not, and vice-versa. For example, subjects often were turned the wrong way when objects flew at them in the flying object room and so did not see the objects. Also, experimenters noted that, when asked the time by the virtual actor, many subjects looked for their watches and were confused by the absence of a virtual watch on their virtual wrists. The authors conclude that this particular metric destroyed the phenomenon it was attempting to measure by calling subjects' attention to the discrepancy between the real and virtual worlds.

Slater and Usoh's results may be summarized as follows (the percentages reported are based on 17 subjects):

Table 1. Slater and Usoh's results.

To what extent did you experience a sense of being "really there" inside the virtual environment?		Were there any circumstances that especially *increased* your sense of being "really there"?		Were there any circumstances that especially *decreased* your sense of being "really there"?	
1) Not at all really there	5.9%	1. Being able to move around	35.3%	1. Outside events (including instructor)	23.5%
2) There to a small extent	11.8%	2. Interacting with objects/doing a task	41.2%	2. Screen / updates / lag / resolution	35.3%
3) There to some extent	29.4%	3. Great concentration	5.9%	3. Things don't behave naturally (laws of physics are violated)	58.8%
4) A definite sense of being there	17.7%	4. Mention of body (exp. group only)	17.7%	4. Things aren't done naturally	52.9%
5) A strong experience of being there	29.4%	5. Being on plank ("fear reaction")	64.7%	5. Body doesn't behave naturally (exp. group only)	17.7%
6) Totally there	5.9%	6. Being upside down	5.9%		

Based on these results, responses to other questions on the questionnaire, and anecdotal reports by subjects and experimenters, they reach the following conclusions:

- Presence of a virtual body increases presence, moreso for females than males.
- Presence increases with perceived risk to the virtual self, interaction with the environment, and sensory fidelity.
- Presence is decreased when attention is drawn to outside events.
- Presence is decreased by unrealistic behavior of objects (including the virtual body) in the environment.
- Presence was lower in individuals who mentioned problems with the display (e.g., update rate and/or resolution).

- Self-report of tendency toward motion sickness was positively related to reported sense of presence.

In a later study, Slater, Usoh, and Steed (1994) had twenty-four subjects engage in a virtual fairy-tale adventure. They manipulated whether gravity applied to objects or not, whether subjects saw a visual cliff, whether a virtual actor followed the subject during part of the adventure, how many environments subjects were exposed to (2, 4, or 6), and whether subjects transitioned from environment to environment via doorways or by donning another — virtual — HMD. The measures of presence in this study were subjects' responses to the following three questions on an anchored 7-point scale:

- In the computer generated world I had a sense of "being there" ... (1 = not at all, 7 = very much).
- To what extent were there times during the experience when the computer-generated world became the "reality" for you, and you almost forgot about the "real world" outside? (1 = at no time, 7 = very much).
- When you think back about your experience, do you think of the computer-generated world as something that you saw, or more as somewhere that you visited? (1 = something that I saw, 7 = somewhere that I visited).

They set the presence score for each person as the number of responses greater than or equal to six on all three questions. This yielded a range of scores for each subject from zero to three. Based on this measure, they found no significant effects (a logistic regression with $\alpha = 0.05$ was used) of any variables except for mode of transition between environments. The visual cliff, the virtual actor, and presence of gravity did not significantly affect subjects' responses. This seems to directly contradict their previous findings (Slater and Usoh, 1993a, 1993b, 1993c, 1993d). Time spent in the environment also was not a significant factor. They expanded their measure to include responses greater than or equal to five and constructed a combined score based on a principal components analysis of the raw scores. In both cases the results (or lack thereof) were the same as for the initial analysis. In light of this, it seems difficult to reconcile this study with their previous work.

Singer and his colleagues (Singer and Witmer, 1995; Singer, Witmer, and Bailey, 1994) have focused on the formal construction and validation of a Presence Questionnaire. They administered this questionnaire to subjects in four experiments, conducted by various researchers and using various environments (one of which was a superset of the VE used in the current experiments). Based on a cluster analysis of these results, they report finding three subscales within the questionnaire that they have labeled "Involved/Control", "Natural", and "Interface Quality". The Involved/Control cluster contains items that ask respondents how well they felt able to control events in the VE, whether the VE seemed responsive to actions, the extent to which the visual aspects were involving, and how involved in the experience the respondent became. The Natural items involve the extent to which interactions felt natural, the extent to which the VE was consistent with reality, and how natural movement control was. Interface Quality items address whether controls interfered with or distracted from task performance, whether the display interfered with or distracted them from task performance, and the extent to which they felt able to concentrate on tasks. Based on statistical analyses of the Presence Questionnaire, they conclude that it is a reliable scale that measures at least some aspects of presence. While the Presence Questionnaire seems a useful tool for studying presence, it has the following disadvantages:

- It necessitates some delay between the experience of presence and subjects' report of presence. This raises the possibility of primacy and recency effects in recall of the experience and allows subjects to forget (perhaps key) aspects of their experience.

- It necessitates interrupting or ending subjects' experience in the VE and removing them from the VE to report their sense of presence.

- It is less efficient than the method (magnitude estimation) proposed for use in the current research, requiring more time for both measurement and analysis.

Independent Variables

Independent variables were chosen on the basis of two general criteria: relevance to the theories of presence reviewed above and relevance to VR system design. One of the primary goals of this research was to provide VR system designers with a useful design tool in the form of empirical models of presence. Each of the eleven independent variables manipulated (with the possible exception of virtual personal risk) is associated with some cost. The cost might be measured monetarily, in processing time, in programming time, or in system weight or complexity, but each is involved in some tradeoff that a VR system engineer must consider during system design. The independent variables were manipulated in a series of three experiments. To the extent that experimental design constraints and variable levels allowed, an attempt was made to group variables in these experiments so that key design tradeoffs (e.g., field of view vs. display resolution) could be directly examined. The following review groups these variables into the same groupings used in the experimental series.

Experiment 1: Field of View, Resolution , Scene Update Rate

These three variables were manipulated in Experiment 1 to determine their joint and separate effects on presence and performance. The primary reason for including these variables in this experiment was because they are all continuous. In addition, there is a direct tradeoff between field of view and resolution in most HMDs: as one adjusts the optics of an HMD to achieve a larger field of view, the size of each pixel in the display is also enlarged and the resolution of the image is decreased. VR system designers must generally decide how much of one is desired at the expense of the other. These three variables have been proposed by a number of researchers as factors that influence the depth of presence in a VE, often under the general rubric of realism or sensory fidelity. Therefore, manipulation of these variables would seem valuable from the standpoint of testing the sensitivity of magnitude estimation as a measure of presence and discriminating among existing taxonomies and theories of presence.

In one oft-cited study, Hatada, Sakata, and Kusaka (1980) studied the effect of field of view of two projected static scenes (a bridge and an open field) on subjective image tilt, head and eye movement, and "sensation of reality" as measured by a subjective seven-point scale ranging from -3 to +3. They describe no statistical analyses, description of the population sampled, or sample size. However, the graph of their results with respect to field of view appears to indicate the beginning of an asymptote for "sensation of reality" at approximately 62˚. There is only one data point after this (at approximately 72˚) and it is therefore difficult to discern whether this asymptote is actual or an artifact resulting from sampling error.

Psotka, Davison, and Lewis (1993) contend that immersion may be defined as the degree of compatibility between the location of self in the real world and the location of self in a virtual world, and equate high compatibility with high immersion or presence. They had 14 subjects view a virtual room from four different geometric fields of view (GFOVs) and two fields of view. The subjects' viewpoint circled the room as they did this and they were later asked to draw the path of the viewpoint on a 2D overhead picture of the room. Their results indicate that subjects appeared to use the frame of the monitor as the frame of reference for their visual field: they treated the contents of the frame as if it were their entire 200˚ field of view. Perceived station point of the viewpoint (in mm) corresponded to (180 ÷ GFOV) times the actual station point of the viewers' eyes. Unfortunately, their paper contains no statistical analysis and they did not attempt to directly measure subjective presence in the VE. Also, the presence of a monitor (and frame) would seem to hinder generalization to immersive VEs.

McGreevy (1993, 1994) conducted ethnographic studies of field geologists at work in real terrain with and without mediation by an HMD. In the first study he placed two subjects in "a Mars-like terrain." The HMD used displayed the input of two video cameras mounted on the subject's head. This HMD was monochrome, had a field of view of 25˚ H x 20˚ V, addressability of 200 lines (6 arcmin/line), and stereoscopic capability; however, the mode used in the studies was biocular (nonstereoscopic). His results are entirely anecdotal and descriptive due to the nature of the studies (ethnography). Subjects reported strongly negative feelings toward the narrow field of view in connection with their ability to navigate, maintain spatial orientation, and establish the context of the environment (i.e., maintain a mental model of the site). They compensated somewhat by increasing scanning behavior and "measuring" distances in paces. They also compensated for limited resolution by moving their heads closer to objects of interest; however some tasks (e.g., differentiation of sand and silt types at a distance) were rendered extremely difficult despite these compensation strategies. Both subjects felt that their task could not be performed without color. Based on these observations McGreevy stresses the importance to presence and performance of continuity of, 1) the environment, 2) objects within the environment, 3) task context and context-component relationships, and, 4) interaction with or observation of the environment. McGreevy's second study (1994) again involved two field geologists, but without any interference or environmental mediation. This data also consisted of interviews and observations. He concludes that these observations also support his earlier hypotheses concerning continuity.

Wells and Venturino (1990) studied the effect of HMD field of view on response time in a visual search task involving target detection in a relatively sparse visual environment with a variable number of distractor targets. The fields of view tested were 20˚ H x 20˚ V, 45˚ H x 42.5˚ V, 60˚ H x 50˚ V, 90˚ H x 60˚ V, and 120˚ H x 60˚ V. However, this manipulation was done in a somewhat unconventional manner: the specified fields of view were more "fields of target visibility" that were superimposed on a 120˚ H x 60˚ V view of the VE, which was visible throughout each experimental session. Because of this, Wells and Venturino's subjects may not have experienced the same negative effects of narrow field of view with regard to spatial orientation and environmental context described by McGreevy (1993, 1994). Wells and Venturino (1990) found that search time increased with decreasing field of view and that this effect interacted with the number of distractors: the negative impact of decreased field of view was greater in conditions with more distractors. In the condition with the greatest number of distractors (three targets and six distractors), performance in the narrowest field of view condition was poorer by roughly 25% in comparison to the widest field of view condition. No significant effect of field of view was found in the condition without distractors.

In a similar study, Piantanida, Boman, Larimer, Gille, and Reed (1992) examined the effect of field of view on a visual search task using a somewhat low-resolution (13 arcmin/pixel) visual display. The fields of view studied were 14˚, 18˚, 41˚, 53˚ (all using a circular aperture), and 80˚

H x 60˚ V (using a rectangular aperture — the limit of the HMD used). Three of the authors served as subjects. They found significant effects of field of view and presence of distractors, but only for distractors that were the same color as the target. Plots of search times by field size (in square degrees) for each subject were well-fit by parallel lines in log-log coordinate plots. The slope of these lines, and thus the exponent of the power function describing the relationship to reaction time was roughly -0.4. This was also true for a reanalysis they performed of the data reported by Wells and Venturino (1990).

Cha, Horch, and Normann (1992) studied field of view and pixelized vision in a mobility task similar to the locomotion task used in the current research. The environment used was relatively sterile, consisting entirely of white backgrounds and black obstacles; however, they found that near-normal walking speeds could be achieved through an obstacle course with a field of view of 30˚ and a 25 x 25 array of pixels. They note that increases in field of view or the number of pixels beyond these levels had little effect on walking speed. An interesting finding reported in this paper is subjects' ability to adapt to the pixelized, narrow-field-of-view vision system used. During 50 traversals of the obstacle course, subjects' average speed increased from 0.2 to 0.6 meters per second, while the average number of contacts with obstacles decreased from nine to three. This finding suggests an ability to adapt to sensory limitations such as those normally associated with current VR technology. Draper, Fujita, and Herndon (1987) found that increasing display resolution (from NTSC to HDTV — approximately doubling the number of raster lines) decreased errors in remote handling and remote inspection using a teleoperation system. In general, it has been a common finding that increasing display resolution (or image quality) leads to corresponding increases in performance, at least for moderately demanding visual tasks (Snyder, 1988).

In the third of three experiments, Arthur and Booth (1993) manipulated two components of lag: scene update rate (a.k.a., frame rate) and tracker lag. Scene update rate was 10, 15 or 30 Hz. Subjects performed a 3D tree-tracing task with stereopsis and head-coupling active throughout. They found that performance suffered with lag: a regression of log response time based on total lag yielded an R^2 of 0.57. This same regression using scene update rate alone yielded an R^2 of 0.45. The nature of the system (fishtank VR) and lack of dynamic stimuli (two static scenes were used) make generalizations to immersive VR tenuous. It seems likely that the effects of lag would be exacerbated by moving stimuli.

Ware and Balakrishnan (1994) report a similar study in which they manipulated frame rate to determine its effects on a 3D positioning task. Subjects moved a cursor to a target along either the X axis or Z axis in a variation of a Fitts' Law task using a hand-held, 6-DOF positioning device. They tested frame rates of 0.66, 1, 2, 3, 5, 10, and 15 Hz and found that average movement time asymptoted at a frame rate of between 5 Hz and 10 Hz, decreasing from seven seconds in the 0.66 Hz condition to roughly three seconds in the 5, 10, and 15 Hz conditions.

It has been stated that conventional wisdom for computer graphics holds that ten updates per second are required for perception of smooth motion, but that researchers in the field often admit informally that this is not really enough (Ware and Balakrishnan, 1994). Data from the simulation community indicate that user performance in and acceptance of a simulation begin to decrease with time delays beyond approximately 100 ms (e.g., Bailey and Knotts, 1987; Merriken, Riccio, and Johnson, 1987). Evidence has been found that delay in a simulator’s visual system has a greater negative impact on user performance than delay in a simulator’s motion system (Frank, 1986; Frank, Casali, and Wierwille, 1988). Increasing lag or decreasing scene update rate has been found, in general, to decrease performance of tracking and remote manipulation tasks (Held and Durlach, 1991).

Experiment 2: Sound, Textures, Head-Tracking, Stereopsis, Virtual Personal Risk

These five variables were manipulated in Experiment 2. They were selected for inclusion in this experiment because they could all be defined as present or absent and thus accommodate a 2^{5-1} fractional factorial design. The first four variables have been proposed by several authors to increase the realism or fidelity of a VE, thus increasing presence. In particular, head-tracking is a variable often cited as increasing presence, with reference made to the importance of the correspondence between what a user does and what a user sees (a.k.a., the connection between efference and afference). The inclusion of sound as a variable goes to the heart of Steuer’s (1992) concept of vividness and sensory breadth: that presence should increase as more senses are stimulated by a VE. Virtual personal risk is another variable proposed by several authors (e.g., Slater and Usoh) as a factor that should increase perceived presence. It was hoped that manipulation of these variables would therefore prove useful in testing the sensitivity and validity of magnitude estimation as a metric of perceived presence.

Arthur and Booth (1993) investigated stereopsis (as applied to fishtank VR) using a 3D tree-tracing task. In the first of three reported experiments, seven subjects used pairwise comparisons to establish subjective preferences concerning "best perception of 3D space" among a factorial combination of stereoscopic and head-coupled conditions. It was found that the order of preference was, 1) head-coupling without stereopsis, 2) head-coupling with stereopsis, 3) stereopsis without head-coupling, and, 4) neither.

Subjects performed the same task in a second experiment . Performance, as measured by time to complete the task and error percentage, was best in the head-coupled stereoscopic condition, followed by the head-coupled only, stereoscopic only, and no depth cue conditions. As the authors point out, the beneficial aspects of head-coupling, both for subjective preference and for performance are likely enhanced by the presentation of static scenes. Moving scenes could do for the user much of what he is forced to do for himself with head movements (e.g., assess parallax) when viewing a static scene.

Chung (1992) reports a study in which he compared head-tracked and non-head-tracked steering modes in a task involving targeting of radiotherapy treatment beams in a VE. Subjects ranked four head-tracked steering modes and three non-head-tracked steering modes in ease of use and preference. Their performance of the task was timed. He found no significant performance or preference differences between head-tracked and non-head-tracked modes, but noted a large amount of inter-subject variability that may have masked these effects. He also found no significant correlation between subjects' performance in a given mode and their rankings of preference and ease of use for that mode.

Ehrlich and Singer (1994) found that neither stereopsis nor head-tracking significantly improved performance of a variety of tasks in a VE. The tasks they used were tracking, locomotion, object manipulation, and distance estimation tasks from the same set of tasks used in the current studies (the Virtual Environment Performance Assessment Battery or VEPAB — proposed by Lampton, Knerr, Goldberg, Bliss, Moshell, and Blau, 1994). Similarly, Lampton, Knerr, et al. (1994) report that, contrary to their expectations, there was no apparent benefit of stereopsis in a distance estimation task even for stimuli at short distances.

Based on the results of a study designed specifically to test for differential effects of stereopsis, Hsu, Pizlo, Babbs, Chelberg, and Delp (1994) suggest that — when confounding influences such as varying viewing conditions, image intensity differences, ghosting, flicker, speed-accuracy tradeoff, subjects' stereoacuity, and the degree of task difficulty are removed — stereopsis increases the sensitivity and specificity of observer performance. The task they used involved detecting subtle features in simulated X-ray transmission images.

Combining a secondary task with a visual search task, Zenyuh, Reising, Walchli, and Biers (1988) found that stereopsis increased accuracy performance on the visual search task, but did not affect response time. The benefit of stereopsis increased with the difficulty of the search task (i.e., as the number of distractors increased). Stereopsis was associated with a 26% increase in accuracy in the condition with the greatest distractor density. Further, they found that including stereopsis as a depth cue resulted in a performance gain roughly equivalent to that produced by inclusion of relative size information.

The interaction of stereopsis with other depth cues was examined by Reinhart, Beaton, and Snyder (1990). Stereopsis, relative size, luminance, and interposition were factorially manipulated in a simple relative depth judgment task. They found that stereopsis did not significantly affect the speed of subjects' judgments, while the other three variables did. Error data showed a similar trend. In contrast, they found that stereopsis dominated subjects' ratings of depth imaging quality: subjects rated the image quality of the display with respect to ability to determine depth much higher in stereoscopic conditions. One notable difference between the studies of Reinhart et al. (1990) and Zenyuh et al. (1988) is that of task difficulty: this further supports the suspicion that beneficial effects of stereopsis increase with the difficulty of a subject's task.

Much work has been done on the achievement of 3D, externalized audio in VEs and it has been shown to enhance localization of virtual stimuli (see Wenzel, 1992). However, there have been no reports to date of examinations of the effects auditory stimuli on perceived presence or performance of nonauditory tasks. Anecdotal evidence that suddenly-deafened adults experience a decreased sense of presence in the real world, or a feeling “that the world had taken on a strange and unreal quality,” has been reported by Gilkey and Weisenberger (1995). Except for the works of Slater and Usoh (1993a, 1993b, 1993c, 1993d, 1994), there is a similar lack of work on effects of texture and virtual personal risk. Evidence from the flight simulation literature seems mixed with regard to the effects of texture-mapping (Padmos and Milders, 1992). Textures *have* been found to improve performance in simulated helicopter flight, nap-of-the-earth flight, and sailing (Padmos and Milders, 1992).

Experiment 3: Interactions, Other Users, Environmental Detail

These three variables were factorially combined in Experiment 3. "Interactions" as used here refers to the number of interactions possible between the user and the VE. Environmental detail is equated with the magnitude of the distancing factor in the VE software’s distancing algorithm. This algorithm replaces an object with a replica of greater or lesser detail (i.e., more or fewer facets) as the user’s viewpoint draws closer to or farther away from the object. As the specified detail level is increased, the distance from the viewpoint at which objects are replaced with more detailed replicas also increases. This is similar to the “level of detail” feature often found in simulator image-generating systems as described by Padmos and Milders (1992). Environmental detail directly affects the number of polygons processed by the image-generating system and therefore the speed with which a VE application runs. This, in turn, directly affects the rate at which a user receives feedback concerning actions (including navigation) taken within the VE. System costs associated with interactions and including other users consist mainly of increased development time and overall complexity of the system needed to program and run the VE application. Environmental detail, like several of the other variables manipulated in the this research, falls into the general category of realism or fidelity. However, the other two variables manipulated in this experiment are not generally viewed as falling into this category. Indeed, Zeltzer (1992) claims that they should have no influence on presence at all, while others (e.g., Heeter, 1992) have stated just the opposite. Manipulation of these variables therefore seems useful in helping to determine which of these views is most correct. It was expected that, to the extent that these variables do affect presence, they would prove useful in testing the measure of presence proposed in this research (magnitude estimation) and, consequently, in testing the relationship between presence and performance. This rationale holds, to some extent, for all eleven variables manipulated in the experimental series.

Other than that already reviewed, there has been very little work done examining the variables manipulated in Experiment 3. Atherton and Caporael (1985) report studying the effects of environmental detail by obtaining subjective judgments of the "form faithfulness" and aesthetic appeal of spheres composed of varying numbers of polygons. They found asymptotes in judged form faithfulness and aesthetic appeal at roughly 200 polygons, with ratings above the midpoint of their scale first occurring at approximately 100 polygons. The two judgments were highly correlated. While computer-supported cooperative work has been a topic of much research in recent years (e.g., Green and Williges, 1995; Wexelblat, 1993), there has apparently been little research done on the effects of other users on perceived presence or on performance of noncooperative tasks. Similarly, no research has yet been reported directly examining the relationship between the interactivity of a VE and the resulting sense of presence.

Time Spent in the Virtual Environment

While it was not manipulated, time spent in the VE was recorded during all three experiments so that an analysis of the effects of this variable could later be done. Other than the literature reviewed above (e.g., Cha et al., 1992; Slater and Usoh, 1994), no research has yet been published containing a direct examination of the effects of time spent in a VE on either presence or performance.

The current state of the literature with respect to theories and data concerning presence, then, may be summarized as follows:

Table 2. Current state of the presence literature.

Parameter	Proposed to Influence Presence By:	Relationship Based Upon Available Evidence Performance	Presence
FOV	Held and Durlach (1992) Barfield and Weghorst (1993) Sheridan (1992) Zeltzer (1992) Steuer (1992) Psotka et al. (1993)	positive (visual search)	positive
Display Resolution	Held and Durlach (1992) Sheridan (1992, 1996) Zeltzer (1992) Steuer (1992) Barfield and Weghorst (1993) Slater and Usoh (1993)	positive (for demanding visual tasks)	positive
Scene Update Rate	Held and Durlach (1992) Carr and England (1993) Sheridan (1996) Steuer (1992) Waldern (1993)	negative	negative
Sound	Sheridan (1992, 1996) Zeltzer (1992) Steuer (1992) Gilkey and Weisenberger (1995)	?	?
Textures	Carr and England (1993) Proposed by many cited authors under the general rubric of "sensory fidelity"	?	?
Head-Tracking	Held and Durlach (1992) Sheridan (1992, 1996)	positive	?
Stereopsis	Carr and England (1993) Barfield and Weghorst (1993)	mixed — positive for demanding tasks	?
Virtual Personal Risk	Slater and Usoh (1993)	?	mixed
Interactions	Sheridan (1992, 1996) Carr and England (1993) Steuer (1992) Slater and Usoh (1993) Singer et al. (1994)	?	positive
Other Users	Heeter (1992) Steuer (1992) Waldern (1993)	?	positive
Environmental Detail	Proposed by many cited authors under the general rubric of "sensory fidelity"	?	?
Time Spent in the VE	Held and Durlach (1992) Loomis (1992) Carr and England (1993) Slater and Usoh (1993)	positive	?

Dependent Variables

A primary goal of the research conducted was to establish a quantitative, replicable measure of the subjective experience of presence in a VE. Another goal was to relate this measure to performance in the VE. The dependent variables needed in the studies therefore included measures designed to satisfy these goals. These are described below.

Magnitude Estimation

It is no more possible to directly measure the sensation of presence in a VE than it is to directly measure sensations of loudness or brightness. A quantitative value may be assigned to stimuli presented to a subject and to responses made by a subject, but the subject remains something of a black box with respect to measurement of the sensations produced by the stimuli that (presumably) form the basis of the response. The magnitude of the sensation corresponding to any given value of a stimulus must be inferred from the subject's behavior. Mapping the relationship between physical stimulus and psychological response is the focus of the field of psychophysics. A great deal of effort has been devoted to this field since its inception in the mid-1800s (see Gescheider, 1985), resulting in a number of psychophysical techniques designed to associate quantitative values with subjective sensations. The purpose of these techniques is to quantitatively map the psychological response produced by physical stimuli. Magnitude estimation is one of these techniques.

Magnitude estimation is a form of "ratio scaling" or "direct estimation" proposed by S. S. Stevens (1953, 1955). In this technique, subjects are presented with a series of stimuli and asked to assign a number to each based on their subjective impressions of the intensity of each stimulus. The only restriction imposed on the numbers assigned is that they be positive. Two ways of implementing this technique are, 1) for the experimenter to define a modulus (or anchor) for the scale (e.g., the experimenter presents a standard stimulus and tells the observer that it has some standard numerical value) and, 2) for the experimenter to allow the observer to define his own modulus. In the latter case, the observer is told to assign any value to the first stimulus that seems appropriate and then assign successive numbers accordingly. This last technique is known as free-modulus magnitude estimation and is generally preferred over an experimenter-defined modulus (Gescheider, 1985; Stevens, 1971). Stevens (1971) recommends one judgment per stimulus per observer, taking the geometric mean of these judgments as the psychological scale value associated with each stimulus when combining data across observers. Use of the geometric mean assumes that the numbers produced by the observers represent a ratio scale (see Gescheider and Bolanowski, 1991b for a discussion of scales of measurement).

Gescheider (1985, pp. 174-181) reviews several examples of research that has successfully used this technique: magnitude estimation has been used to create psychological scales of warmth, pain, physiological and psychological stress, fear, and even drunkenness. Roberts (1979, p. 184) cites studies in which its use has been extended to the study of pleasantness of musical selections, seriousness of crimes, and even the desirability of wrist watches. Luce and Krumhansl (1988) include a table (reproduced below) of exponents of the power function relating subjective magnitude estimation to physical stimulus magnitude.

Table 3. Representative exponents of the power functions relating subjective magnitude to stimulus magnitude.

Continuum	Measured Exponent	Stimulus Condition
Loudness	0.67	Sound pressure of 3000 Hz tone
Vibration	0.95	Amplitude of 60 Hz on finger
Vibration	0.6	Amplitude of 250 Hz on finger
Brightness	0.33	5˚ target in dark
Brightness	0.5	Point source
Brightness	0.5	Brief flash
Brightness	1.0	Point source briefly flashed
Lightness	1.2	Reflectance of gray papers
Visual Length	1.0	Projected line
Visual Area	0.7	Projected square
Redness (Color Saturation)	1.7	Red-gray mixture
Taste	1.3	Sucrose
Taste	1.4	Salt
Taste	0.8	Saccharine
Smell	0.6	Heptane
Cold	1.0	Metal contact on arm
Warmth	1.6	Metal contact on arm
Warmth	1.3	Irradiation of skin, small area
Warmth	0.7	Irradiation of skin, large area
Discomfort, Cold	1.7	Whole body irradiation
Discomfort, Warmth	0.7	Whole body irradiation
Thermal Pain	1.0	Radiant heat on skin
Tactual Roughness	1.5	Rubbing emery cloths
Tactual Hardness	0.8	Squeezing rubber
Finger Span	1.3	Thickness of blocks
Pressure on Palm	1.1	Static force on skin
Muscle Force	1.7	Static contractions
Heaviness	1.45	Lifted weights
Viscosity	0.42	Stirring silicone fluids
Electric Shock	3.5	Current through fingers
Vocal Effort	1.1	Vocal sound pressure
Angular Deceleration	1.4	5-second rotation
Duration	1.1	White noise stimuli

The exponents in this table represent Ñ in Stevens' now-familiar power law: $B = é@^{Ñ}$, where B is the psychological scale value, é is a constant, and @ is the value of the physical stimulus. While Luce and Krumhansl (1988) advise caution in attaching meaning to the values of these exponents (they note a number of extraneous factors that can influence exponent values), this table serves as a good illustration of the wide variety of sensations to which magnitude estimation has been applied. Zwislocki and Goodman (1980) have even gone a step further: they provide evidence that, at least for relatively simple psychological scales of loudness and line length, magnitude estimation produces not just a ratio scale, but an absolute scale. They present data indicating that numbers themselves acquire subjective magnitudes some time before the age of six and that these magnitudes are fixed after that age. Their contention is that the subjective magnitudes of numbers are then matched to the subjective magnitudes of stimuli presented in a free-modulus magnitude estimation paradigm in what is essentially a cross-modality matching task, with "numerosity" as one of the modalities. They review a number of studies of loudness and show that a tendency toward an absolute scale is one of the sources of bias in magnitude estimation experiments that have used an experimenter-defined modulus. The notion of absolute sensory magnitudes and absolute magnitude estimation for line length and loudness has also been supported by Collins and Gescheider (1989).

In spite of this, one may question the meaning of the numbers assigned by individuals participating in a magnitude estimation experiment: because an observer assigns a value of 10 to one stimulus and a value of 100 to another stimulus, can it really be assumed that the sensation produced by the second stimulus is ten times greater than the sensation produced by the first? The preponderance of the evidence gathered to date indicates that the answer is yes (Roberts, 1979). A number of researchers have tested this assumption and found it to be valid, at least for group data. Gescheider (Gescheider, 1985; Gescheider and Bolanowski, 1991b) cites a number of studies that have supported this assumption by testing the additivity of magnitude estimation data. The general finding for data averaged over several subjects is that the estimated magnitude of two stimuli presented together (e.g., the loudness of two binaurally presented tones) is equal to the sum of the magnitude estimates resulting from each presented separately. In one key paper, Zwislocki (1983) demonstrated additivity of magnitude estimates for group data with respect to loudness and line length and showed that deviations from additivity in individual data were systematic. Individuals who demonstrated a bias in assigning numbers to perceived loudness demonstrated the same bias in assigning numbers to line length (as measured by the exponent of the power function relating magnitude estimates to stimulus magnitude). It has been suggested that some individuals exhibit systematic bias in assigning numbers to sensations and that this bias is a constant for the individual — that it does not vary with sensory modality. Although these differences have been shown to be small and to average out in group data, it has been recommended that a measure of distortion in the assignment of numbers be determined for each observer in a magnitude estimation experiment. This measure is usually based on magnitude estimates of line length and can be used to determine a bias correction factor for each observer, if needed (Collins and Gescheider, 1989; Gescheider, 1985; Gescheider and Bolanowski, 1991a). Further evidence for the validity of magnitude estimation and for its production of ratio-scale data comes from studies that have compared magnitude estimation with other psychophysical techniques (e.g., Verillo, 1991) and compared psychophysical functions resulting from magnitude estimation across multiple sensory modalities (e.g., Bolanowski, Zwislocki, and Gescheider, 1991).

Some examples of the use of magnitude estimation by human factors engineers include research on perceived image quality (e.g., Ansley, 1991; Sebok, 1991), perceived image motion (Miller, 1993), and, in an application particularly relevant to the current research, the perceived reality of a star field (Kinkade, Snyder, and Greening, 1963). Based on the broad success of magnitude estimation in such diverse applications, it seems reasonable to predict that it would be successful in mapping perceived presence in a VE.

Performance

Measurement of performance implies that some task is performed on which performance is measured. The task selection criteria used in the current experiments included the following:

- The task(s) performed should be the same in all three experiments to allow comparison of data across experiments and satisfy sequential experimentation constraints.

- The task(s) performed should be generalizable to other tasks and applications for which VR is commonly used.

- The task(s) performed should be sensitive to manipulation of the independent variables used and all variables must fit plausibly into the task(s).

- The task(s) should involve a relatively simple VE so that actual scene update rate can remain at the maximum scene update rate as specified in the experimental design (of particular concern when adding textures or increasing environmental detail).

- The task(s) performed should be relatively short.

This last criteria deserves some justification. Regan and Price (Regan, 1994; Regan and Price, 1994) have investigated the frequency of occurrence and severity of side effects associated with the use of an immersive VR system. Out of 146 subjects, 61% reported symptoms of malaise at some point during a 20-minute immersive and 10-minute postimmersive period. These symptoms ranged from dizziness, stomach awareness, headaches, eyestrain, and lightheadedness to severe nausea. The symptoms caused 5% of the subjects to withdraw from the experiment before completing their 20-minute immersive period. The highest percentage of subjects reporting symptoms at any point was 53% after 20 minutes of immersion (the limit of immersion in the study). Only 28% of subjects reported any sensation of nausea. It should be noted that their experimental procedure focused subjects' attention on possible symptoms and thus possibly increased subjects' awareness of these symptoms. However, Lampton, Kolasinski, Knerr, Bliss, Bailey, and Witmer (1994) report a series of experiments that did *not* focus subjects' attention on symptoms of malaise. They conducted four experiments and recorded the incidence with which subjects withdrew from the studies because of symptoms similar to simulator sickness. They also administered simulator sickness questionnaires following participation in each experiment. The experiments differed in the type of display used (a Virtual Research Flight Helmet vs. a Fakespace two-color BOOM), duration of immersion (from 20 min. to 80 min.), and subjects' task and environment (the first two experiments used the VEPAB, while the second two studies used a building walk-through). They found that between 4% and 16% of participants withdrew from the experiments because of symptoms similar to simulator sickness, that this withdrawal usually (94% of the time) occurred within the first 10 minutes of immersion, and that total severity scores on the simulator sickness questionnaire increased with time immersed. It appears from the graph that they provide that an asymptote in these scores may be reached some time between 30 and 80 minutes. In all of the experiments, subjects scored higher on oculomotor and disorientation subscales than on a nausea subscale. They found that display type affected the pattern and level of symptoms reported, while task and environment did not. In spite of these results, they note that most users reported enjoying the VE experiments while also reporting some level of discomfort. The results of Lampton, Kolasinski, et al. (1994) and Regan and Price (1994) seem more than sufficient reason to attempt to limit the amount of time subjects are required to spend immersed in a VE.

The task selection criteria used lead to the conclusion that it would be best to have subjects perform multiple tasks. Performing multiple tasks seemed especially fitting when considering the generalizability of experimental results: most VEs and VR applications are quite complex and involve performance of a variety of subtasks. Lampton, Knerr, Goldberg, Bliss, Moshell, and Blau (1994) propose a standardized set of 21 tasks (the Virtual Environment Performance Assessment Battery or VEPAB) to facilitate comparison of results between laboratories, hardware and software configurations, and across experiments. Task categories proposed include vision, locomotion, manipulation, tracking, and reaction time. Within each category they propose a variety of tasks. They evaluated all proposed tasks in two experiments and with two different input devices, but with visual display only. Their criteria for the usefulness of the tasks proposed were performance sensitivity to differences in input and output devices and sensitivity to practice effects. They obtained good results with most tasks, noting large differences in individual ability and large variation in individual ability. They report that the variation in individual ability decreased with time, but was not eliminated. A subset of these tasks was chosen as the overall task that subjects performed in the current series of experiments.

METHOD

Experimental Design

Three experiments were conducted under the constraints of a sequential experimentation paradigm (Han, 1991a, 1991b; Han, Williges, and Williges, 1990; Williges and Williges, 1989; Williges, Williges, and Han, 1992, 1993). Eleven variables were manipulated in three experiments. These variables were: scene update rate, visual display resolution, field of view, sound, textures, head-tracking, stereopsis, virtual personal risk, interactions, presence of a second user, and environmental detail. In addition, time spent in the VE was recorded for later analysis. The experiments and the manipulation of these variables are described below.

Experiment 1

Experiment 1 factorially combined scene update rate, visual display resolution (measured in arcminutes per pixel), and field of view. There were three levels of scene update rate (8 Hz, 12 Hz, and 16 Hz), two levels of visual display resolution (corresponding to pixel addressabilities of 320 H x 200 V and 640 H x 480 V), and three levels of field of view (48˚ H x 36˚ V, 36˚ H x 27˚ V, and 24˚ H x 18˚ V). There were a total of 18 conditions in this experiment (3^2 x 2). Figures 1 and 2 show sample scenes from the VE in the two levels of visual display resolution, while Figures 3 and 4 show the same scene in the medium and low field of view conditions. Three levels of scene update rate and field of view were employed so that the quadratic effects of each could be tested. The two levels of visual display resolution tested represent the full range of capability of the VE hardware/software system used in the experiments. The factors not manipulated in this experiment were set at the following levels: sound = present, textures = absent, head-tracking = present, stereopsis = absent, virtual personal risk = absent, interactions = medium, second user = absent, and environmental detail = medium. This experimental design is summarized in Table 4.

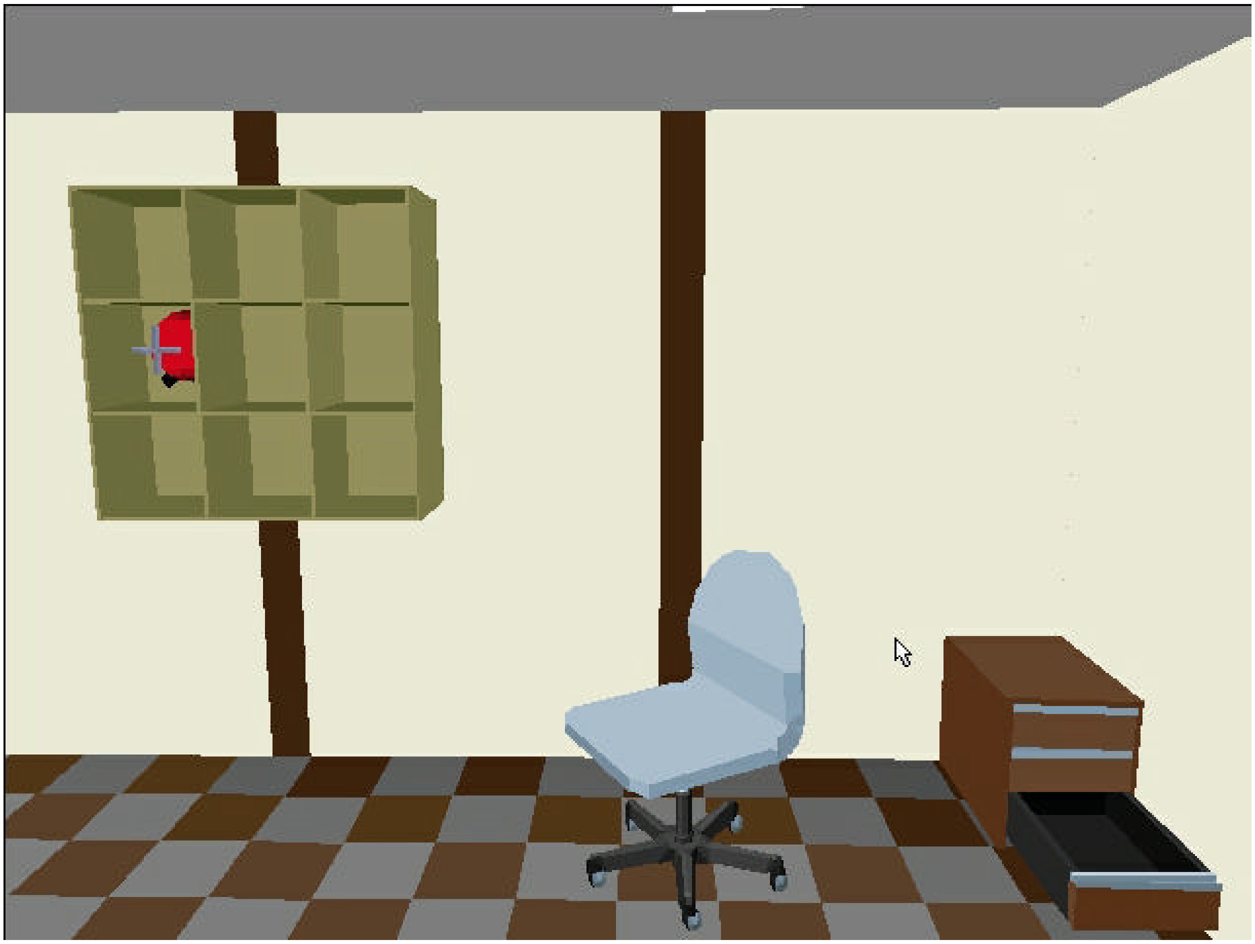


Figure 1. Sample scene from the VE in the high visual display resolution condition.

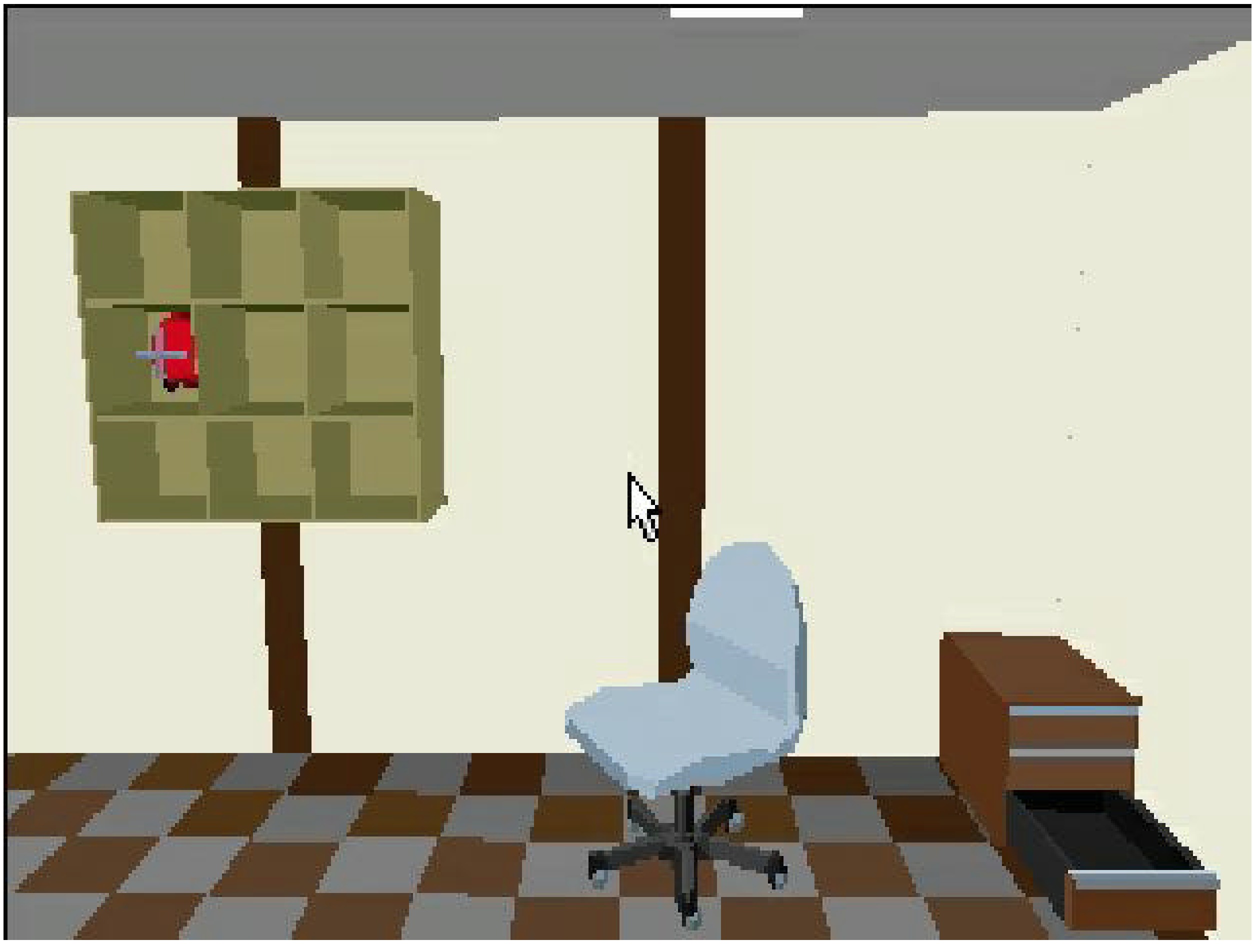


Figure 2. Sample scene from the VE in the low visual display resolution condition.

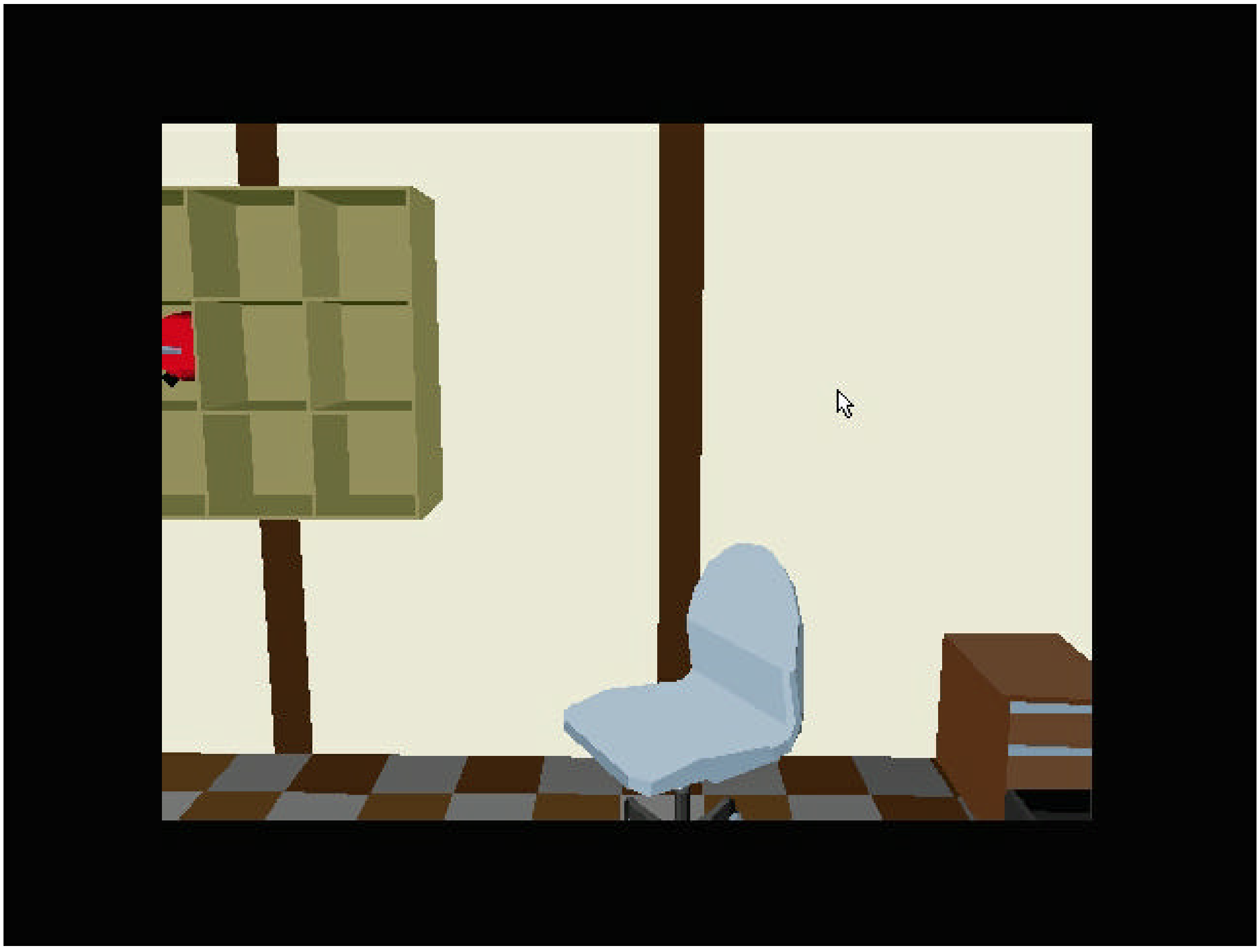


Figure 3. Sample scene from the VE in the medium field of view condition.

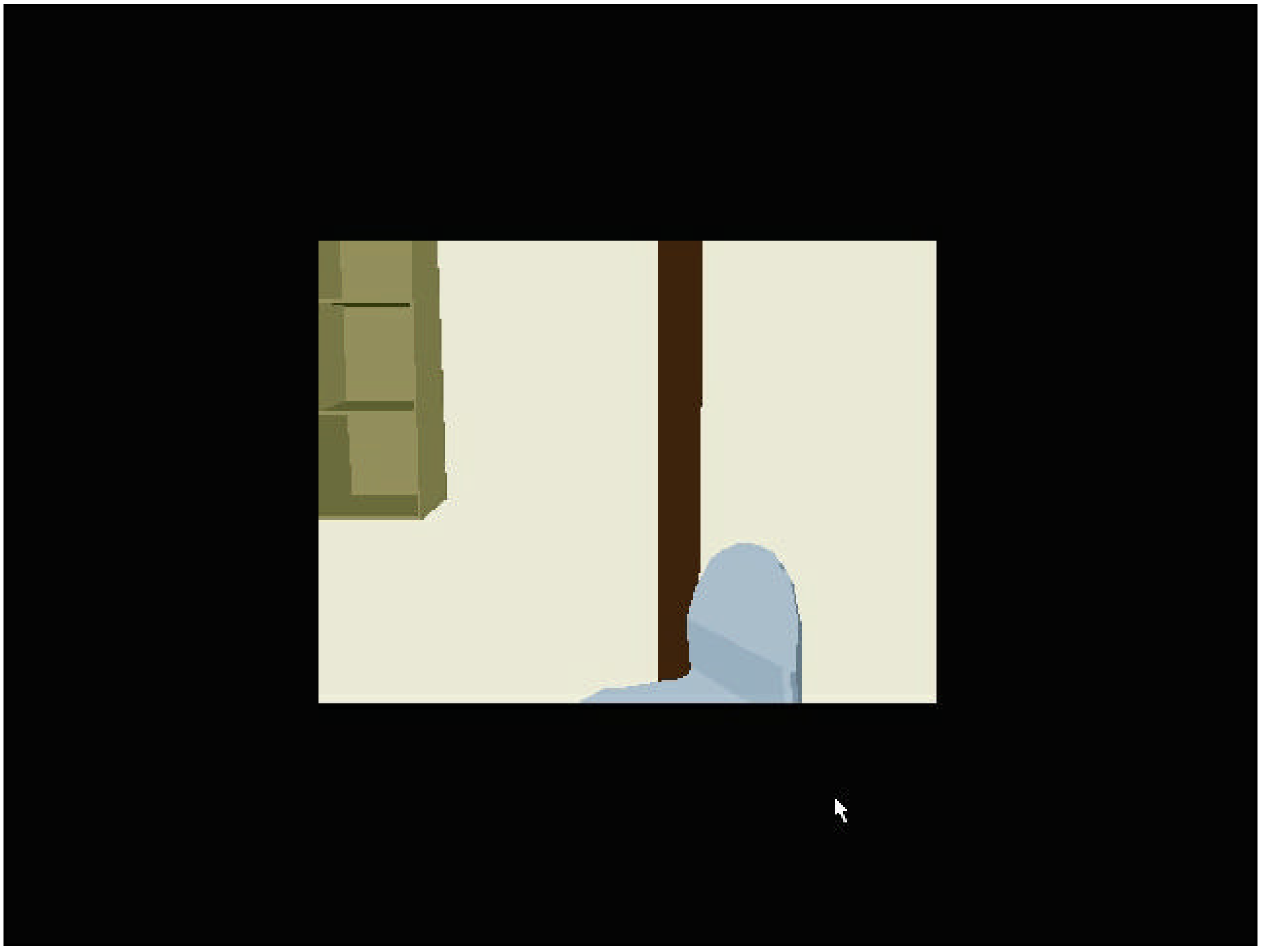


Figure 4. Sample scene from the VE in the low field of view condition.

Table 4. Design of Experiment 1.

Factor Levels										
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
-1	0	-1	1	0	1	0	0	0	0	0
-1	0	0	1	0	1	0	0	0	0	0
-1	0	+1	1	0	1	0	0	0	0	0
-1	1	-1	1	0	1	0	0	0	0	0
-1	1	0	1	0	1	0	0	0	0	0
-1	1	+1	1	0	1	0	0	0	0	0
0	0	-1	1	0	1	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0
0	0	+1	1	0	1	0	0	0	0	0
0	1	-1	1	0	1	0	0	0	0	0
0	1	0	1	0	1	0	0	0	0	0
0	1	+1	1	0	1	0	0	0	0	0
+1	0	-1	1	0	1	0	0	0	0	0
+1	0	0	1	0	1	0	0	0	0	0
+1	0	+1	1	0	1	0	0	0	0	0
+1	1	-1	1	0	1	0	0	0	0	0
+1	1	0	1	0	1	0	0	0	0	0
+1	1	+1	1	0	1	0	0	0	0	0

x_1 = scene update rate: -1 = low, 0 = medium, +1 = high
x_2 = visual display resolution: 0 = low, 1 = high
x_3 = field of view: -1 = low, 0 = medium, +1 = high
x_4 = sound: 0 = absent, 1 = present
x_5 = textures: 0 = absent, 1 = present
x_6 = head-tracking: 0 = absent, 1 = present
x_7 = stereopsis: 0 = absent, 1 = present
x_8 = virtual personal risk: 0 = absent, 1 = present
x_9 = interactions: -1 = low, 0 = medium, +1 = high
x_{10} = second user: 0 = absent, 1 = present
x_{11} = environmental detail: -1 = low, 0 = medium, +1 = high

☐ Indicates data point common to all three experiments

Experiment 2

Experiment 2 combined sound, textures, head-tracking, stereopsis, and virtual personal risk in a half-replicate fractional factorial design. Each of the variables in this study had two levels: present or absent. There were a total of sixteen conditions in the experiment (2^{5-1}). In this design, higher-order interaction effects (i.e., three-way and four-way) were confounded with main effects and two-way interactions. The sacrifice of these higher-order interactions was intentional and acceptable given that the goal of the research was the construction of second-order empirical models. The nonmanipulated factors in this experiment were set at the following levels: scene update rate = low (8 Hz), visual display resolution = high (640 H x 480 V), field of view = high

(48° H x 36° V), interactions = medium, second user = absent, and environmental detail = medium. This experimental design is summarized in Table 5.

Table 5. Design of Experiment 2.

Factor Levels										
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
-1	1	+1	0	0	0	0	0	0	0	0
-1	1	+1	1	1	0	0	0	0	0	0
-1	1	+1	1	0	1	0	0	0	0	0
-1	1	+1	1	0	0	1	0	0	0	0
-1	1	+1	1	0	0	0	1	0	0	0
-1	1	+1	0	1	1	0	0	0	0	0
-1	1	+1	0	1	0	1	0	0	0	0
-1	1	+1	0	1	0	0	1	0	0	0
-1	1	+1	0	0	1	1	0	0	0	0
-1	1	+1	0	0	1	0	1	0	0	0
-1	1	+1	0	0	0	1	1	0	0	0
-1	1	+1	1	1	1	1	0	0	0	0
-1	1	+1	1	1	1	0	1	0	0	0
-1	1	+1	1	1	0	1	1	0	0	0
-1	1	+1	1	0	1	1	1	0	0	0
-1	1	+1	0	1	1	1	1	0	0	0

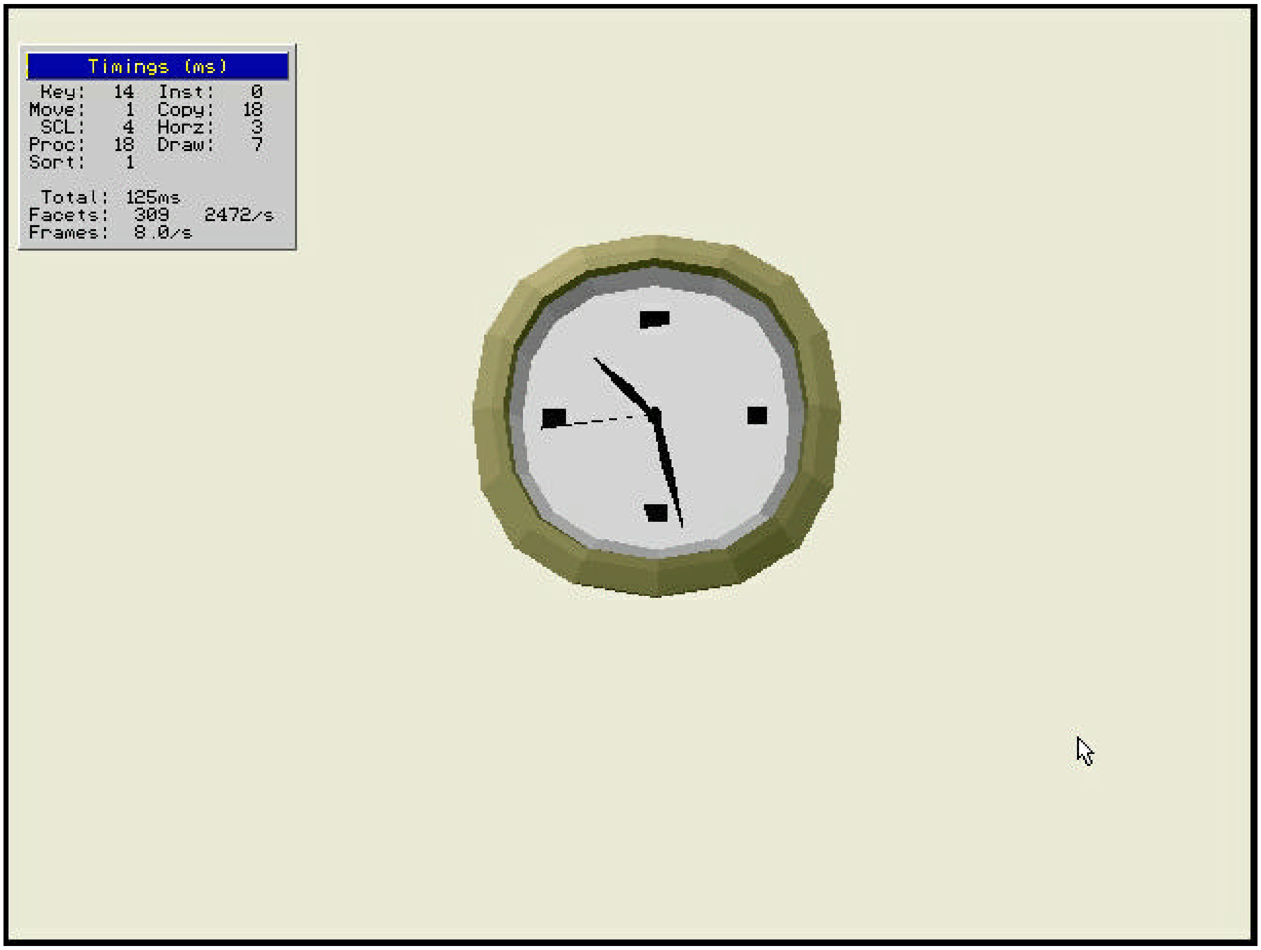
See bottom of Table 4 for identification of factors and levels.

Defining Relationship: $i = x_4 + x_5 + x_6 + x_7 + x_8 = 0 \pmod{2}$

☐ Indicates data point common to all three experiments

Experiment 3

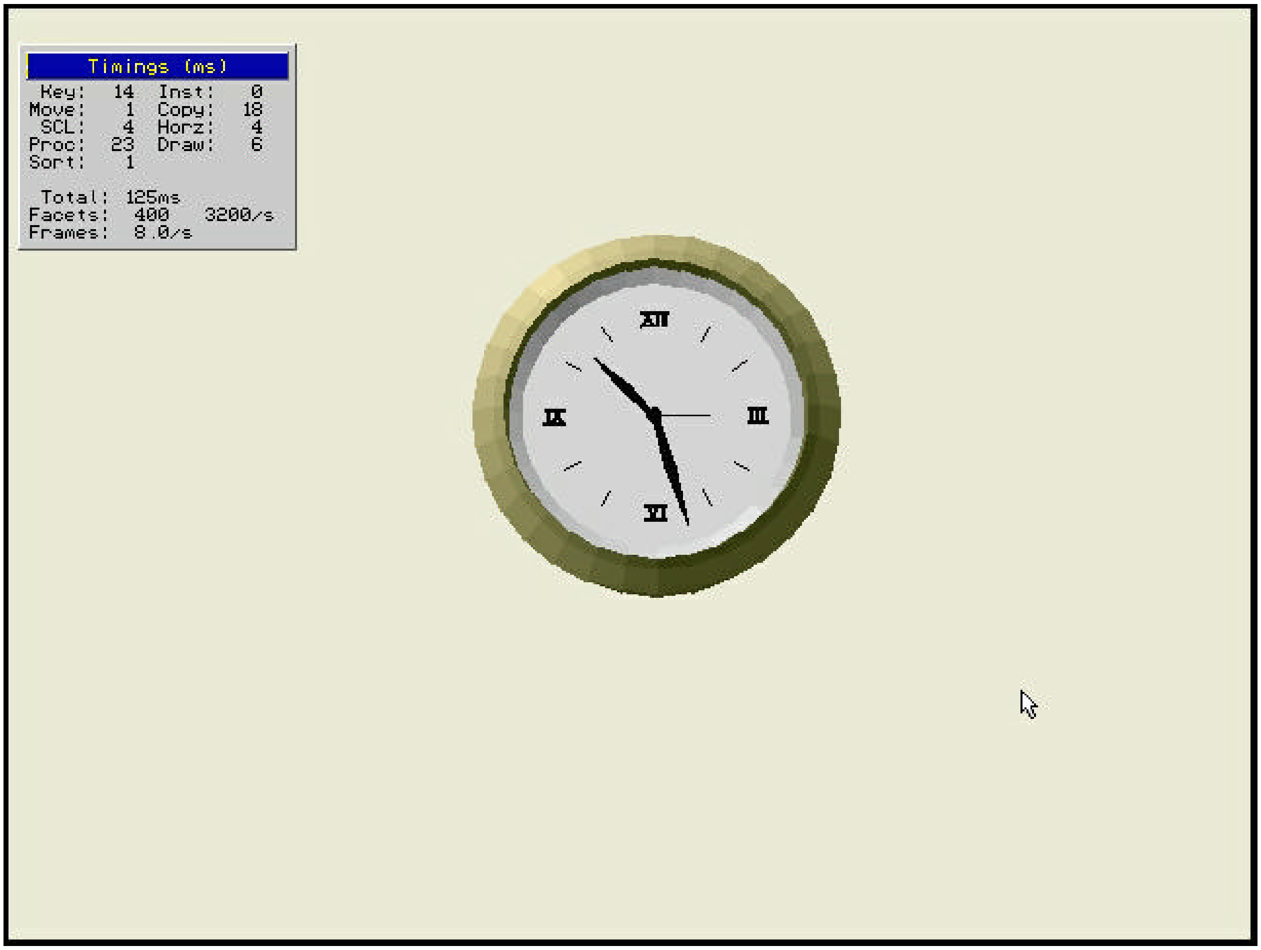
Experiment 3 factorially combined the number of different interactions possible in the VE (6, 12, and 18), presence of a second user in the VE (present or absent), and environmental detail (-7, 0, and +7: the magnitude of the factor applied to the VE software's distancing algorithm). Figures 5 and 6 show a view of a clock in the VE in the low and high environmental detail conditions, respectively. The facet count is shown next to "Facets" in the dialog boxes shown in the upper left of these figures. This experiment had a total of 18 conditions (3^2 x 2). The nonmanipulated factors in this experiment were set at the following levels: scene update rate = low (8 Hz), visual display resolution = high (640 H x 480 V), field of view = high (48° H x 36° V), sound = present, textures = absent, head-tracking = present, stereopsis = absent, and virtual personal risk = absent. This experimental design is summarized in Table 6.



```
     Timings (ms)
 Key:   14  Inst:   0
Move:    1  Copy:  18
 SCL:    4  Horz:   3
Proc:   18  Draw:   7
Sort:    1

 Total:  125ms
Facets:   309   2472/s
Frames:   8.0/s
```

Figure 5. View of a clock in the low environmental detail condition.



```
      Timings (ms)
 Key:    14  Inst:    0
Move:     1  Copy:   18
 SCL:     4  Horz:    4
Proc:    23  Draw:    6
Sort:     1

 Total:  125ms
Facets:   400   3200/s
Frames:   8.0/s
```

Figure 6. View of a clock in the high environmental detail condition.

Table 6. Design of Experiment 3.

Factor Levels										
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
-1	1	+1	1	0	1	0	0	-1	0	-1
-1	1	+1	1	0	1	0	0	-1	0	0
-1	1	+1	1	0	1	0	0	-1	0	+1
-1	1	+1	1	0	1	0	0	-1	1	-1
-1	1	+1	1	0	1	0	0	-1	1	0
-1	1	+1	1	0	1	0	0	-1	1	+1
-1	1	+1	1	0	1	0	0	0	0	-1
-1	1	+1	1	0	1	0	0	0	0	0
-1	1	+1	1	0	1	0	0	0	0	+1
-1	1	+1	1	0	1	0	0	0	1	-1
-1	1	+1	1	0	1	0	0	0	1	0
-1	1	+1	1	0	1	0	0	0	1	+1
-1	1	+1	1	0	1	0	0	+1	0	-1
-1	1	+1	1	0	1	0	0	+1	0	0
-1	1	+1	1	0	1	0	0	+1	0	+1
-1	1	+1	1	0	1	0	0	+1	1	-1
-1	1	+1	1	0	1	0	0	+1	1	0
-1	1	+1	1	0	1	0	0	+1	1	+1

See bottom of Table 4 for identification of factors and levels.

☐ Indicates data point common to all three experiments

Subjects

Each experiment employed twelve subjects. Subjects were not allowed to participate in more than one experiment. Subjects were screened for normal visual acuity, normal color vision, and normal stereo acuity. Advertisements for subjects were placed on Virginia Tech newsgroups and throughout the Virginia Tech campus. Subjects ranged in age from sixteen to forty-two with an average age of twenty-two. Eight of the thirty-six subjects were female. Subjects were paid $5 an hour for their participation.

Task

The subtasks chosen from those proposed by Lampton, Knerr, et al. (1994) are shown in Table 7.

Table 7. Experimental subtasks.

Task Category	Task Name	Task Description
Vision	Distance Estimation	Indicate when the image of a human figure, moving toward the viewer from an initial distance of 40 ft. is 20, 10, 5, and 2.5 ft. away.
Manipulation	Bins	Move a ball located in a vertical rack of open containers (bins), pull it out of the original bin, and push it into a target bin.
Locomotion	Turns	Move through a series of corridors requiring alternating left and right turns.
Vision	Search	Detect a target moving about the walls, floor, or ceiling of a room.
Reaction Time	Choice	Indicate in which of four boxes an "X" has appeared.

These subtasks, performed in the order shown above, constituted the experimental task (i.e., a trial) for each treatment condition in each experiment. The differences between the tasks used in the present research and those described by Lampton, Knerr, et al. (1994) were minor and derived from suggestions for changes made by these authors based upon their results or from a desire to limit the duration of immersion in the VE (and the length of experimental sessions). For the Turns and Choice tasks, the dependent variables consisted of time to perform the task and error rate (errors were counted as the number of contacts with walls in the Turns task). For the Distance Estimation task the dependent variables were the actual distances at each of four estimated distances. For the Search task and the Bins task the dependent variable was response time.

Environment

The VE used in all experiments had floors with a checkerboard pattern, walls eight feet high with narrow vertical stripes every five feet, ceilings with horizontal light panels every ten feet, and corridors three feet wide. All measurements specified are, of course, virtual. Figure 7 shows a cut-away, overhead of view of that part of the VE containing the first three tasks. Figure 8 shows a cut-away, overhead of view of that part of the VE containing the last two tasks. Each subject's viewpoint was attached to an invisible “body” that measured (virtually) 16.5” front-to-back and side-to-side.

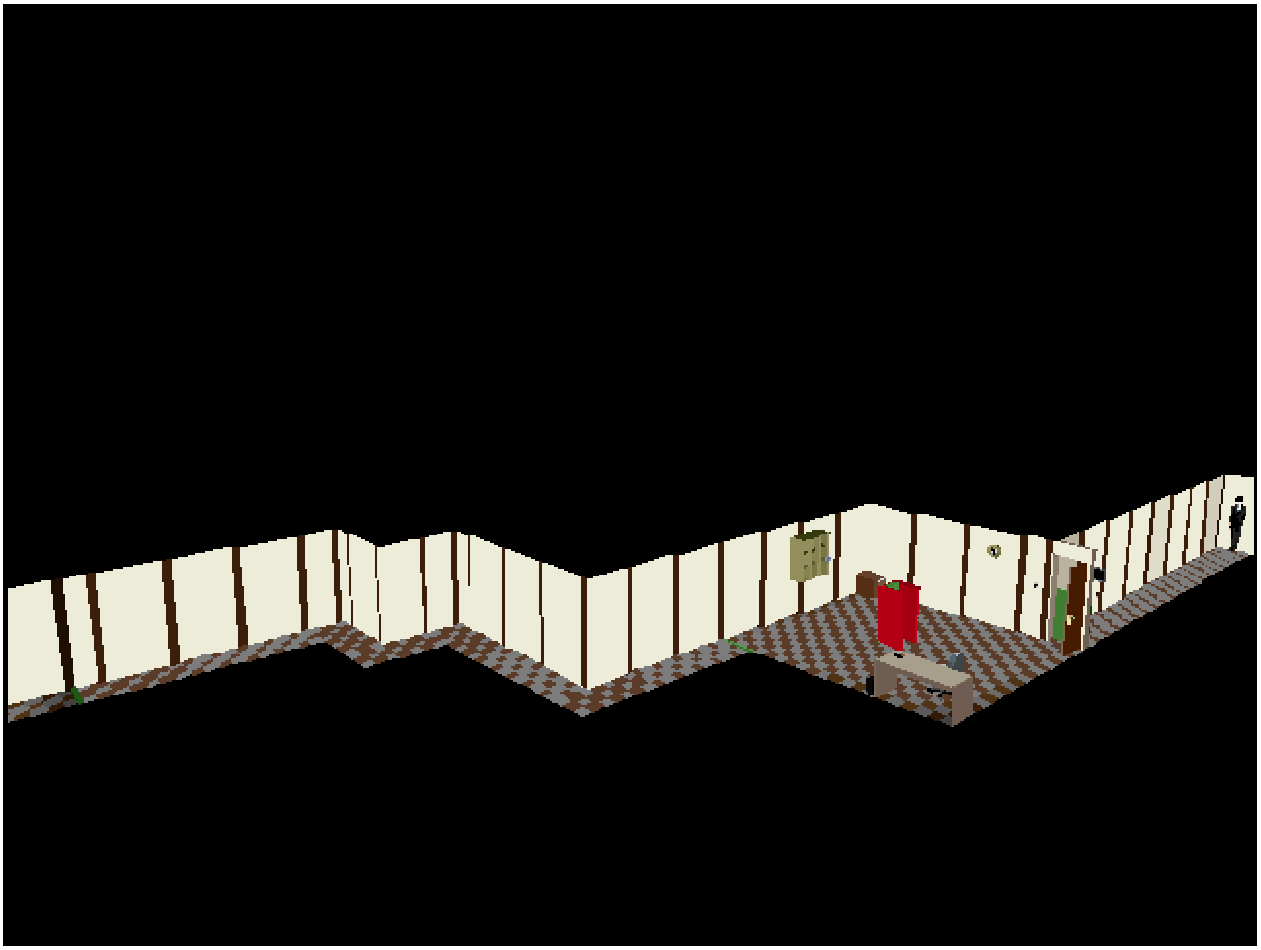


Figure 7. Overhead, cut-away view of VE containing the first three tasks.

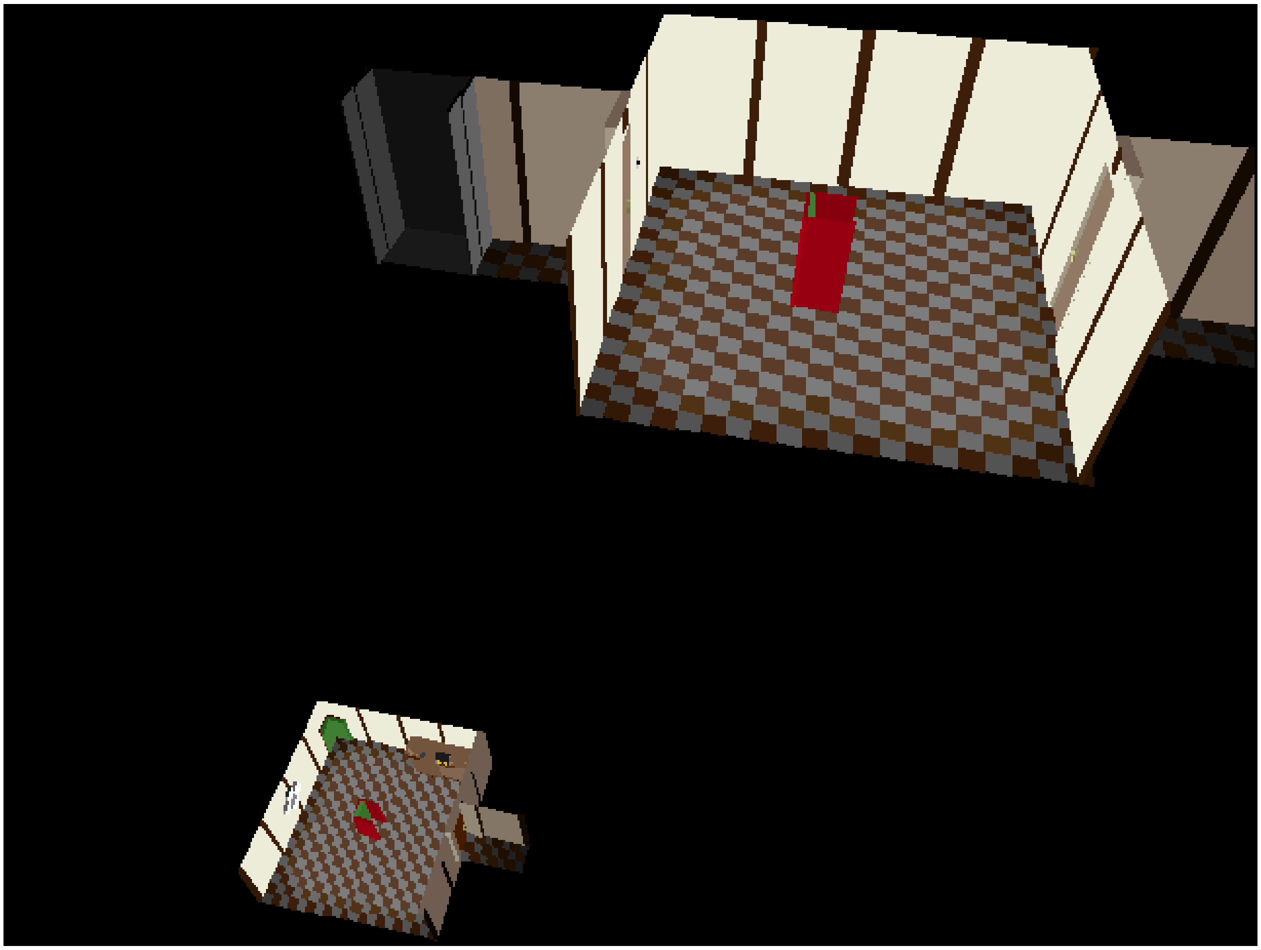


Figure 8. Overhead, cut-away view of VE containing the last two tasks.

The figure of 16.5” is the same as that used by Lampton, Knerr, et al. (1994) and is based on the 50th percentile value for elbow-to-elbow breadth for adult males in the U.S. population (Sanders and McCormick, 1987). Viewpoint and eye level were set at the subject's actual standing eye level as measured at the start of the experiment. This virtual body interacted with the environment by colliding with walls, doorframes, and movable objects. Controlling the virtual body through head-tracking and manipulation of a 3D control device, subjects were able to adjust their viewpoint with four degrees of freedom: Z-axis rotation (i.e., roll) and Y-axis translation (i.e., elevation) were turned off. Speed of rotation and translation (i.e., sensitivity of the software to input from the 3D control device and the position sensor) for all axes was set via software to roughly correspond to that normally associated with walking. During immersion subjects stood at a swiveling platform adjusted to a height comfortable to them (approximately standing elbow height) at the start of the experiment. Their left hands rested on the 3D controller while a standard 2D mouse was held in their right hands. Subjects used the mouse to control a standard arrow cursor and left-clicked with the mouse to interact with objects beneath this cursor.

In the sound = present condition, subjects were given auditory feedback when they bumped into objects and when they clicked on interactable objects. Also, various objects in the VE made context-appropriate sounds (e.g., wall-clocks ticked). In the textures = present condition, the bins, doors, walls, and other objects within the environment were given context-appropriate textures (e.g., wood grain for the doors) to the extent that preservation of maximum scene update rate allowed this (see Figures 9 and 10). In the head-tracking = present condition, subjects were able to control their viewpoint rotationally by moving and turning their heads. In the head-tracking = absent condition, viewpoint control was possible via manipulation of the 3D controller only. In the virtual personal risk = present condition, the rear doors of an elevator that carried subjects from the room containing the Search task to the room containing the Choice task were absent. In this condition these doors were replaced by a yellow and black striped warning bar and subjects could view the room and hallway below as they descended the five (virtual) stories down to the next task (see Figures 11 and 12).

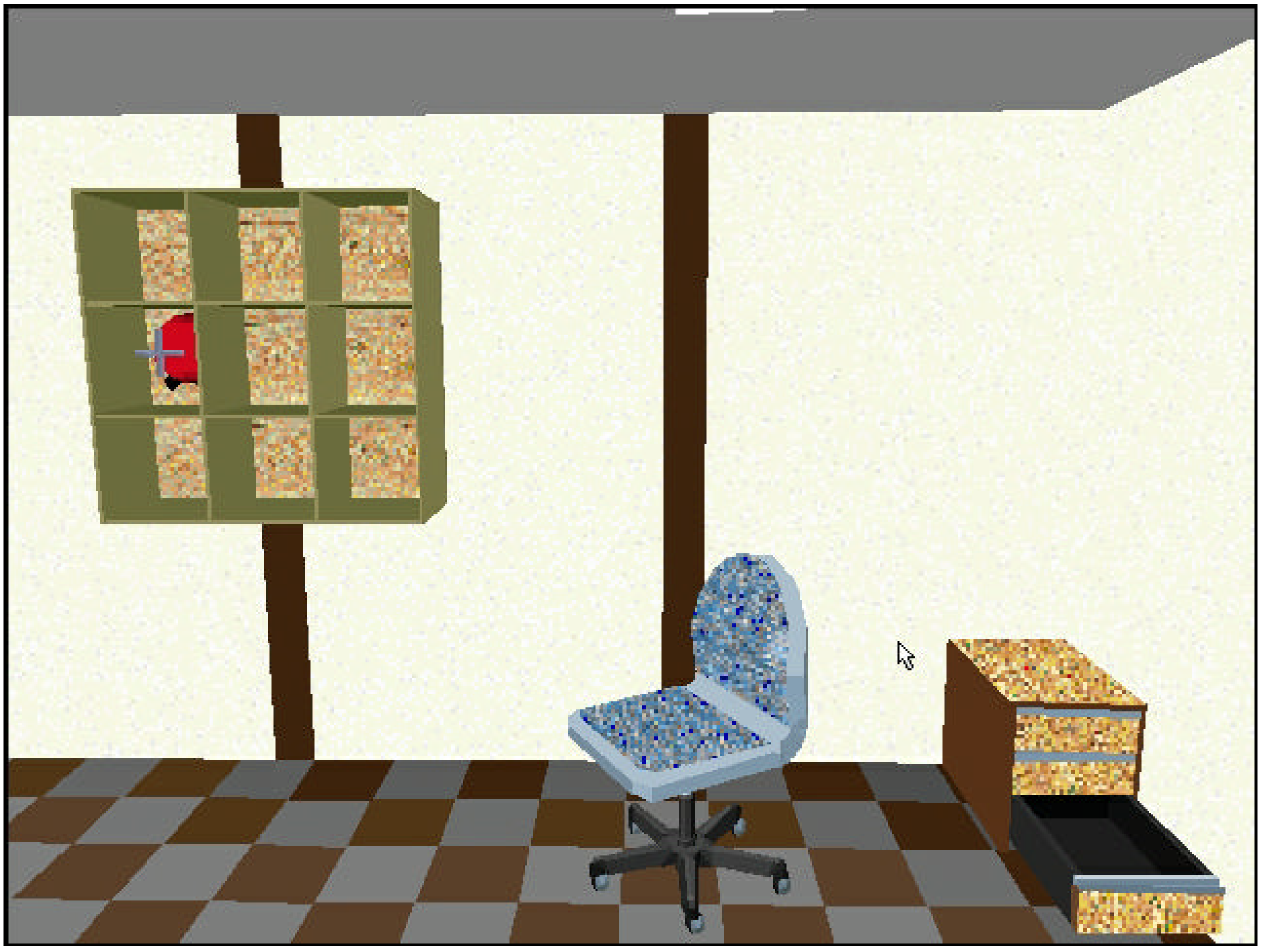


Figure 9. Scene from VE with textures on.

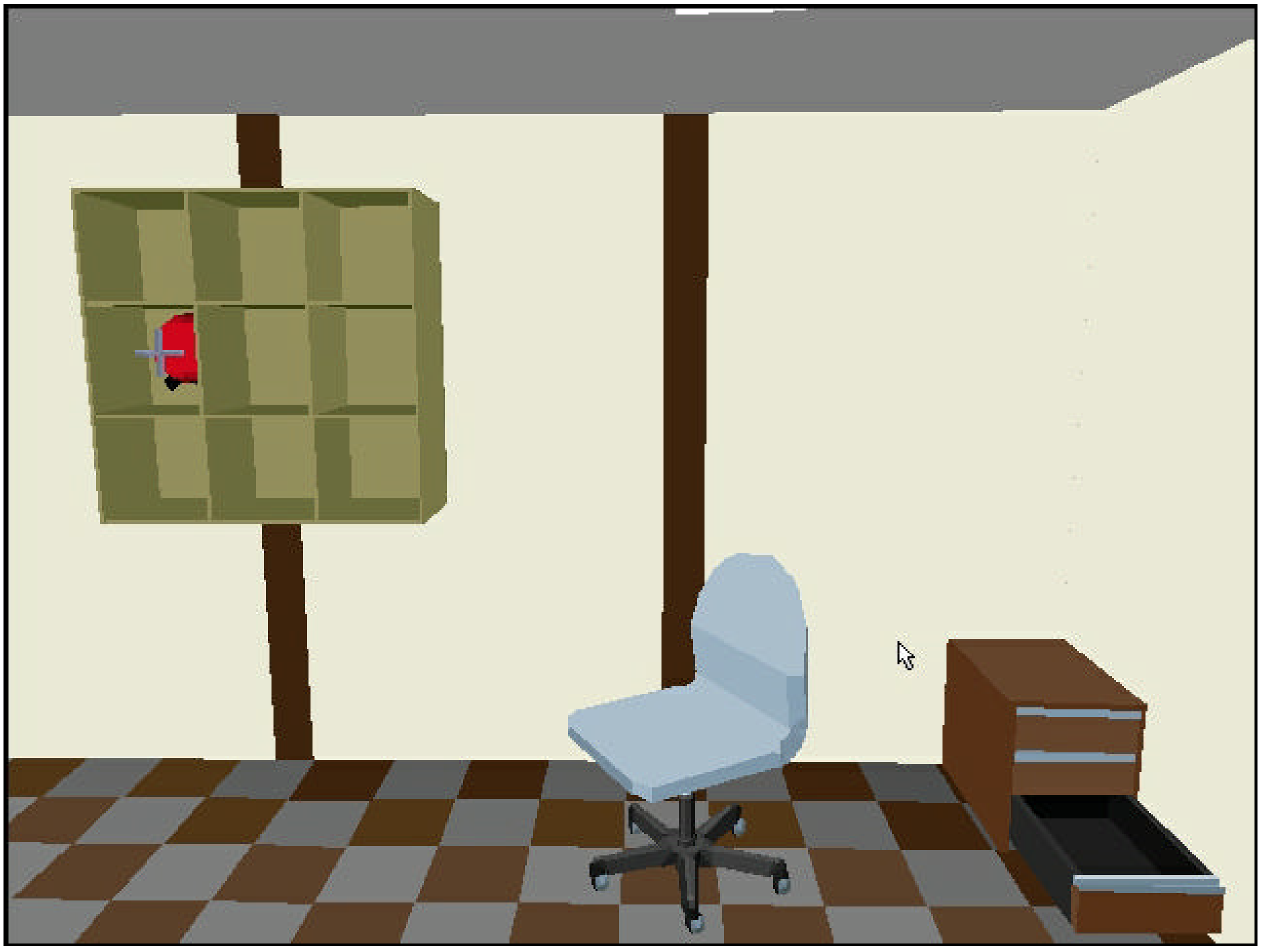


Figure 10. Same scene with textures off.

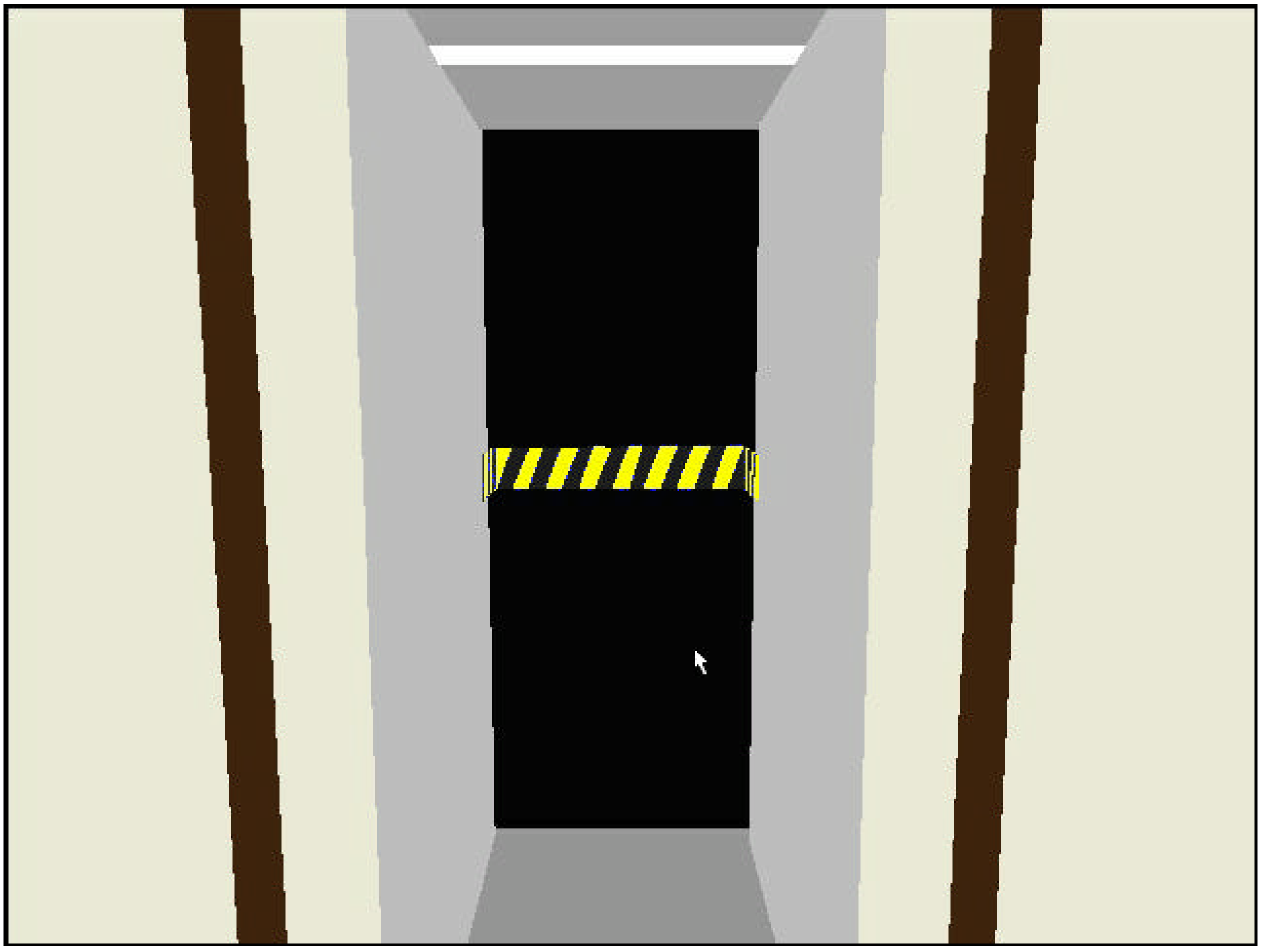


Figure 11. View moving into elevator with virtual personal risk present.

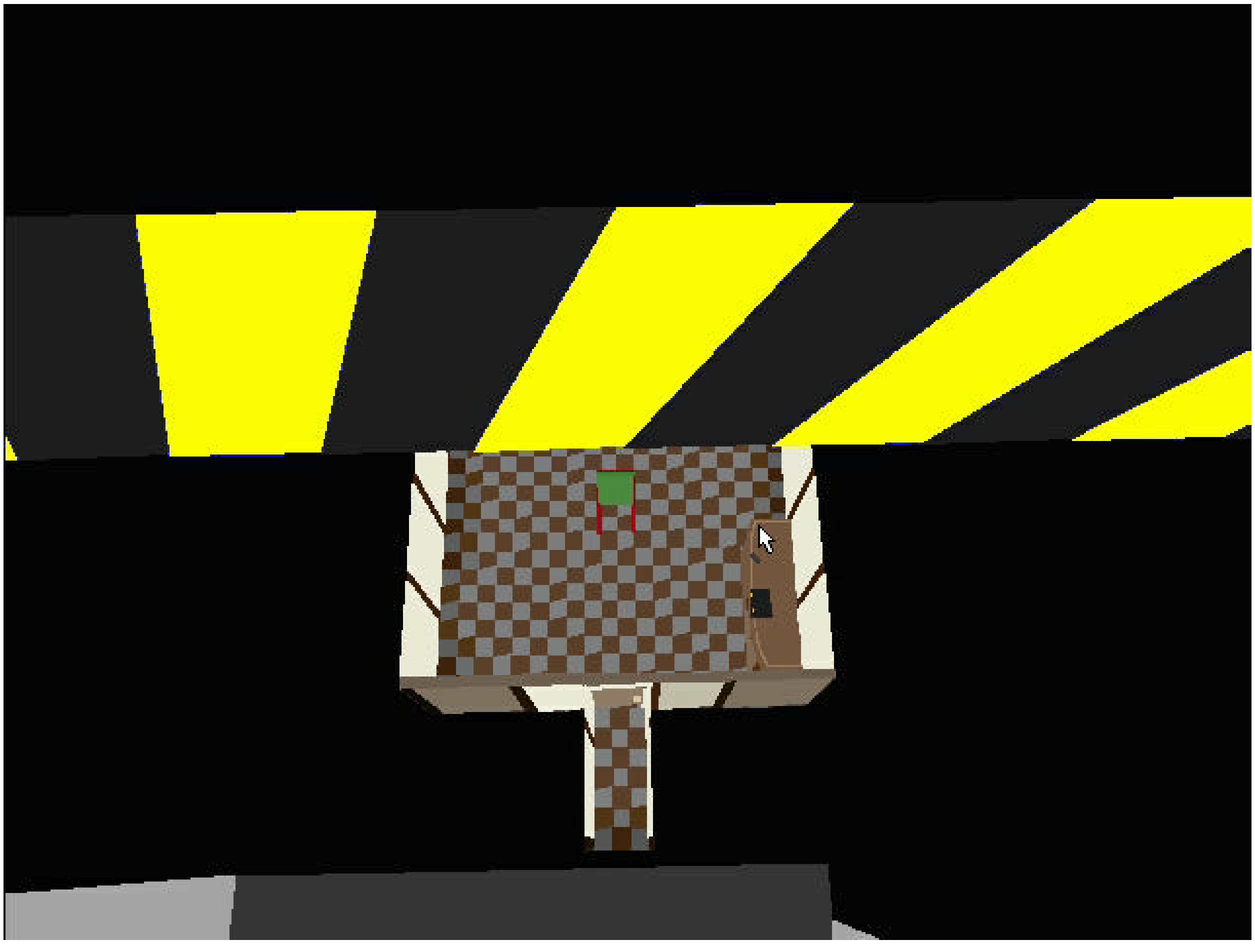


Figure 12. View from elevator during descent.

While the number of objects in the environment remained constant in all conditions, the extent of interaction with these objects varied. For example in the high interactions condition, subjects were able to open and close doors and drawers, click on light switches to turn lights on and off, open and close a briefcase, etc. (see Table 8). In the low interactions condition, the drawers, light switches, briefcase, etc. did respond to attempts to activate them. Subjects were told with which objects interaction was possible during a demonstration trial. Subjects were not allowed to interact extraneously with the VE during timed tasks and were informed of this at the start of the experiment. In the second user = present condition, the experimenter controlled and was represented by the human figure in the distance estimation task. At the completion of this task the experimenter virtually accompanied the subject through the rest of the trial, interacting with the subject as if both were present in the VE.

Table 8. Possible interactions and interactions conditions.

Possible Interaction	Interaction Condition		
	Low	Medium	High
Button press in distance estimation task	Active	Active	Active
3D cursor manipulation	Active	Active	Active
Elevator door	Active	Active	Active
Start blocks	Active	Active	Active
Ball stop in search task	Active	Active	Active
Reaction time response	Active	Active	Active
Chair swivel	Inactive	Active	Active
Cabinet doors	Inactive	Active	Active
Light switches	Inactive	Active	Active
Wood doors	Inactive	Active	Active
Drawer open and close	Inactive	Active	Active
Chair push	Inactive	Active	Active
Calculator	Inactive	Inactive	Active
Briefcase lock tumblers	Inactive	Inactive	Active
Briefcase latch	Inactive	Inactive	Active
Briefcase handle	Inactive	Inactive	Active
Stapler	Inactive	Inactive	Active
Briefcase lid	Inactive	Inactive	Active

Distance Estimation

In the distance estimation task, a human figure appeared at the end of a forty-foot corridor. Subjects were informed in the written instructions for this task and during the demonstration trial that the figure was six feet tall. The figure began to move toward the subject and the subject was asked to click on a button on a response panel in the VE when the figure appeared to be 20, 10, 5, and 2.5 ("arm's length") feet away. An example of a subject’s view of this task is shown in Figure 13.

Figure 13. Example of subject's view during distance estimation task.

Bins

In this task, the participant faced a three by three stack of open-ended, box-like compartments (bins) much like Tic-Tac-Toe squares. At the beginning of the task, inputs to the 3D controller transitioned from control of the subject's viewpoint to control of a 3D cross. A red ball appeared in one of the bins (randomly determined, but not the middle bin) and an X appeared in another bin two bins away (to make the required movement distance constant). The 3D cross was moved until it contacted the ball (at which point auditory feedback was given and the cross changed from blue to yellow). Upon contact with the ball, the cross locked onto the ball and was then used to move the ball to the bin containing the X. The task ended when the ball contacted the X at the back of the bin (again accompanied by auditory and visual feedback). An example of a subject's view of this task is shown in Figure 14.

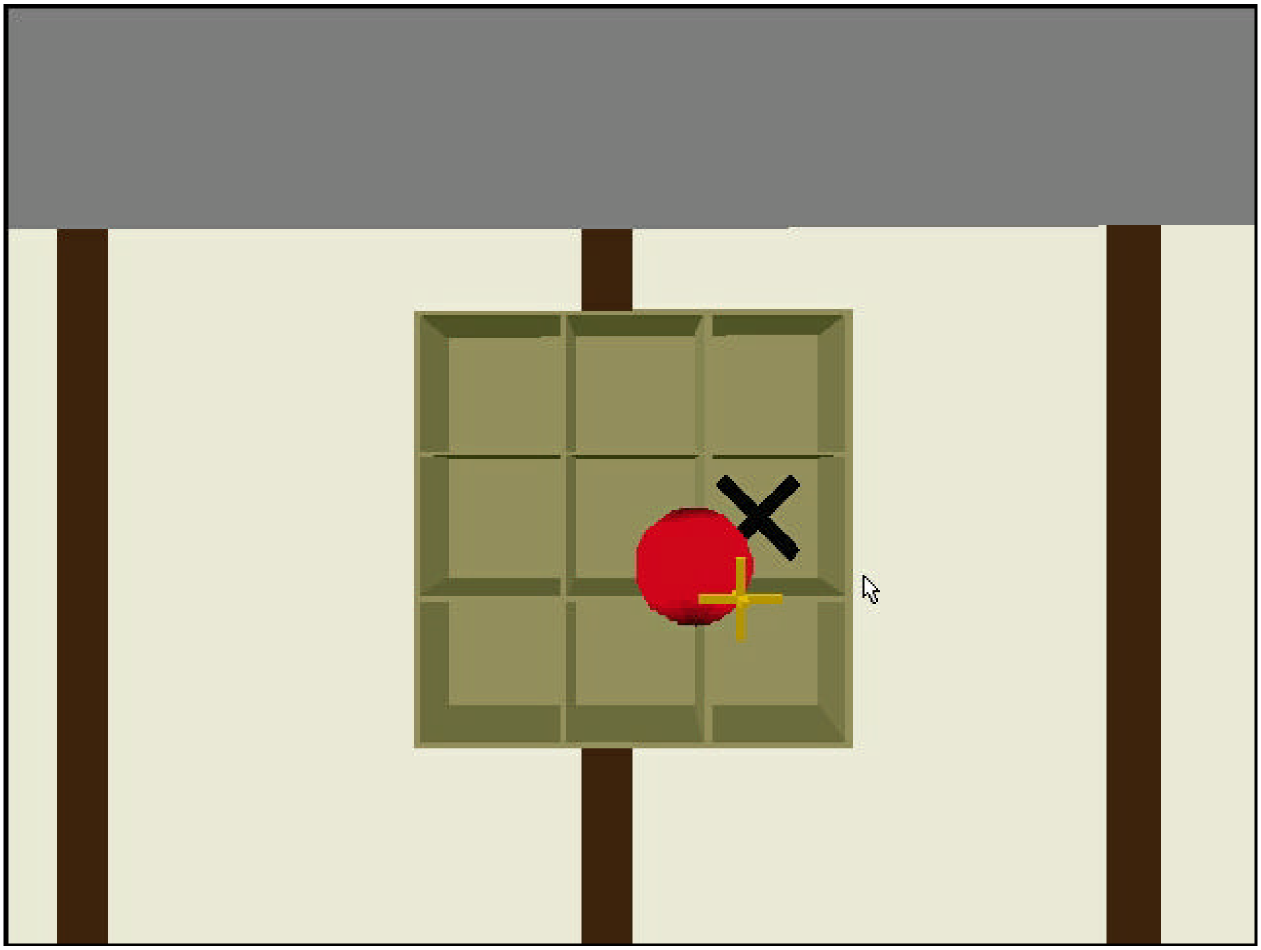


Figure 14. Example of subject's view during the bins task.

Turns

Subjects were required to navigate through a continuous narrow corridor formed by straight-aways joining alternate 90˚ turns to the left and right. There were six turns in the corridor. The lengths of the straight-aways varied, with two 20-ft. segments alternating with two 10-ft. segments. An example of a subject's view of the start of this task is shown in Figure 15.

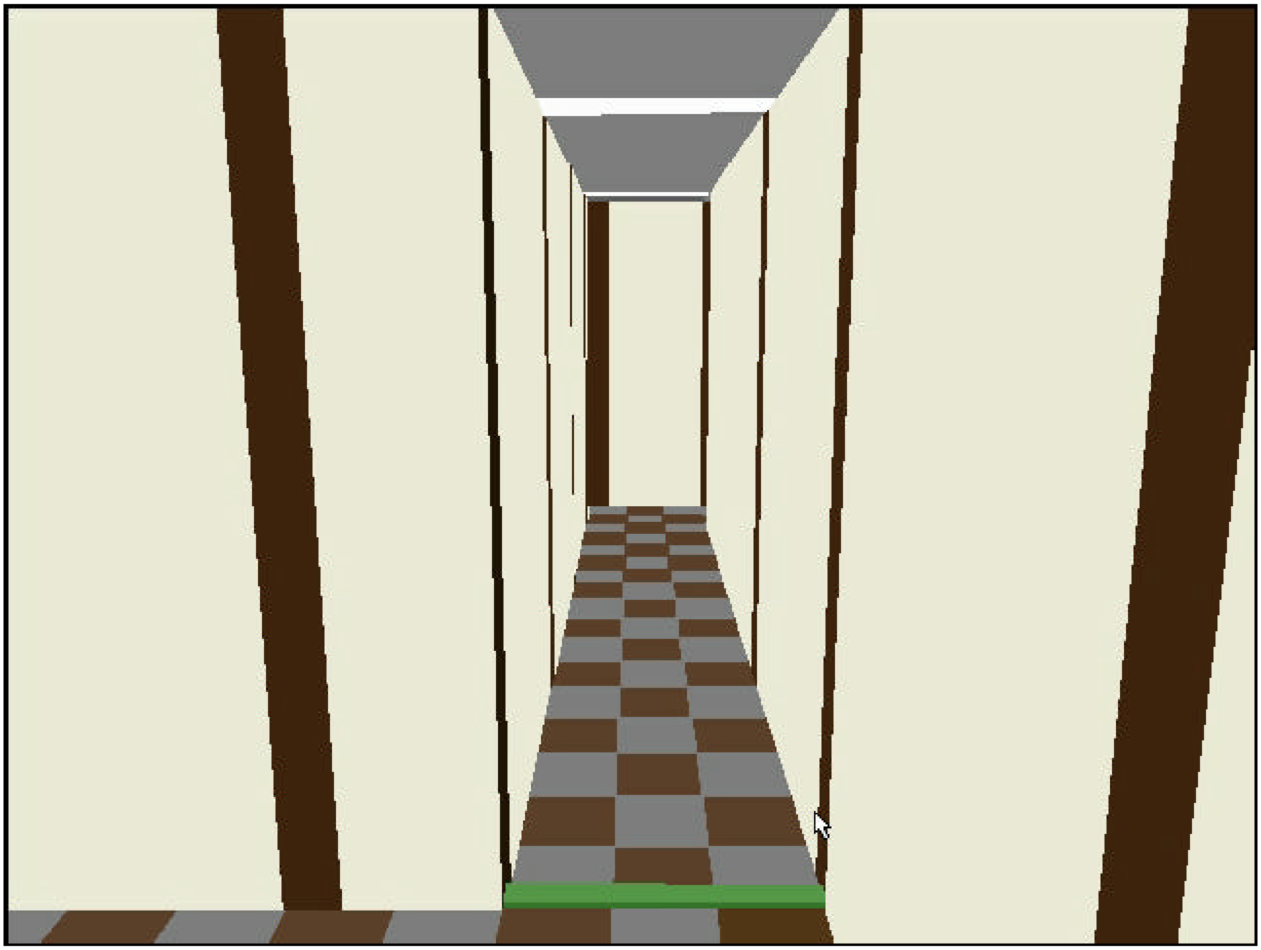


Figure 15. Example of subject’s view at the beginning of the turns task.

Search

Subjects went (virtually) to the center of a room. At the start of the task, a red ball appeared near the floor, ceiling, or walls, and moved slowly (1 ft. per second) around the room. Subjects searched for the ball by making head and body movements or by controlling their viewpoint using the 3D controller (depending on head-tracking condition). The task ended (accompanied by auditory feedback) when the subject clicked on the ball with the mouse. The starting location of the ball was determined randomly with the restriction that it could not appear in the subject's initial field of view. An example of a subject’s view of this task is shown in Figure 16.

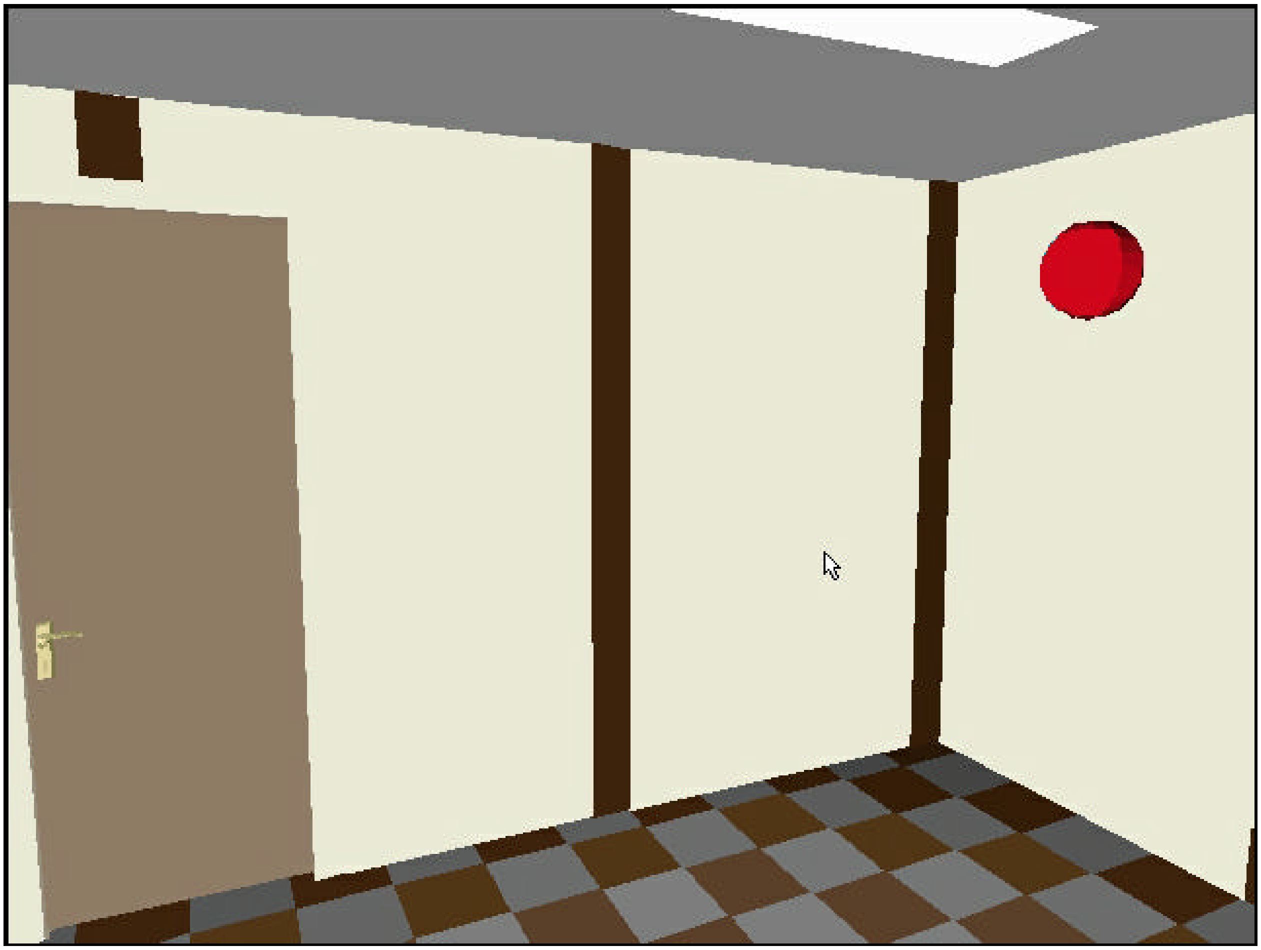


Figure 16. Example of subject’s view during search task.

Choice

An X appeared on one of four blocks (which block was randomly determined) and subjects were asked to click with the mouse on the X or the block on which it appeared as soon as they could. An example of a subject’s view of this task is shown in Figure 17.

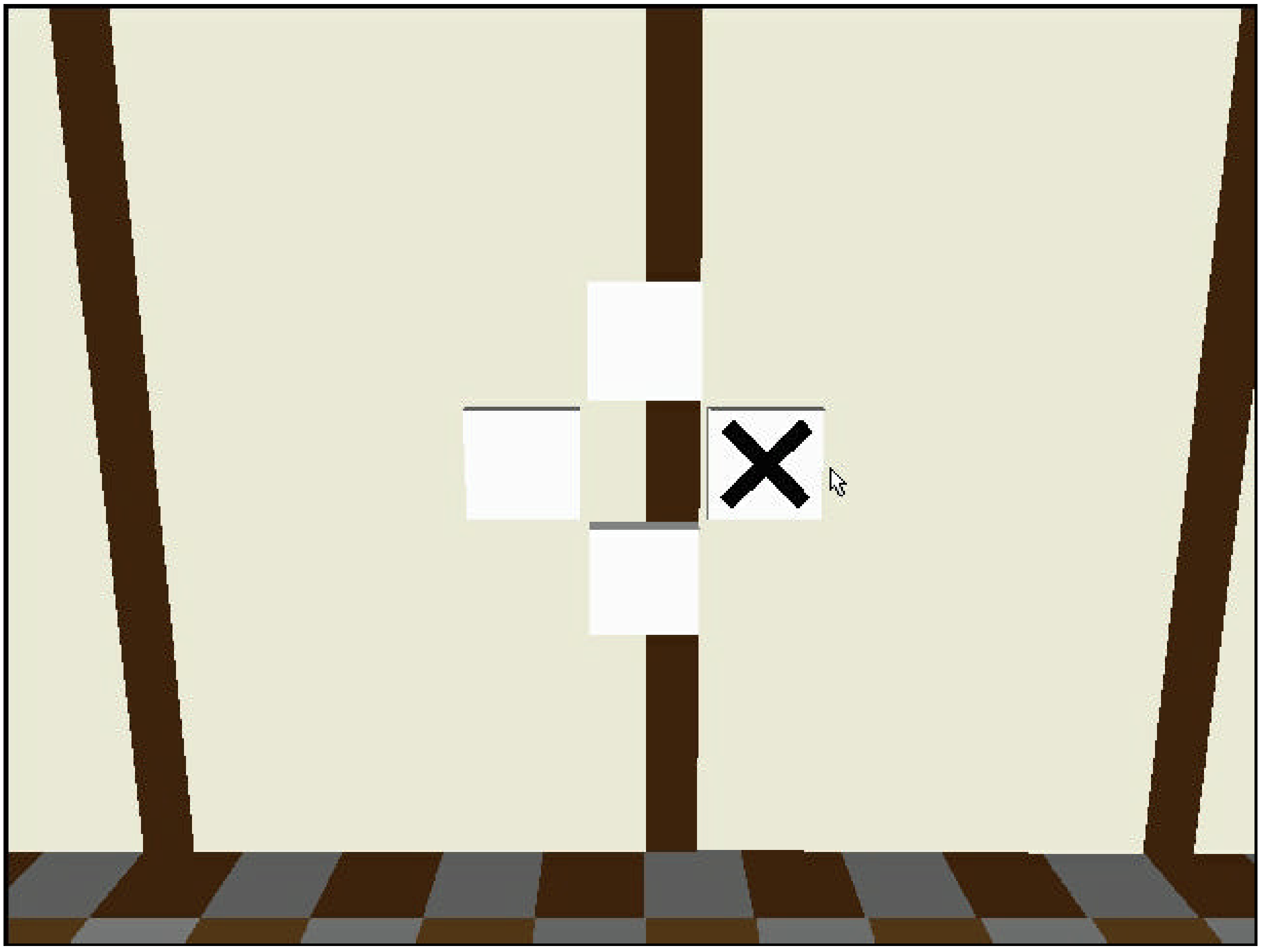


Figure 17. Example of subject’s view during choice reaction time task.

Equipment

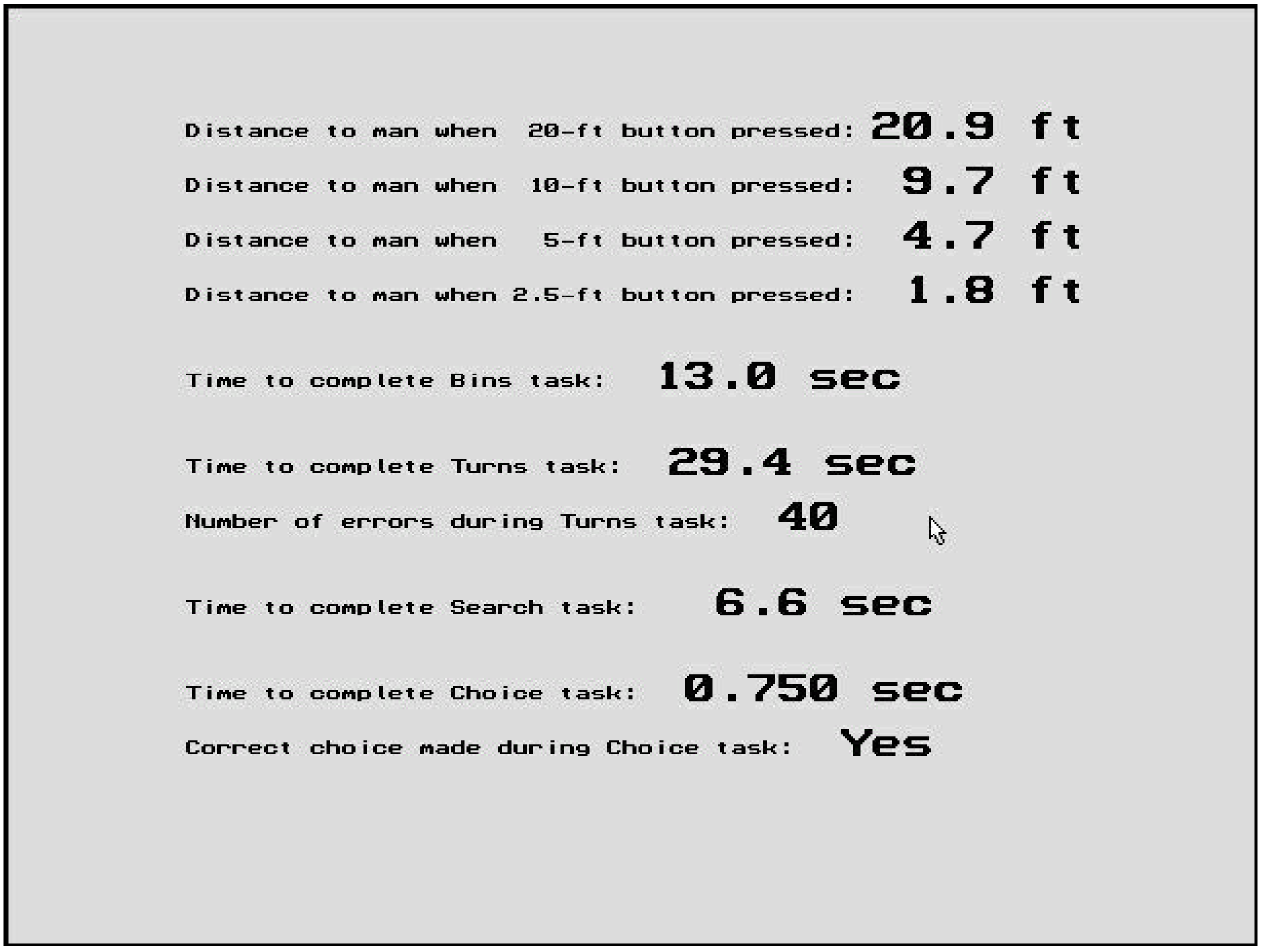
Two Pentium-based personal computers were used to run the software that interfaced the VR system peripherals and generated the VE. Head-tracking was done using Ascension Technologies' Flock of Birds, a DC-pulsed magnetic tracking system. Visual and auditory presentation of the VE was done via a VR4 head-mounted display manufactured by Virtual Research Systems. Navigation through the environment and object manipulation was accomplished through the use of a Logitech Magellan — a 6-DOF, stationary control device. A standard Microsoft, 2-DOF mouse was used to interact with objects in the VE. The Magellan and the Microsoft mouse (and the participant’s hands) were supported by a rotating platform. This system is shown in Figure 18.

Figure 18. Experimental apparatus: VR4 worn on the head, magnetic tracking receiver mounted on top of VR4, Magellan held in the left hand, Microsoft mouse held in the right hand, and rotating platform supporting hand-held input devices.

The software used to generate the VE and control each experiment was Superscape VRT produced by Superscape VR plc.

Procedure

Each experimental session began with the subject reading an introduction to the study (provided in **APPENDIX A**). After this, they were screened for normal visual acuity, color perception, and stereoscopic perception, and their standing eye height was measured so that their virtual eye height could be set accordingly. They then read and sign an informed consent form (provided in **APPENDIX B**) and read the experimental instructions (provided in **APPENDIX C**). Subjects became familiar with the concept and process of magnitude estimation by estimating the lengths of twenty lines presented in random order on twenty sheets of paper. This practice session had a number of purposes in addition to familiarizing subjects with magnitude estimation: these included stabilizing the results of subsequent scaling (Zwislocki, 1983), reducing the regression effect (subjects' tendency to avoid assigning extreme values and regress toward their individual mean assigned values (Zwislocki, 1983)), and obtaining a baseline psychophysical scale for each subject that could later be compared with previous research (e.g., Collins and Gescheider, 1989) and with each individual's scale of presence (Gescheider, 1985). The instructions to subjects for performing magnitude estimations were based on prototypical instructions provided by Stevens (1971) and Gescheider (1985). Next, subjects were shown how to don the HMD and given a demonstration trial in which the experimenter went through the VE from beginning to end, demonstrating each task in sequence and verbally pointing out objects with which interaction was possible. Each subject then went through one practice trial. Subjects were told that they could ask any questions they wished to during the demonstration and practice trials, but that the experimenter would attempt to remain silent during the data trials to follow. The experimental condition of the demonstration and data trials was the same as that common to all three experiments: scene update rate = 8 Hz, visual display resolution = high, field of view = high, sound = present, textures = absent, head-tracking = present, stereopsis = absent, virtual personal risk = absent, interactions = medium, second user = absent, and environmental detail = medium. Following this practice trial, subjects were given a short (3-5 minutes) break during which they could take off the HMD, sit down, get a drink of water, etc. Subjects were given similar breaks after every fourth data trial in all three experiments. Trials averaged roughly seven minutes in length. Subjects were given feedback at the end of each trial concerning their performance on each task (see Figure 19 for an example). Subjects underwent a number of trials equivalent to the number of conditions in the experiment, one from each condition. The order of presentation of these conditions was determined randomly. The length of each experimental session varied from subject to subject and between experiments. The time spent in the HMD averaged approximately 2.5 hours. Subjects were closely monitored throughout each session for adverse reactions to the VR system and verbally reminded by the experimenter of their freedom to withdraw if they demonstrated or reported any adverse side effects of immersion. Of the 42 subjects who passed the vision screening, six did not complete the experimental session. Five of these ceased participation because of motion-sickness-like symptoms, invariably within the first 20 minutes after donning the HMD. The other subject stopped participation after the practice trial, complaining that the HMD could not be adjusted so that it was physically comfortable.



Distance to man when 20-ft button pressed: 20.9 ft

Distance to man when 10-ft button pressed: 9.7 ft

Distance to man when 5-ft button pressed: 4.7 ft

Distance to man when 2.5-ft button pressed: 1.8 ft

Time to complete Bins task: 13.0 sec

Time to complete Turns task: 29.4 sec

Number of errors during Turns task: 40

Time to complete Search task: 6.6 sec

Time to complete Choice task: 0.750 sec

Correct choice made during Choice task: Yes

Figure 19. Example of feedback given to subjects at the end of each trial.

RESULTS

Magnitude Estimation

It has been shown that the distribution of estimates people give in magnitude estimation experiments is usually positively skewed. These distributions are typically lognormal, with error increasing as stimulus magnitude increases (see Mansfield, 1974; Stevens, 1971; Stevens, 1975), and so the geometric mean is taken rather than the arithmetic mean, since this produces an unbiased estimate of the expected value of a lognormal variate (Cross, 1974). This is fairly standard procedure in the analysis of magnitude estimation data, but it begs the question of what to do when one wishes to do something other than plot a regression line through measures of central tendency at different stimulus values. A number of authors (see Engen, 1972; Lane, Catania, and Stevens, 1961; Morrison, 1995; Stevens, 1971) have proposed and used a data reduction process for free-modulus magnitude estimation data designed to, 1) adjust the data to a common modulus (also reducing variability in the data arising purely from subjects' selection of different moduli and number ranges), and, 2) provide data suitable for analysis using parametric statistics (e.g., normally distributed). This procedure, called modulus equalization by Stevens (1971), is described as follows:

1. Convert all responses to their logarithmic values. This converts the distribution of responses from lognormal to normal.

2. Calculate the mean of each individual's responses. This yields each individual's modulus.

3. Calculate the mean of all responses. This yields the common modulus.

4. Subtract the result of Step 3 from the result of Step 2 for each individual. This yields a set of constants representing the offsets of each individual modulus from the common modulus.

5. Subtract the constant obtained in Step 4 for each individual from that individual's scores (obtained in Step 1). This yields a dataset that is normally distributed and adjusted to a single, common modulus.

6. Take the antilog of all values obtained in Step 5. This yields a ratio-scale dataset (i.e., one with an origin of zero in which the ratios of scale values are meaningful) in the same units as individuals' original responses (albeit still units of subjective magnitude).

An example of this procedure, taken from Experiment 1, is shown in Tables 9-12. Table 9 shows the raw presence estimates of all subjects in the experiment. Table 10 shows the logarithmic values of these estimates (Step 1 above), the individuals' moduli (Step 2 above), the common modulus (Step 3 above), and individuals' offsets from the common modulus (Step 5 above). Table 11 shows the resulting data set with all scores adjusted to the common modulus. Table 12 shows the final dataset after taking the antilog of the values in Table 11.

Table 9. Raw presence estimates from Experiment 1.

	Subject											
Trial	1	2	3	4	5	6	7	8	9	10	11	12
1	3.00	3.00	11.00	8.30	2.75	0.25	7.00	1.75	0.50	10.00	7.00	3.00
2	5.00	5.00	11.50	8.50	3.50	3.00	10.00	5.00	4.00	25.00	4.00	5.00
3	20.00	7.00	11.50	9.00	6.75	4.00	10.00	3.00	4.00	20.00	7.00	8.00
4	3.00	5.00	7.00	5.00	3.70	1.00	12.00	4.00	0.66	18.00	3.00	2.00
5	8.00	7.00	10.50	9.00	4.50	5.00	12.00	2.00	7.00	15.00	9.00	7.00
6	15.00	8.00	15.00	9.00	6.80	7.00	35.00	4.00	3.00	30.00	7.00	9.00
7	3.00	4.00	8.00	8.00	5.00	0.25	5.00	2.00	0.75	10.00	2.00	3.00
8	6.00	5.00	13.50	8.00	2.75	4.00	25.00	4.00	3.00	15.00	4.00	7.00
9	17.00	10.00	10.00	9.00	5.70	4.00	20.00	1.50	4.00	25.00	8.00	7.00
10	3.00	1.00	9.50	7.00	5.00	0.50	8.00	3.00	0.66	25.00	7.00	4.00
11	7.00	7.00	11.00	8.50	6.25	5.00	15.00	4.50	6.00	15.00	5.00	7.00
12	10.00	7.00	16.00	9.00	5.00	6.00	40.00	3.00	9.00	40.00	10.00	8.00
13	2.00	6.00	9.50	7.50	2.00	0.50	8.00	1.00	1.00	12.00	3.00	4.00
14	8.00	4.00	12.00	7.00	4.50	2.00	12.00	4.00	4.00	25.00	9.00	5.00
15	24.00	8.00	11.50	8.00	7.70	5.00	10.00	5.00	9.00	30.00	8.00	7.00
16	7.00	5.00	5.00	8.00	5.00	0.50	20.00	2.50	1.25	20.00	7.00	3.00
17	9.00	6.00	12.00	8.00	5.50	3.00	50.00	5.00	5.00	15.00	7.00	7.00
18	20.00	10.00	16.00	8.50	8.00	7.00	30.00	3.00	7.00	35.00	9.00	9.00

Table 10. Example of Steps 1 through 4 in modulus equalization procedure.

	Subject											
Trial	1	2	3	4	5	6	7	8	9	10	11	12
1	0.48	0.48	1.04	0.92	0.44	-0.60	0.85	0.24	-0.30	1.00	0.85	0.48
2	0.70	0.70	1.06	0.93	0.54	0.48	1.00	0.70	0.60	1.40	0.60	0.70
3	1.30	0.85	1.06	0.95	0.83	0.60	1.00	0.48	0.60	1.30	0.85	0.90
4	0.48	0.70	0.85	0.70	0.57	0.00	1.08	0.60	-0.18	1.26	0.48	0.30
5	0.90	0.85	1.02	0.95	0.65	0.70	1.08	0.30	0.85	1.18	0.95	0.85
6	1.18	0.90	1.18	0.95	0.83	0.85	1.54	0.60	0.48	1.48	0.85	0.95
7	0.48	0.60	0.90	0.90	0.70	-0.60	0.70	0.30	-0.12	1.00	0.30	0.48
8	0.78	0.70	1.13	0.90	0.44	0.60	1.40	0.60	0.48	1.18	0.60	0.85
9	1.23	1.00	1.00	0.95	0.76	0.60	1.30	0.18	0.60	1.40	0.90	0.85
10	0.48	0.00	0.98	0.85	0.70	-0.30	0.90	0.48	-0.18	1.40	0.85	0.60
11	0.85	0.85	1.04	0.93	0.80	0.70	1.18	0.65	0.78	1.18	0.70	0.85
12	1.00	0.85	1.20	0.95	0.70	0.78	1.60	0.48	0.95	1.60	1.00	0.90
13	0.30	0.78	0.98	0.88	0.30	-0.30	0.90	0.00	0.00	1.08	0.48	0.60
14	0.90	0.60	1.08	0.85	0.65	0.30	1.08	0.60	0.60	1.40	0.95	0.70
15	1.38	0.90	1.06	0.90	0.89	0.70	1.00	0.70	0.95	1.48	0.90	0.85
16	0.85	0.70	0.70	0.90	0.70	-0.30	1.30	0.40	0.10	1.30	0.85	0.48
17	0.95	0.78	1.08	0.90	0.74	0.48	1.70	0.70	0.70	1.18	0.85	0.85
18	1.30	1.00	1.20	0.93	0.90	0.85	1.48	0.48	0.85	1.54	0.95	0.95
Individual Modulus	0.86	0.73	1.03	0.90	0.67	0.31	1.17	0.47	0.43	1.30	0.77	0.73
Common Modulus = 0.78												
Individual Offset	-0.08	0.05	-0.25	-0.12	0.11	0.48	-0.39	0.31	0.35	-0.51	0.01	0.05

Table 11. Common-modulus presence data from Experiment 1 after modulus equalization (still in log units).

	Subject											
Trial	1	2	3	4	5	6	7	8	9	10	11	12
1	0.40	0.52	0.79	0.80	0.55	-0.13	0.46	0.55	0.05	0.49	0.85	0.53
2	0.62	0.75	0.81	0.81	0.65	0.95	0.61	1.01	0.95	0.88	0.61	0.75
3	1.22	0.89	0.81	0.83	0.94	1.08	0.61	0.79	0.95	0.79	0.85	0.96
4	0.40	0.75	0.60	0.58	0.68	0.48	0.69	0.91	0.17	0.74	0.49	0.35
5	0.82	0.89	0.77	0.83	0.76	1.17	0.69	0.61	1.20	0.66	0.96	0.90
6	1.10	0.95	0.93	0.83	0.94	1.32	1.15	0.91	0.83	0.96	0.85	1.01
7	0.40	0.65	0.65	0.78	0.81	-0.13	0.31	0.61	0.23	0.49	0.31	0.53
8	0.70	0.75	0.88	0.78	0.55	1.08	1.01	0.91	0.83	0.66	0.61	0.90
9	1.15	1.05	0.75	0.83	0.86	1.08	0.91	0.49	0.95	0.88	0.91	0.90
10	0.40	0.05	0.73	0.72	0.81	0.17	0.51	0.79	0.17	0.88	0.85	0.66
11	0.76	0.89	0.79	0.81	0.90	1.17	0.79	0.96	1.13	0.66	0.71	0.90
12	0.92	0.89	0.95	0.83	0.81	1.25	1.21	0.79	1.31	1.09	1.01	0.96
13	0.22	0.83	0.73	0.75	0.41	0.17	0.51	0.31	0.35	0.56	0.49	0.66
14	0.82	0.65	0.83	0.72	0.76	0.78	0.69	0.91	0.95	0.88	0.96	0.75
15	1.30	0.95	0.81	0.78	0.99	1.17	0.61	1.01	1.31	0.96	0.91	0.90
16	0.76	0.75	0.45	0.78	0.81	0.17	0.91	0.71	0.45	0.79	0.85	0.53
17	0.87	0.83	0.83	0.78	0.85	0.95	1.31	1.01	1.05	0.66	0.85	0.90
18	1.22	1.05	0.95	0.81	1.01	1.32	1.09	0.79	1.20	1.03	0.96	1.01

Table 12. Final presence data from Experiment 1 (after conversion back into units of subjective magnitude).

	Subject											
Trial	1	2	3	4	5	6	7	8	9	10	11	12
1	2.49	3.35	6.20	6.28	3.52	0.75	2.85	3.58	1.12	3.06	7.16	3.39
2	4.15	5.58	6.48	6.43	4.48	8.96	4.08	10.22	8.98	7.65	4.09	5.65
3	16.61	7.81	6.48	6.81	8.65	11.95	4.08	6.13	8.98	6.12	7.16	9.04
4	2.49	5.58	3.94	3.78	4.74	2.99	4.89	8.18	1.48	5.51	3.07	2.26
5	6.64	7.81	5.91	6.81	5.77	14.94	4.89	4.09	15.72	4.59	9.21	7.91
6	12.46	8.92	8.45	6.81	8.71	20.91	14.27	8.18	6.74	9.18	7.16	10.17
7	2.49	4.46	4.51	6.05	6.41	0.75	2.04	4.09	1.68	3.06	2.05	3.39
8	4.98	5.58	7.60	6.05	3.52	11.95	10.20	8.18	6.74	4.59	4.09	7.91
9	14.12	11.16	5.63	6.81	7.30	11.95	8.16	3.07	8.98	7.65	8.18	7.91
10	2.49	1.12	5.35	5.29	6.41	1.49	3.26	6.13	1.48	7.65	7.16	4.52
11	5.81	7.81	6.20	6.43	8.01	14.94	6.12	9.20	13.48	4.59	5.11	7.91
12	8.31	7.81	9.01	6.81	6.41	17.92	16.31	6.13	20.22	12.24	10.23	9.04
13	1.66	6.69	5.35	5.67	2.56	1.49	3.26	2.04	2.25	3.67	3.07	4.52
14	6.64	4.46	6.76	5.29	5.77	5.97	4.89	8.18	8.98	7.65	9.21	5.65
15	19.93	8.92	6.48	6.05	9.86	14.94	4.08	10.22	20.22	9.18	8.18	7.91
16	5.81	5.58	2.82	6.05	6.41	1.49	8.16	5.11	2.81	6.12	7.16	3.39
17	7.47	6.69	6.76	6.05	7.05	8.96	20.39	10.22	11.23	4.59	7.16	7.91
18	16.61	11.16	9.01	6.43	10.25	20.91	12.23	6.13	15.72	10.71	9.21	10.17

This procedure was used on all magnitude estimation data collected in the experimental series. It was used to adjust the presence scores from each experiment to the common modulus of that experiment and was similarly used in the analysis of line length estimation data. All graphs and analyses of magnitude estimation data reported below were performed on estimates adjusted to a common modulus by the procedure just described.

Line Length Estimation

An analysis was done of the line length estimation data collected from each subject during the magnitude estimation practice conducted at the start of each experimental session. This was done to examine the distribution of the data, both raw and after adjustment to a common modulus, and to compare the exponent of the resulting power function to the figure of 1.0 obtained in several other examinations of this psychophysical function (Collins and Gescheider, 1989; Gescheider, 1985; Zwislocki, 1983).

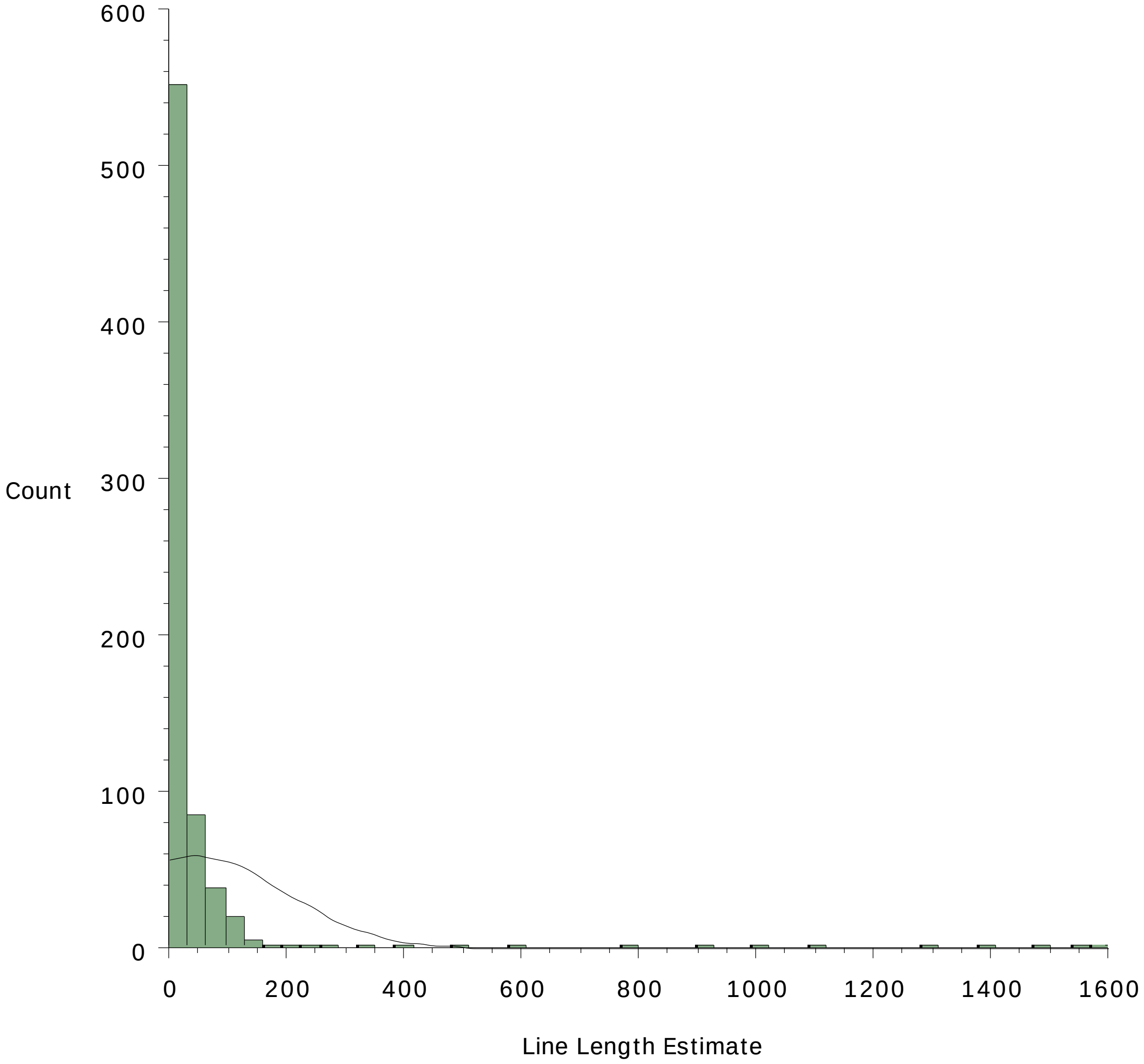


Figure 20. Distribution of line length estimates in comparison to normal.

Figure 20 shows the distribution of the raw line length estimates of all thirty-six subjects. Superimposed on this histogram is a normal probability density function with the same mean and standard deviation. As can be seen in this figure, the data are quite positively skewed, as would be expected based on published research and the assumed lognormality of these estimates.

Figure 21 shows the distribution of the line length estimates of all thirty-six subjects after adjustment to a common modulus (the common modulus in this case was 11.6). Superimposed on this histogram is a normal probability density function with the same mean and standard deviation. As shown in this figure, the data now appear less skewed and more normally distributed.

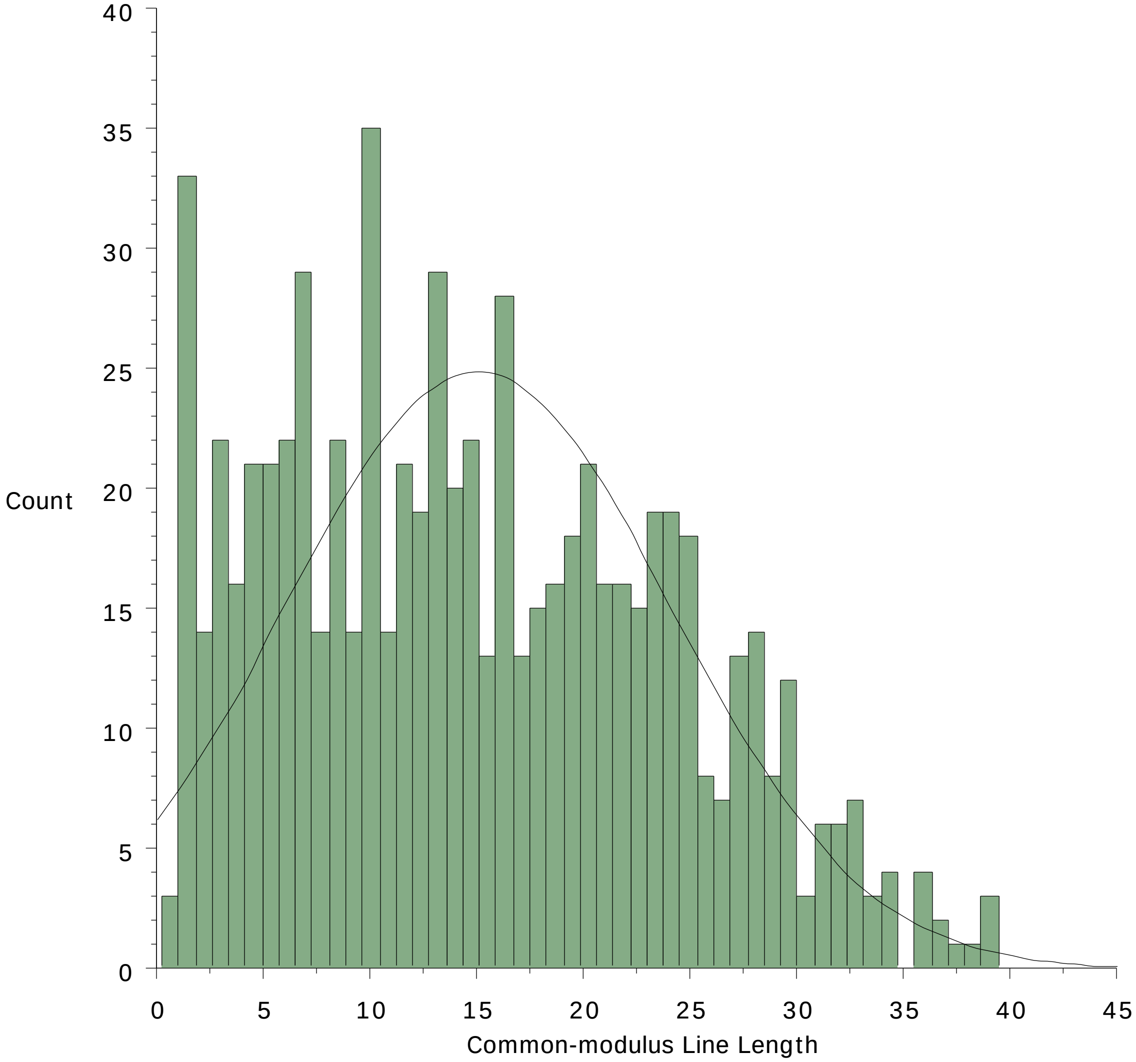


Figure 21. Distribution of common-modulus line length estimates in comparison to normal.

A linear regression was performed with log line length as the regressor and the log of the arithmetic mean of subjects' common-modulus line length estimates as the response. The reason for using logarithmic values in this case is to make the slope of the resulting regression equation equivalent to the exponent in the power function relating estimated (or perceived) line length to actual line length. This exponent is then directly comparable to similar values obtained by other researchers. The resulting regression equation is $y = 1.002x + 0.63$ with $R^2 = 0.998$. The line produced by this equation is shown in Figure 22 superimposed upon the data.

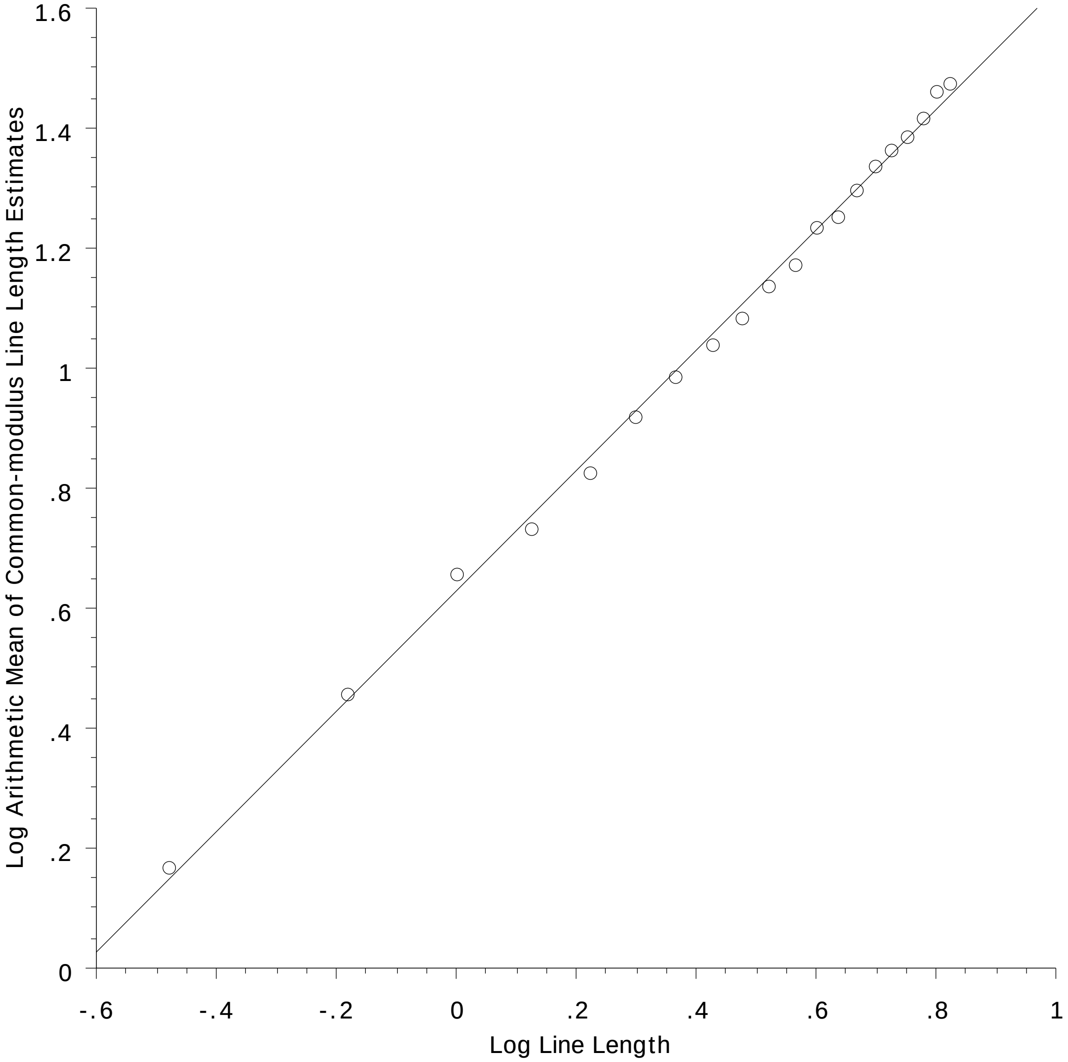


Figure 22. Plot of the power function relating estimated line length to actual line length.

Presence

An analysis similar to the analysis of line length described above was performed on all 624 presence estimates obtained in the experimental series to examine the distribution of this data, both raw and after adjustment to a common modulus. Simple linear regressions on this data (like that done on line length estimates) were not performed primarily because there exists no published body of research establishing exponents of psychophysical power functions for comparison (in contrast to the psychophysical relationship between actual and perceived line length).

Figure 23 shows the distribution of the raw presence estimates of all thirty-six subjects. Superimposed on this histogram is a normal probability density function with the same mean and standard deviation.

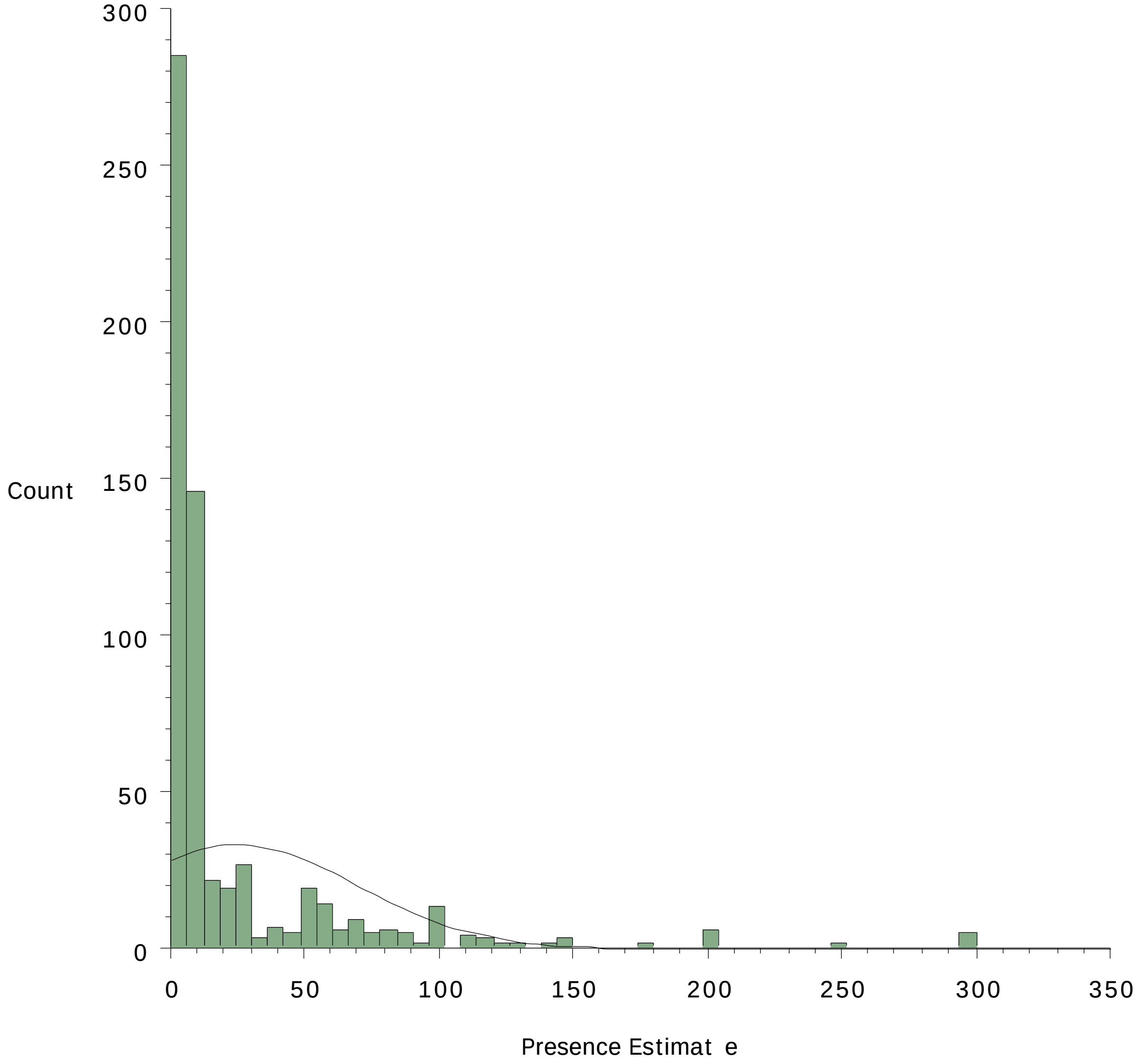


Figure 23. Distribution of presence estimates in comparison to normal.

Figure 24 shows the distribution of the presence estimates of the thirty-six subjects after adjustment to a common modulus (the common modulus in this case was 8.91). Again, superimposed on this histogram is a normal probability density function with the same mean and standard deviation.

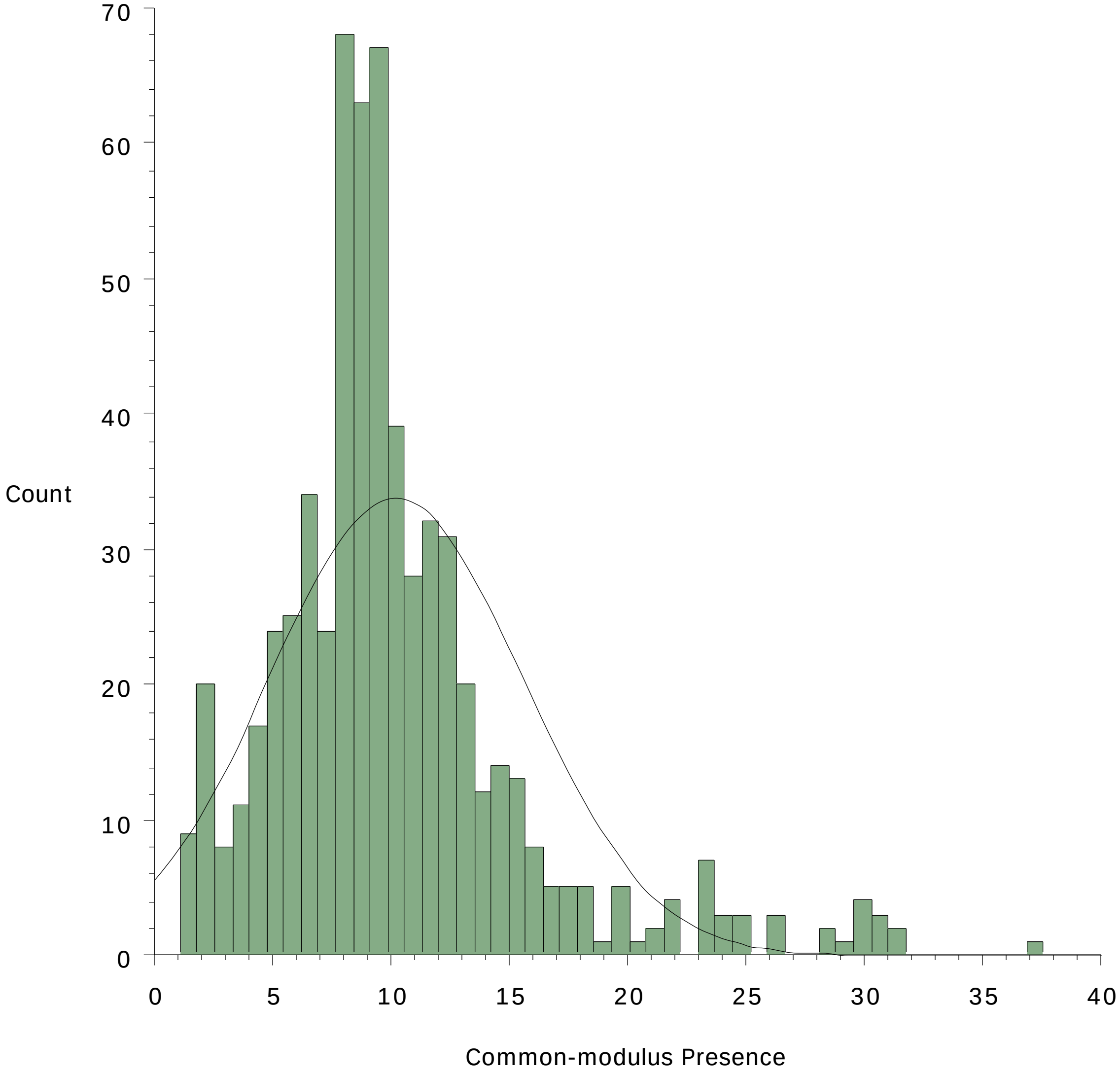


Figure 24. Distribution of common-modulus presence estimates in comparison to normal.

Similar to the distribution of line length estimates and in accordance with the assumption of lognormality, these data appear very positively skewed prior to adjustment (Figure 23), but near normal afterward (Figure 24).

Experiment 1

Performance

Analyses of variance (ANOVAs) were conducted for each dependent and independent variable. Actual distance was treated as an independent variable in conducting the ANOVAs for distance estimation performance. The ANOVA summary tables referred to throughout this section may be found in **APPENDIX D**.

The ANOVA summary table for the effects of scene update rate, visual display resolution and field of view on distance estimation is shown in Table 41. Actual distance was treated as an independent variable for purposes of this analysis. As shown in this table, the effects of actual distance, visual display resolution, field of view, and all interactions among these three variables were statistically significant with $\alpha \leq 0.05$. These effects are shown in Figures 25 through 27. These figures are in the format used by Lampton, Knerr, et al. (1994) in the paper in which they originally proposed this task. The straight lines in the figures represent perfect distance estimation performance.

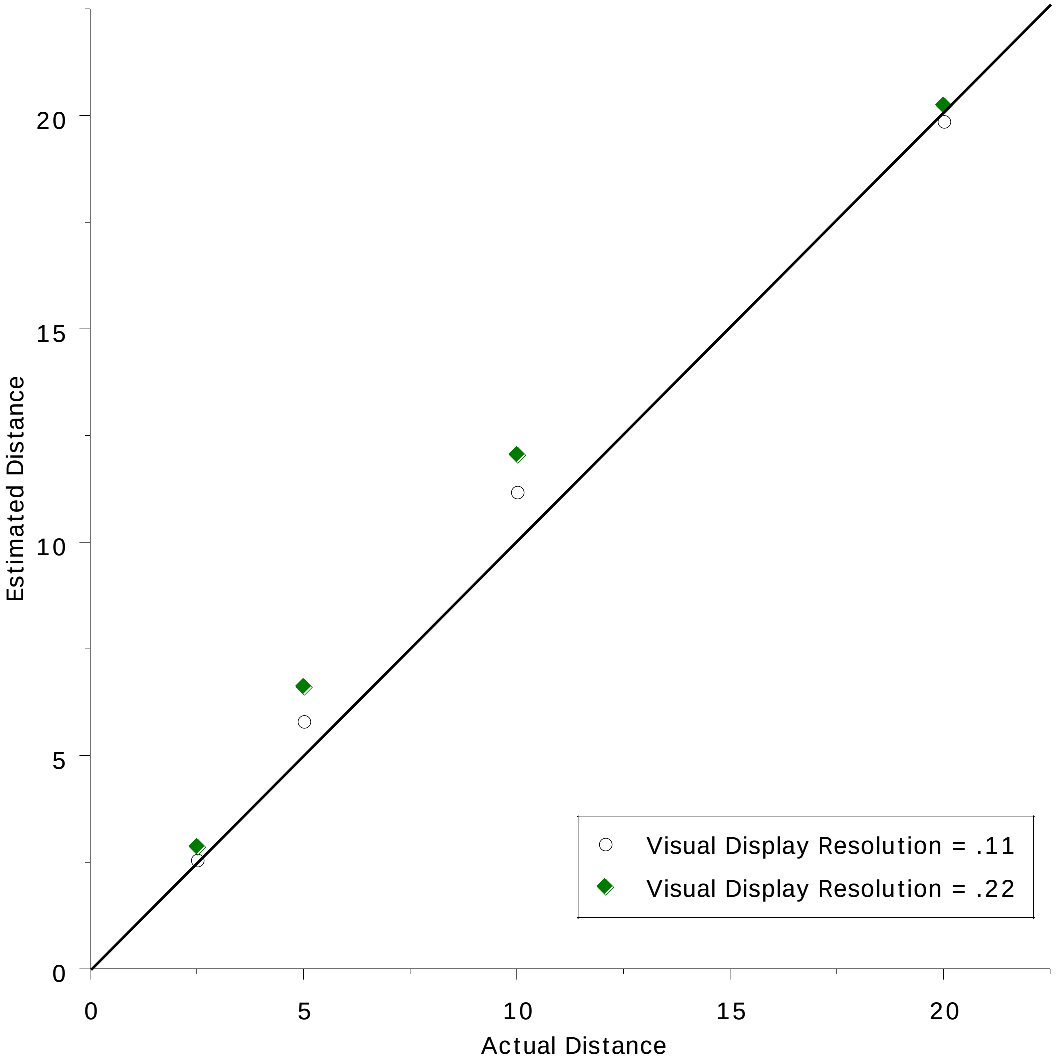


Figure 25. Effect of visual display resolution on distance estimation.

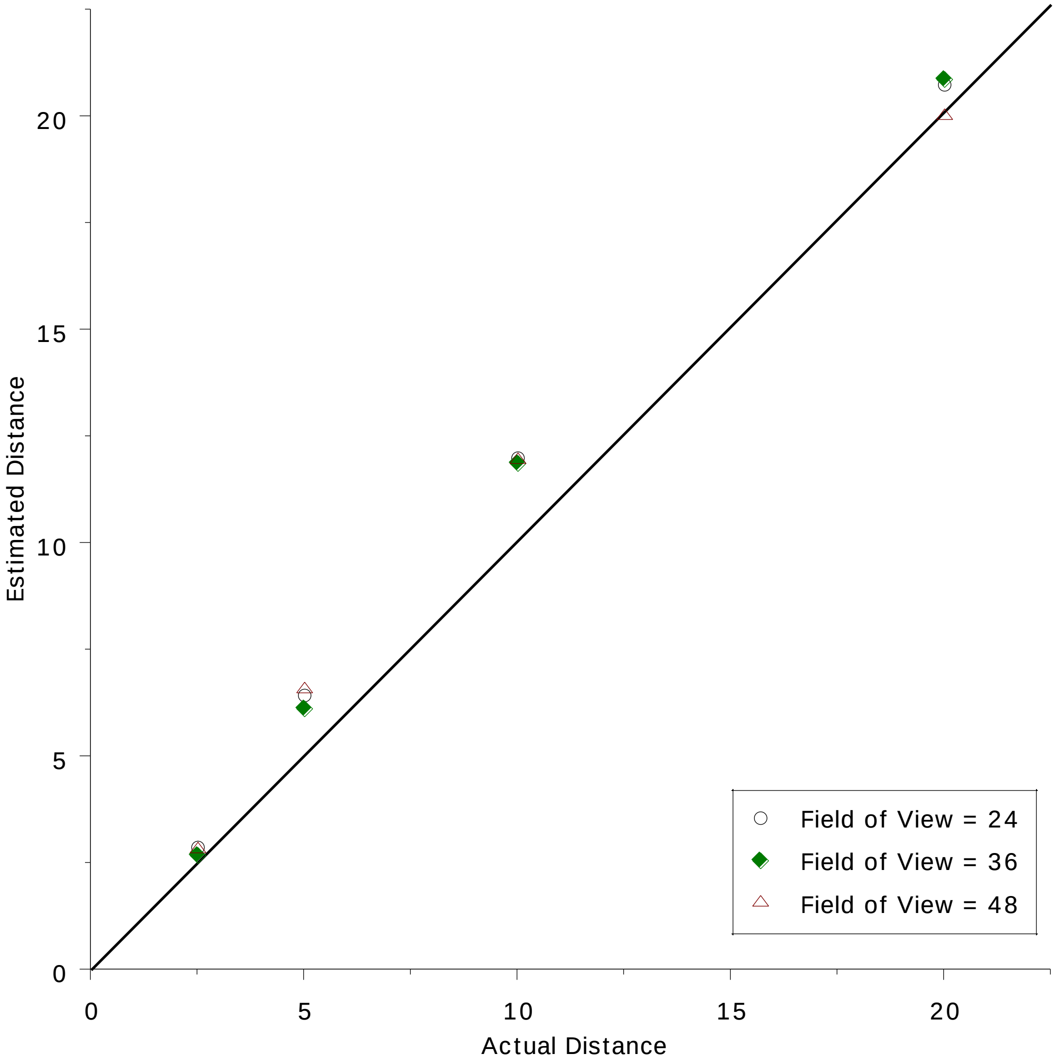


Figure 26. Effect of field of view on distance estimation.

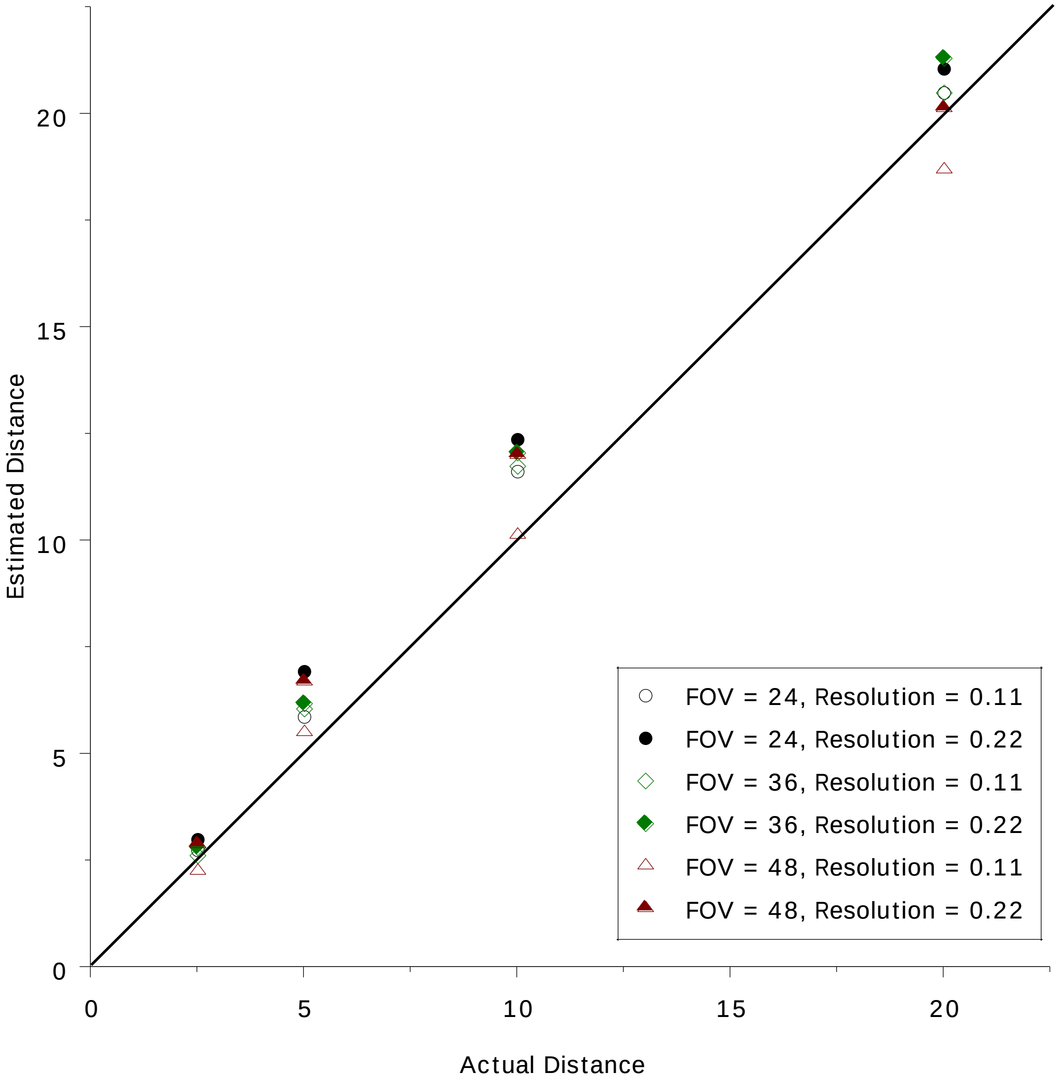


Figure 27. Effect of the interaction between field of view (FOV) and visual display resolution (Resolution) on distance estimation.

The ANOVA summary table for the effects of scene update rate, visual display resolution and field of view on time to complete the Bins task and errors committed during completion of the Turns task are shown in Tables 42 and 43. There were no statistically significant effects of these three variables or their interactions on either of these two performance measures.

The ANOVA summary table for the effects of these same variables on time to complete the Turns task is shown in Table 44. The effect of scene update rate was statistically significant with $\alpha \leq 0.05$ and the effect of field of view was significant with $\alpha = 0.06$. These effects are shown in Figures 28 and 29, respectively. Error bars represent the standard error of the mean in all figures.

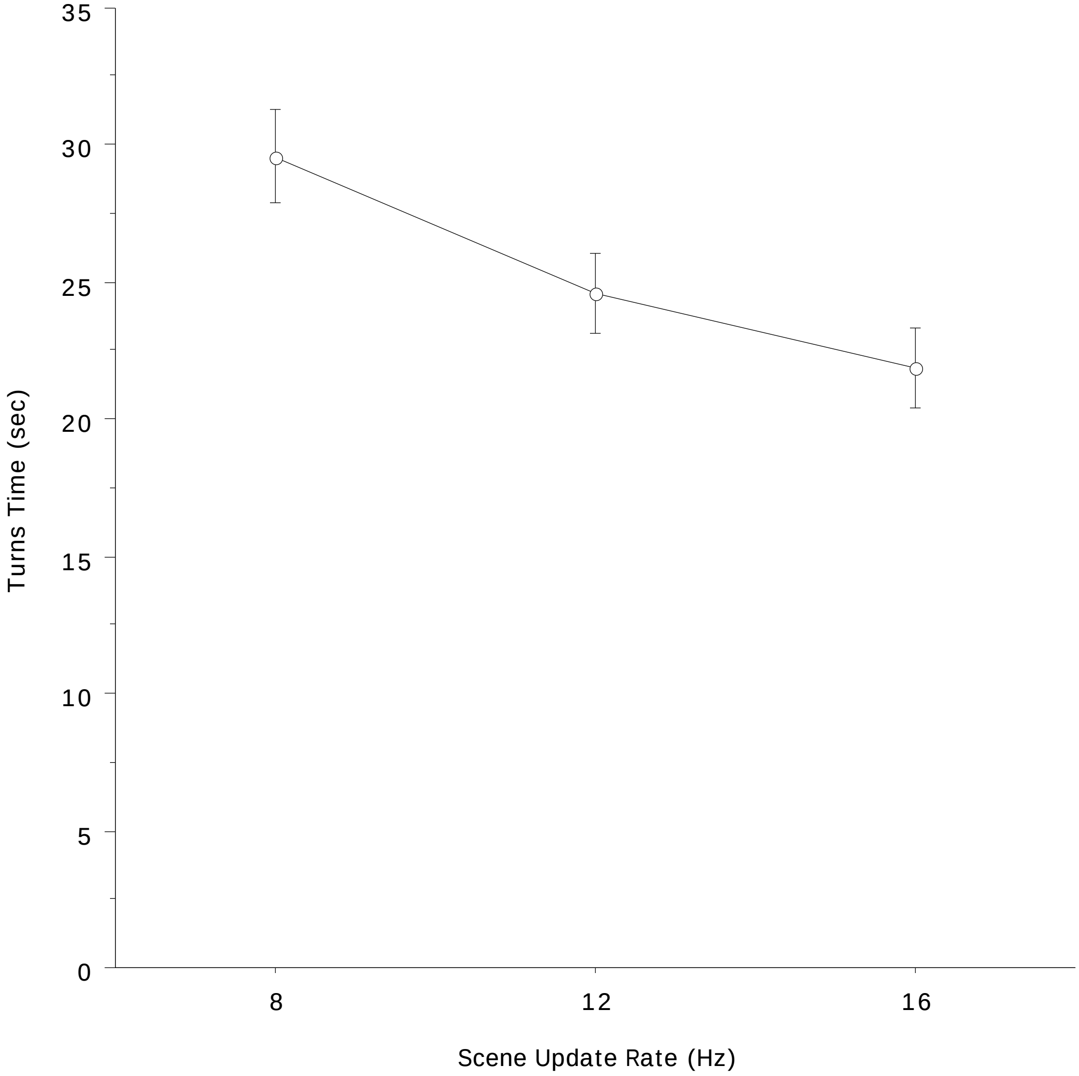


Figure 28. The effect of scene update rate on time to complete the Turns task.

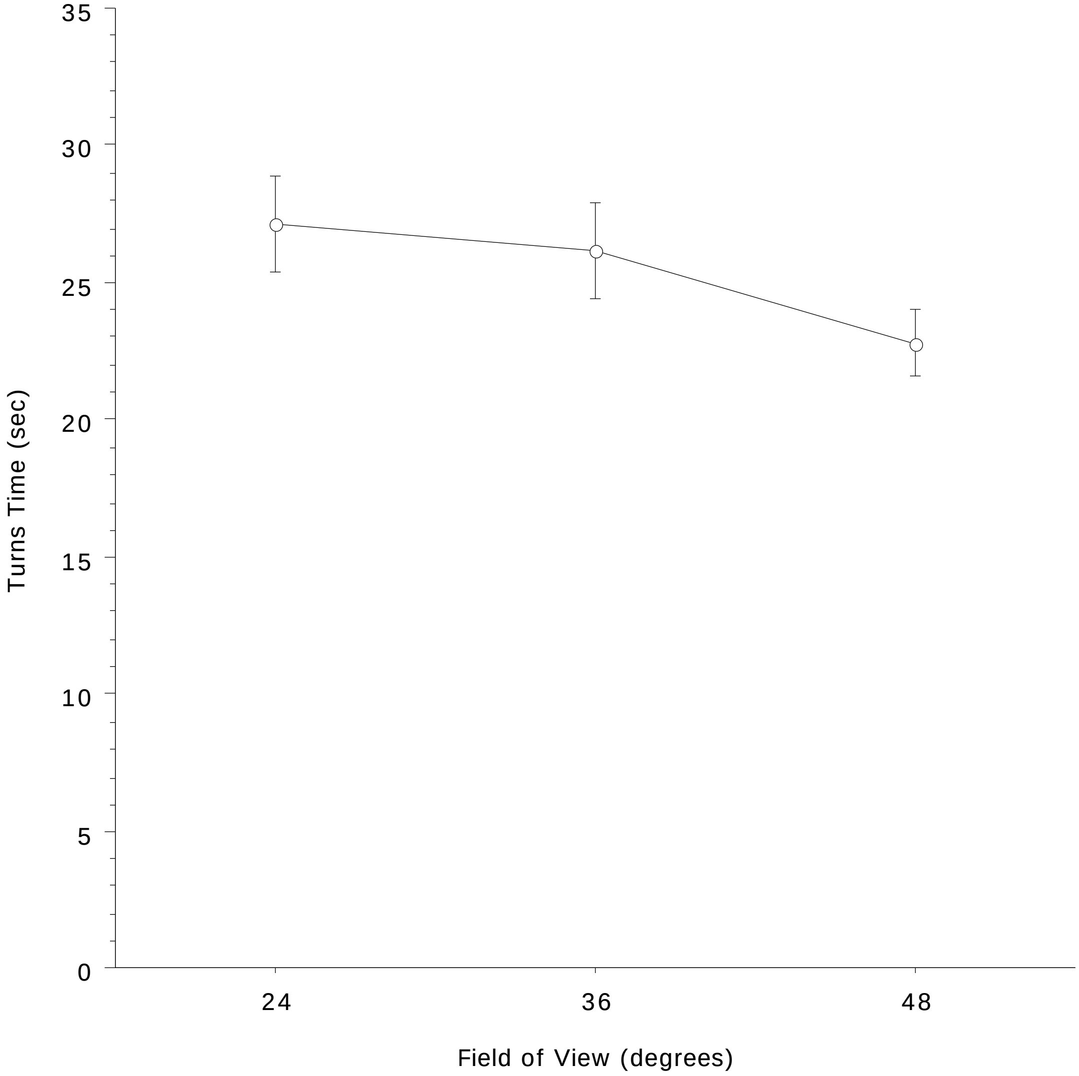


Figure 29. The effect of field of view on time to complete the Turns task.

Table 45 shows the ANOVA summary table for the effects of scene update rate, visual display resolution and field of view on time to complete the Search task . The effects of visual display resolution, field of view, and the interaction between the two were statistically significant with $\alpha \leq 0.05$. The mean search time in the low visual display resolution condition was 11.95 seconds with a standard error of 1.08 seconds. The mean search time in the high visual display resolution condition was 8.81 seconds with a standard error of 0.45 seconds. The effects of field of view and the interaction between field of view and visual display resolution are shown in Figures 30 and 31, respectively.

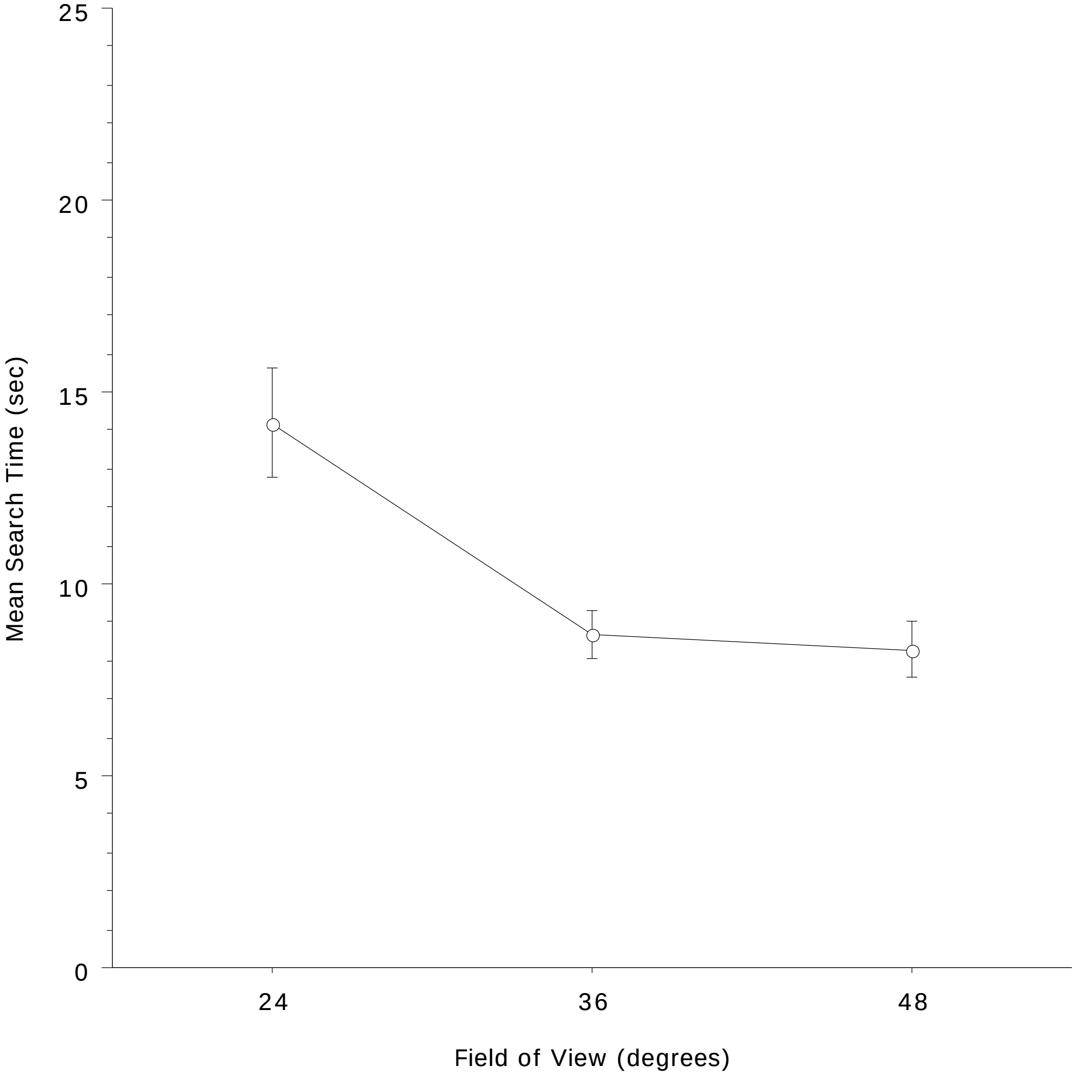


Figure 30. Effect of field of view on search time.

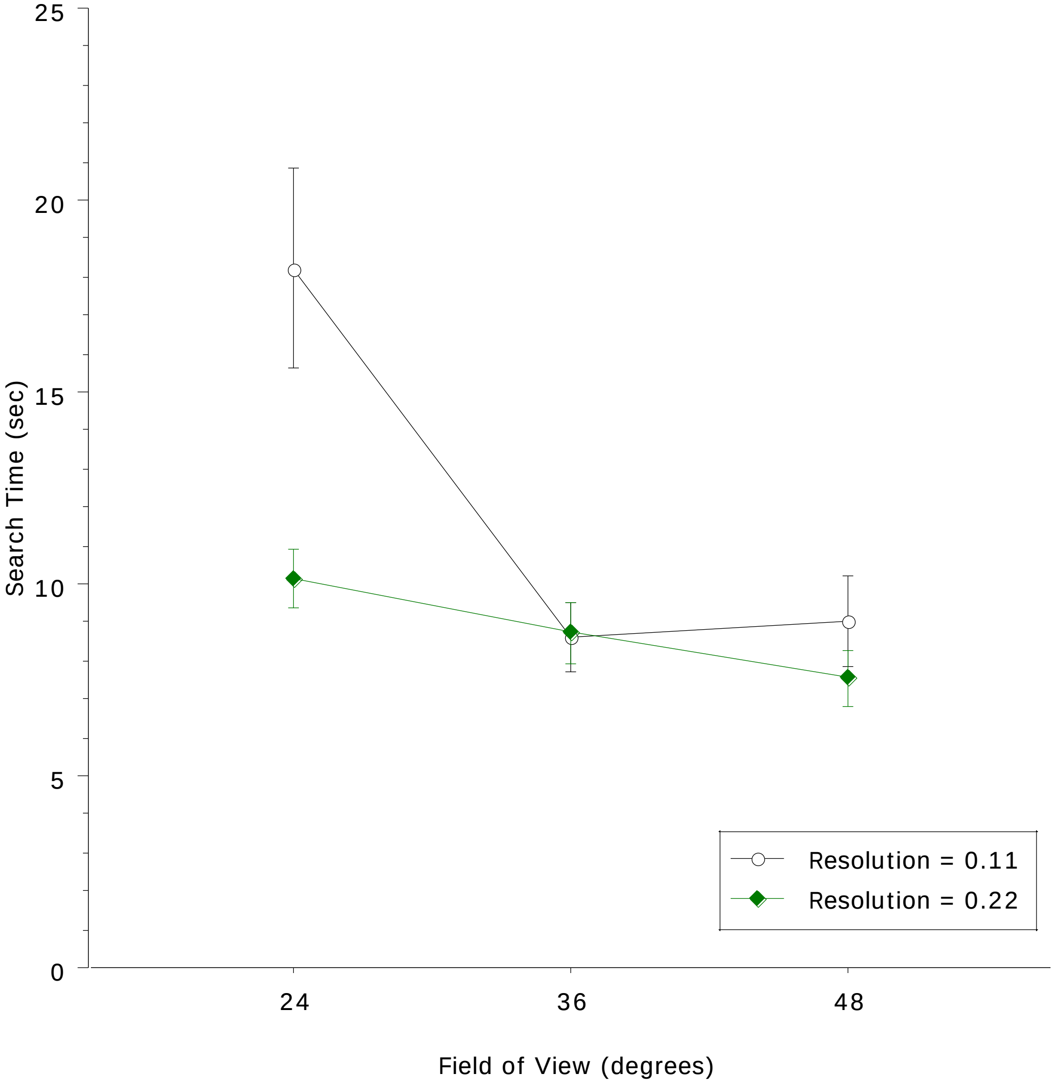


Figure 31. Effect of interaction between field of view and visual display resolution on search time.

The ANOVA summary table for the effects of scene update rate, visual display resolution and field of view on time to complete the Choice task is shown in Table 46. The effect of scene update rate was statistically significant with $\alpha \leq 0.05$. This effect is shown in Figure 32. No errors were made by any subjects in the completion of this task.

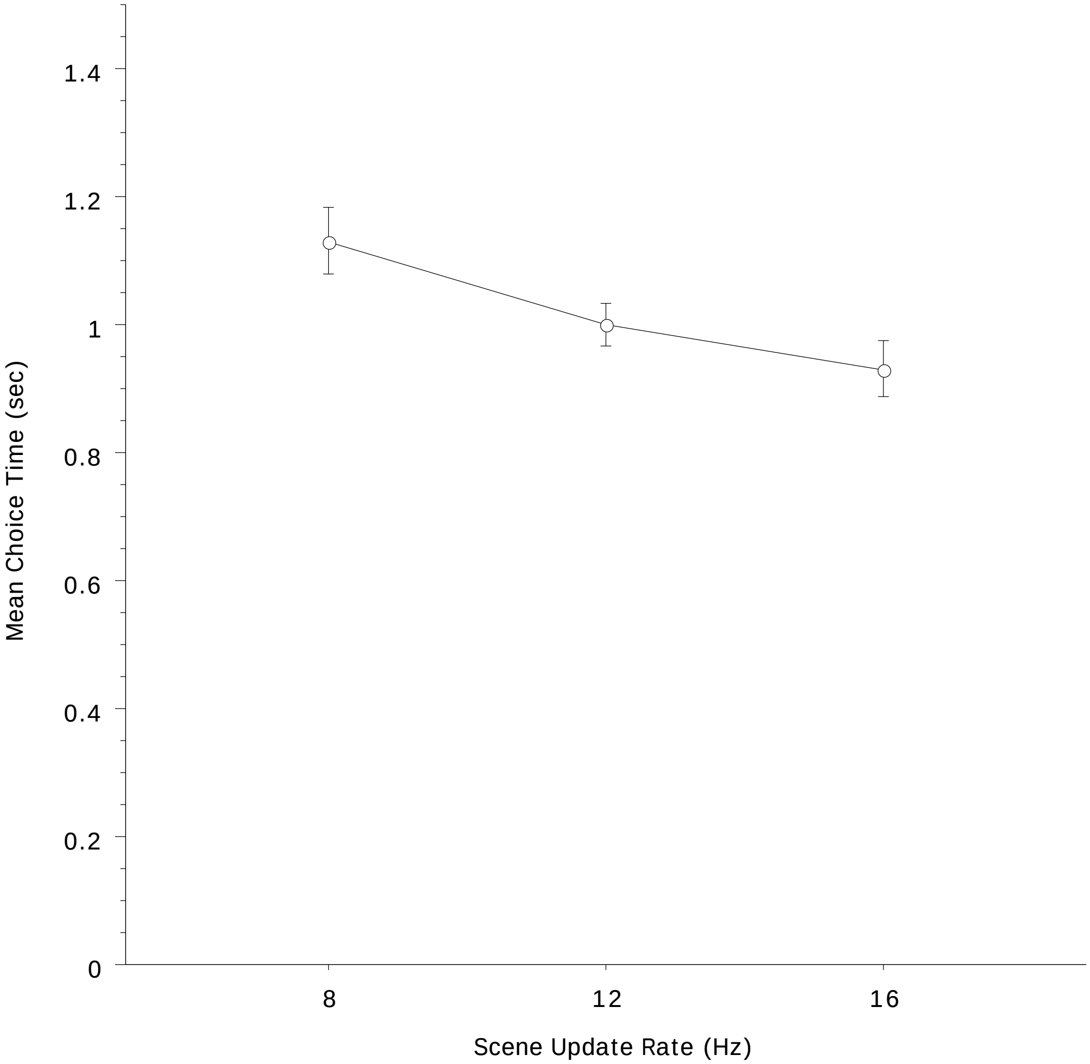


Figure 32. Effect of scene update rate on choice time.

Second-order polynomial regressions were calculated using the independent variables in this experiment — plus minutes spent in the VE and perceived presence — as regressors to predict each of the dependent variables measuring performance. All regressions reported use the same notation as in Table 4: x_1 = scene update rate, x_2 = visual display resolution, x_3 = field of view, x_4 = sound, x_5 = textures, x_6 = head-tracking, x_7 = stereopsis, x_8 = virtual personal risk, x_9 = interactions, x_{10} = second user, x_{11} = environmental detail. Additionally, x_{12} = minutes spent in the VE, x_{13} = perceived presence, and x_{14} = actual distance (in distance estimation models only) .

In performing these regressions, all data and regressors were first standardized. This was done for two reasons: 1) so that the resulting parameter estimates (the $\hat{\beta}$s) would all be in the same units (i.e., z-scores), allowing direct comparison of the influence of the various regressors on the predicted response in each regression question, and, 2) to reduce multicollinearity among the regressors resulting from inclusion of second-order terms (Montgomery and Peck, 1992). Regressions were calculated for all possible combinations of all regressors for each predicted response. The regression equations reported below represent, for each predicted response, the model representing the combination of smallest residual mean square (equivalent to selecting the model with the largest adjusted R^2) and smallest prediction error sum of squares (PRESS). These models are shown in Tables 13 through 18.

Table 13. Regression equation predicting distance estimation in Experiment 1.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_1	0.02	0.01	3.42	0.001
x_2	0.06	0.01	7.79	0.0001
x_3	-0.03	0.01	-4.15	0.0001
x_{14}	1.02	0.01	97.33	0.0001
x_{12}	-0.03	0.01	-4.08	0.0001
x_1^2	0.04	0.01	3.54	0.0004
x_3^2	0.02	0.01	1.93	0.054
x_{12}^2	0.04	0.01	6.12	0.0001
x_{13}^2	-0.02	0.004	-5.23	0.0001
x_1x_2	0.02	0.01	2.13	0.034
x_2x_3	0.02	0.01	2.70	0.007
x_{14}^2	-0.09	0.01	-8.79	0.0001
x_2x_{14}	0.02	0.01	2.81	0.005
x_3x_{14}	-0.01	0.01	-1.69	0.091

$MS_E = 0.056$, PRESS = 49.73, $R^2 = 0.944$

$F_{14,\,850}$ for H_0: No linear relationship = 1028.46, $p < 0.0001$

Table 14. Regression equation predicting time to complete the Bins task in Experiment 1.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_1	-0.13	0.07	-1.88	0.062
x_2	0.09	0.07	1.27	0.205
x_{12}	-0.17	0.07	-2.50	0.013
x_{13}	0.10	0.07	1.48	0.141
x_{12}^2	0.06	0.04	1.45	0.148

$MS_E = 0.960$, PRESS = 212.28, $R^2 = 0.058$

$F_{5,211}$ for H_0: No linear relationship = 2.58, $p < 0.028$

Table 15. Regression equation predicting time to complete the Turns task in Experiment 1.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_1	-0.25	0.06	-4.20	0.0001
x_3	-0.11	0.06	-1.76	0.080
x_{12}	-0.25	0.06	-4.08	0.0001
x_3^2	-0.18	0.06	-3.22	0.002
x_{12}^2	0.26	0.05	5.76	0.0001

$MS_E = 0.778$, PRESS = 174.16, $R^2 = 0.237$

$F_{5, 211}$ for H_0: No linear relationship = 13.08, $p < 0.0001$

Table 16. Regression equation predicting errors in the Turns task in Experiment 1.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_{13}	-0.15	0.07	-2.32	0.021
x_3^2	-0.17	0.06	-2.69	0.008
x_{12}^2	0.21	0.05	4.17	0.0001

$MS_E = 0.901$, PRESS = 197.79, $R^2 = 0.107$

$F_{3,\,213}$ for H_0: No linear relationship = 8.52, $p < 0.0001$

Table 17. Regression equation predicting time to complete the Search task in Experiment 1.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_2	-0.15	0.06	-2.30	0.023
x_3	-0.14	0.08	-1.60	0.112
x_{12}	-0.14	0.07	-2.16	0.032
x_{13}	-0.25	0.11	-2.34	0.020
x_{13}^2	0.07	0.04	2.06	0.041
x_2x_3	0.15	0.06	2.42	0.016

$MS_E = 0.831$, PRESS = 185.80, $R^2 = 0.188$

$F_{6,\,210}$ for H_0: No linear relationship = 8.12, $p < 0.0001$

Table 18. Regression equation predicting time to complete the Choice task in Experiment 1.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_1	-0.21	0.07	-3.21	0.002
x_2	0.15	0.07	2.17	0.031
x_{12}	-0.08	0.07	-1.24	0.218
x_{13}	-0.07	0.07	-0.98	0.330

$MS_E = 0.934$, PRESS = 205.31, $R^2 = 0.079$

$F_{4,\ 212}$ for H_0: No linear relationship = 4.56, $p < 0.002$

Presence

The ANOVA summary table for the effects of scene update rate, visual display resolution and field of view on presence is shown in Table 47. As shown in this table, the main effects of scene update rate, visual display resolution, and field of view — but none of their interactions — were all significant with $\alpha \leq 0.05$. Mean presence in the low visual display resolution condition was 6.40 with a standard error of 0.34. Mean presence in the high visual display resolution condition was 7.85 with a standard error of 0.41. The effects of scene update rate and field of view are displayed in Figures 33 and 34.

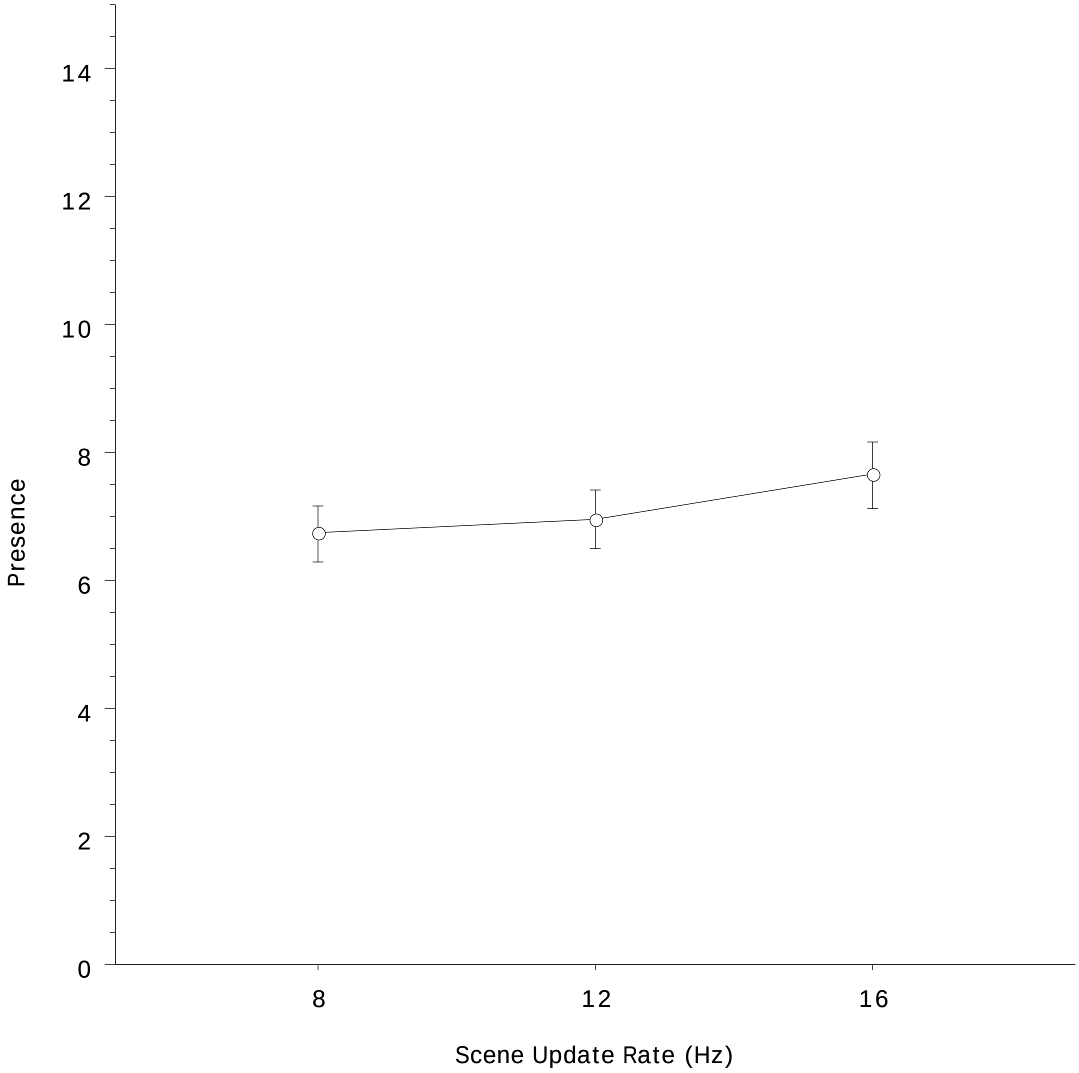


Figure 33. Effect of scene update rate on presence.

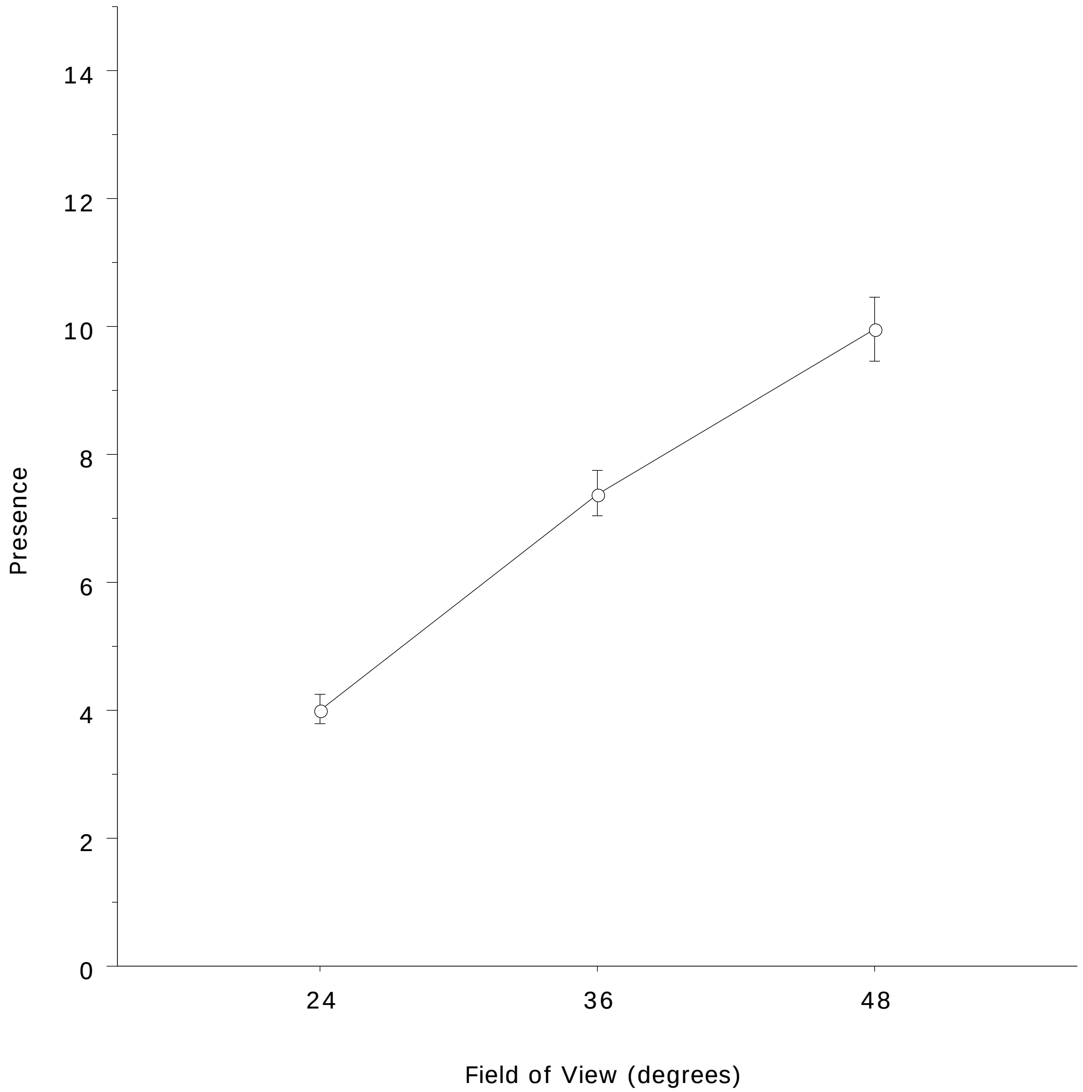


Figure 34. Effect of field of view on presence.

A second-order polynomial regression was done in the same manner as those done for performance to predict perceived presence from the independent variables manipulated in this experiment (plus minutes spent in the VE). This resulted in the equation contained in Table 19.

Table 19. Regression equation predicting perceived presence in Experiment 1.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_1	0.10	0.05	1.99	0.048
x_2	0.17	0.05	3.35	0.001
x_3	0.59	0.05	11.61	0.0001
x_{12}	0.21	0.05	4.12	0.0001
x_{12}^2	-0.08	0.03	-2.26	0.025

$MS_E = 0.548$, PRESS = 120.56, $R^2 = 0.462$

$F_{5,\,211}$ for H_0: No linear relationship = 36.26, $p < 0.0001$

Experiment 2

Performance

As for Experiment 1, analyses of variance (ANOVAs) were conducted for each dependent and independent variable. The ANOVA summary table for the effects of sound, textures, head-tracking, stereopsis, and virtual personal risk on performance of the distance estimation task is shown in Table 48. The effects of textures, stereopsis, the interaction between sound and stereopsis, and the interaction among actual distance, sound, and textures were all statistically significant with $\alpha \leq 0.05$ (in addition to the effect of actual distance). These effects are shown in Figures 35 through 38.

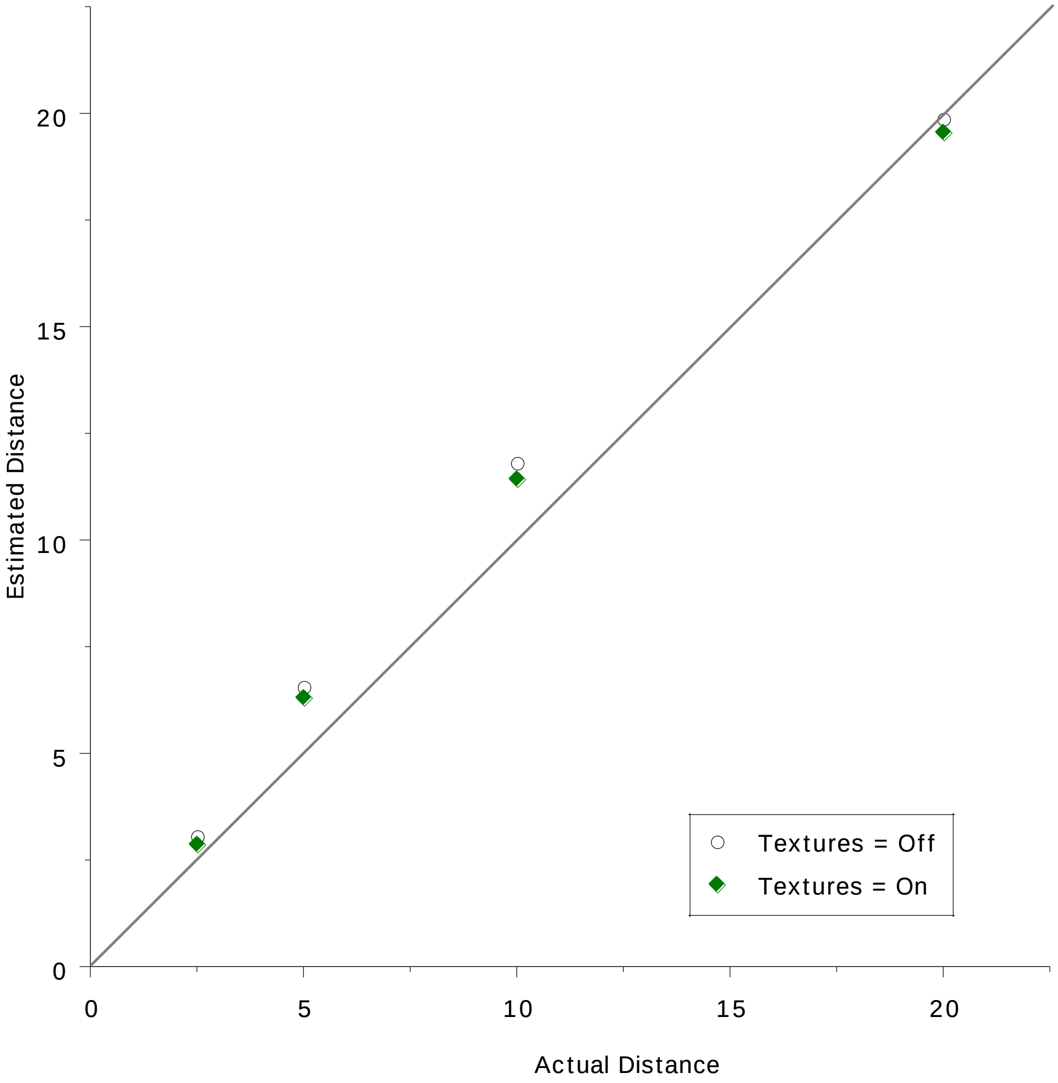


Figure 35. Effect of textures on distance estimation.

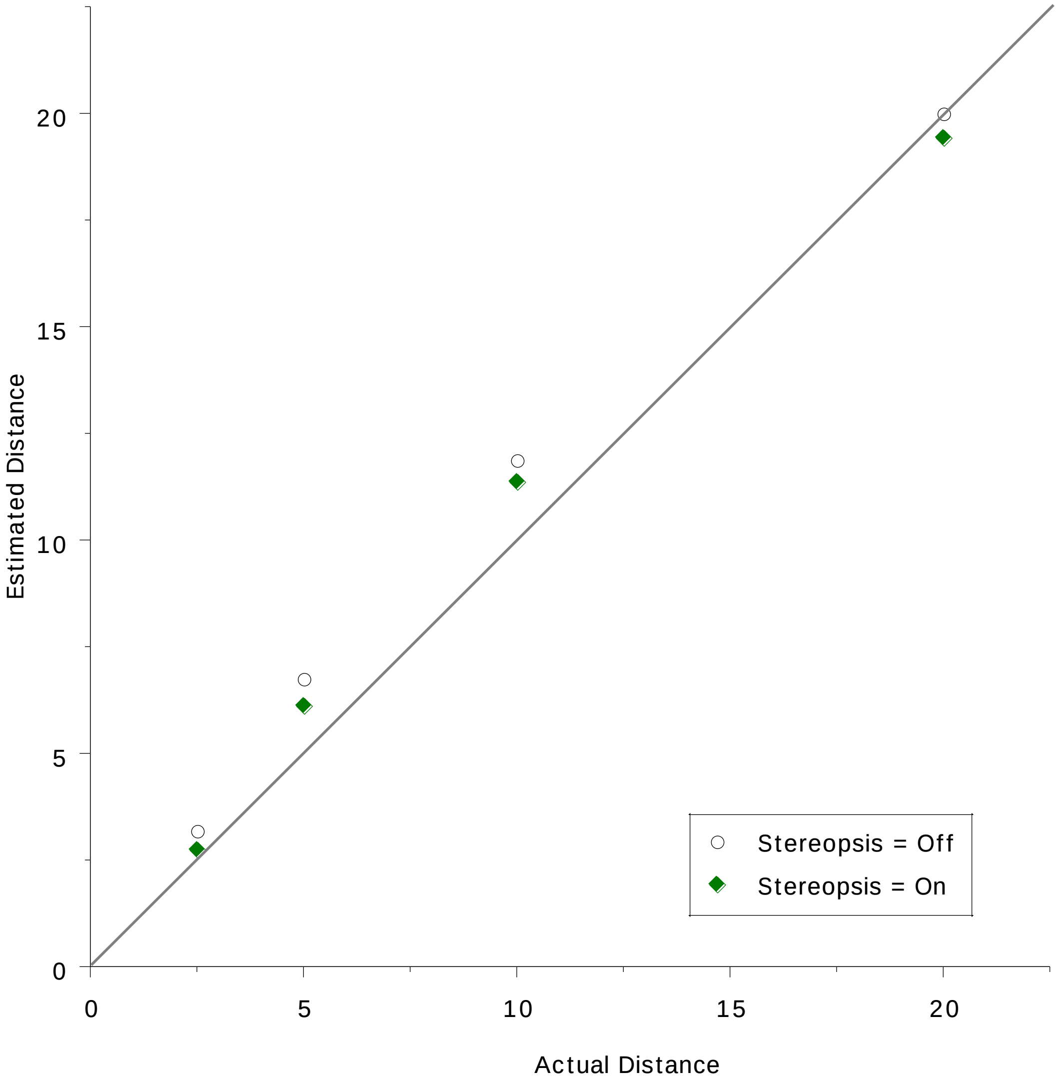


Figure 36. Effect of stereopsis on distance estimation.

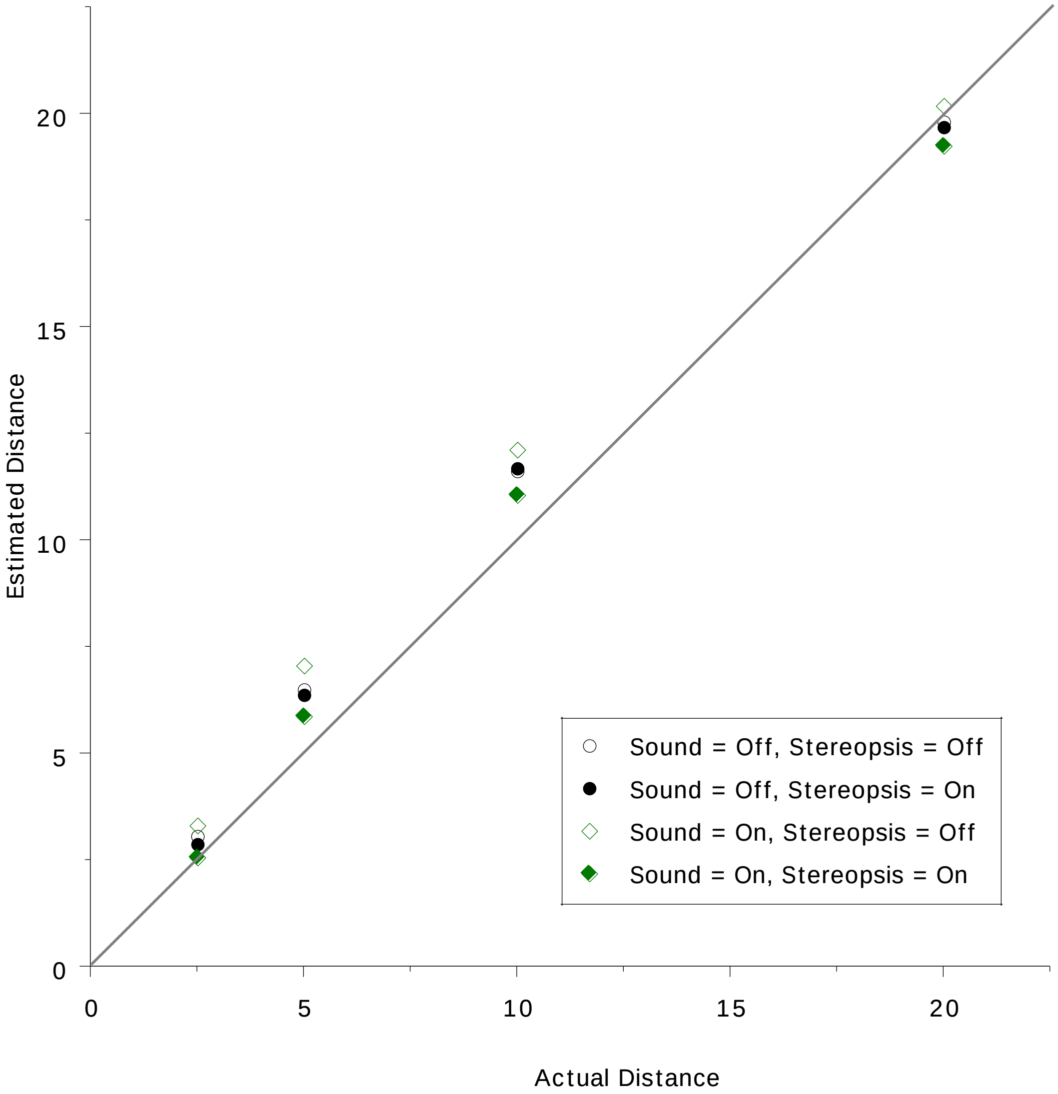


Figure 37. Effect of interaction between sound and stereopsis on distance estimation.

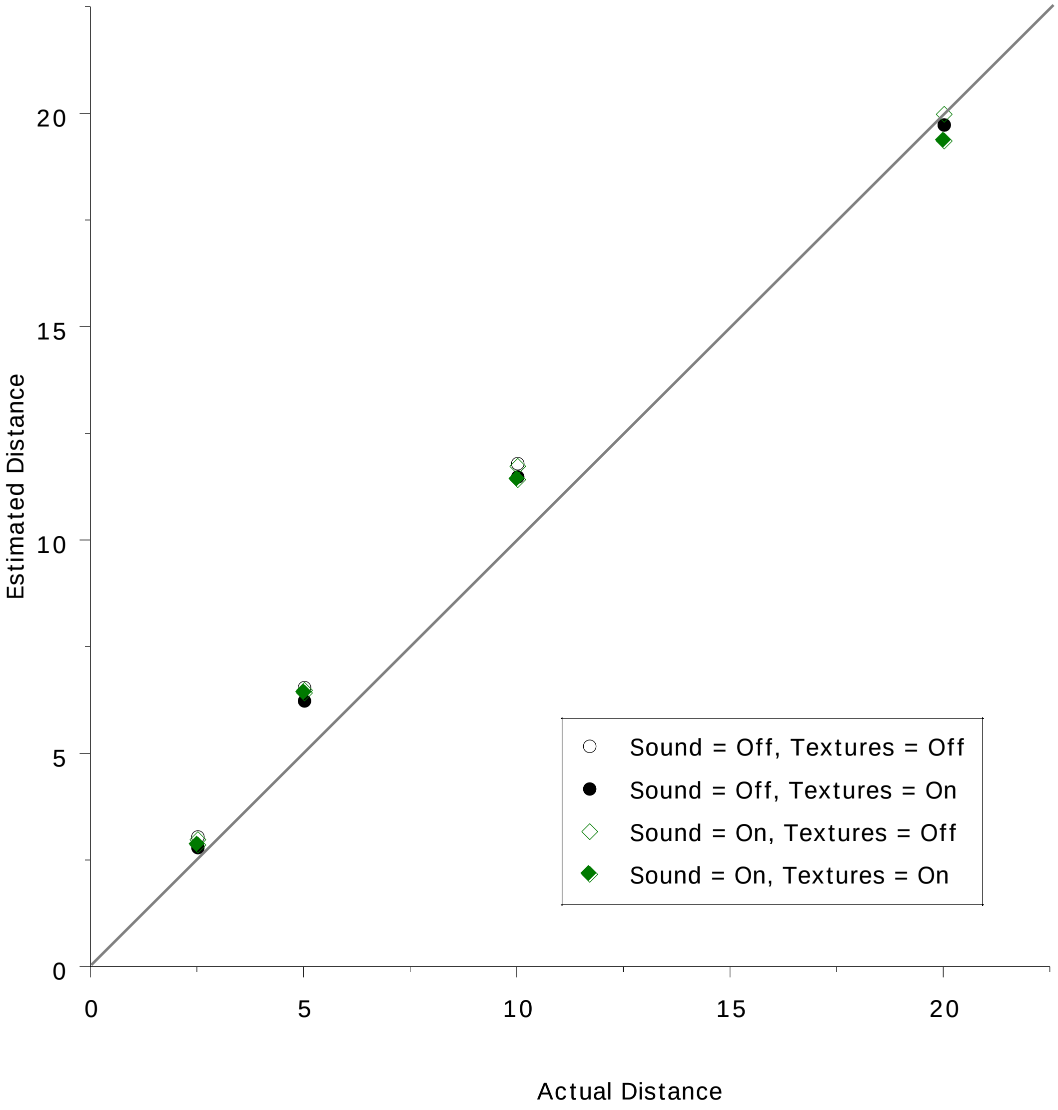


Figure 38. Effect of interaction among actual distance, sound and textures on distance estimation.

The ANOVA summary table for the effects of sound, textures, head-tracking, stereopsis, and virtual personal risk on time to complete the Bins task is shown in Table 49. The effects of the interaction between sound and stereopsis, and the interaction between textures and head-tracking were statistically significant with $\alpha \leq 0.05$. These effects are shown in Figures 39 and 40.

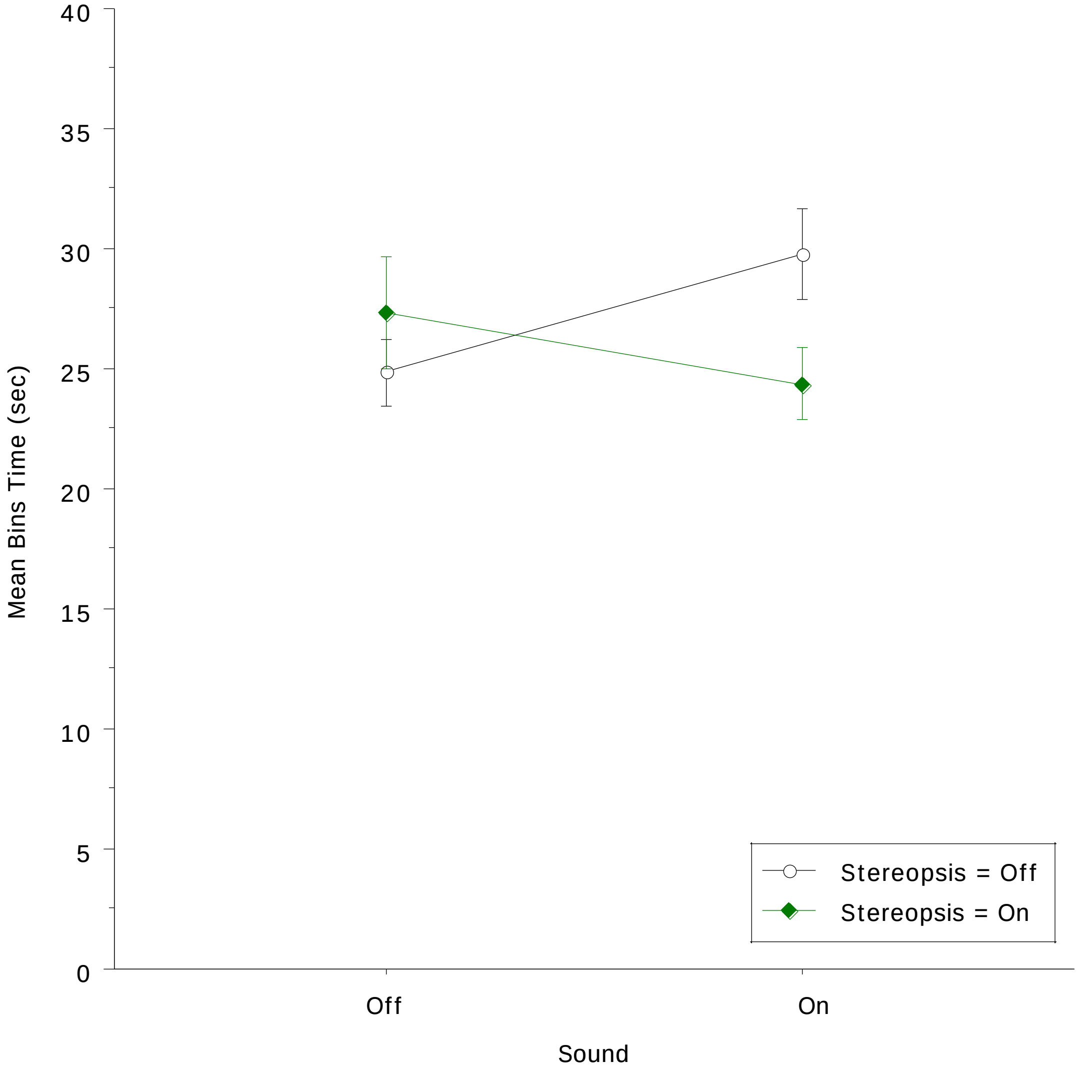


Figure 39. Effect of interaction between sound and stereopsis on time to complete the Bins task.

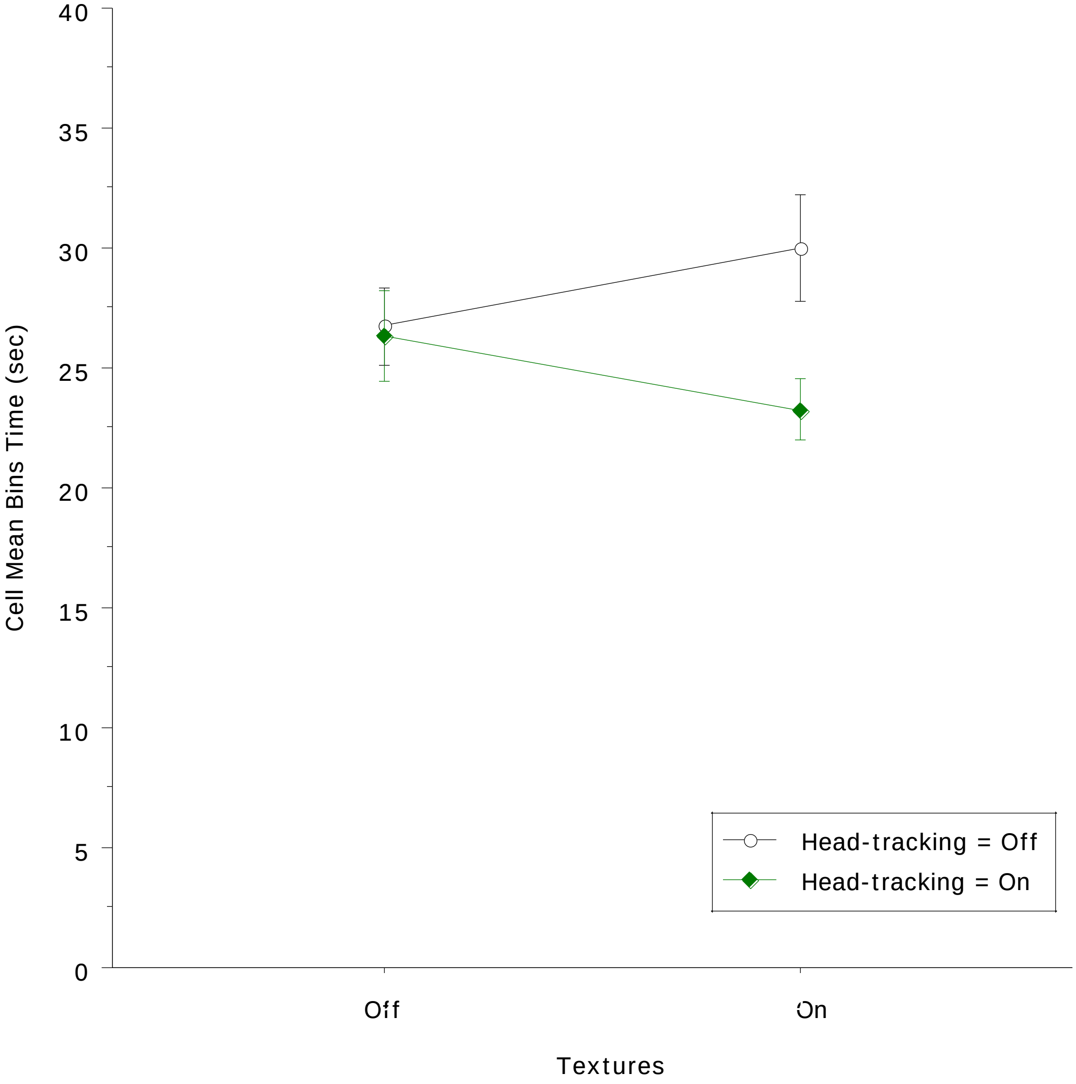


Figure 40. Effect of interaction between textures and head-tracking on time to complete the Bins task.

The ANOVA summary table for the effects of these same five variables on errors made in completing the Turns task is shown in Table 50. The effects of head-tracking, virtual personal risk, and the interaction between sound and textures were statistically significant with $\alpha \leq 0.05$. The effects of head-tracking and virtual personal risk are shown in Table 20. The interaction between sound and textures is shown in Figure 41.

Table 20. Effect of head-tracking and virtual personal risk on errors in the Turns task.

	Off		On	
Effect	Mean	Standard Error	Mean	Standard Error
Head-tracking	60.04	3.66	44.01	3.86
Virtual personal risk	47.06	3.89	56.99	6.17

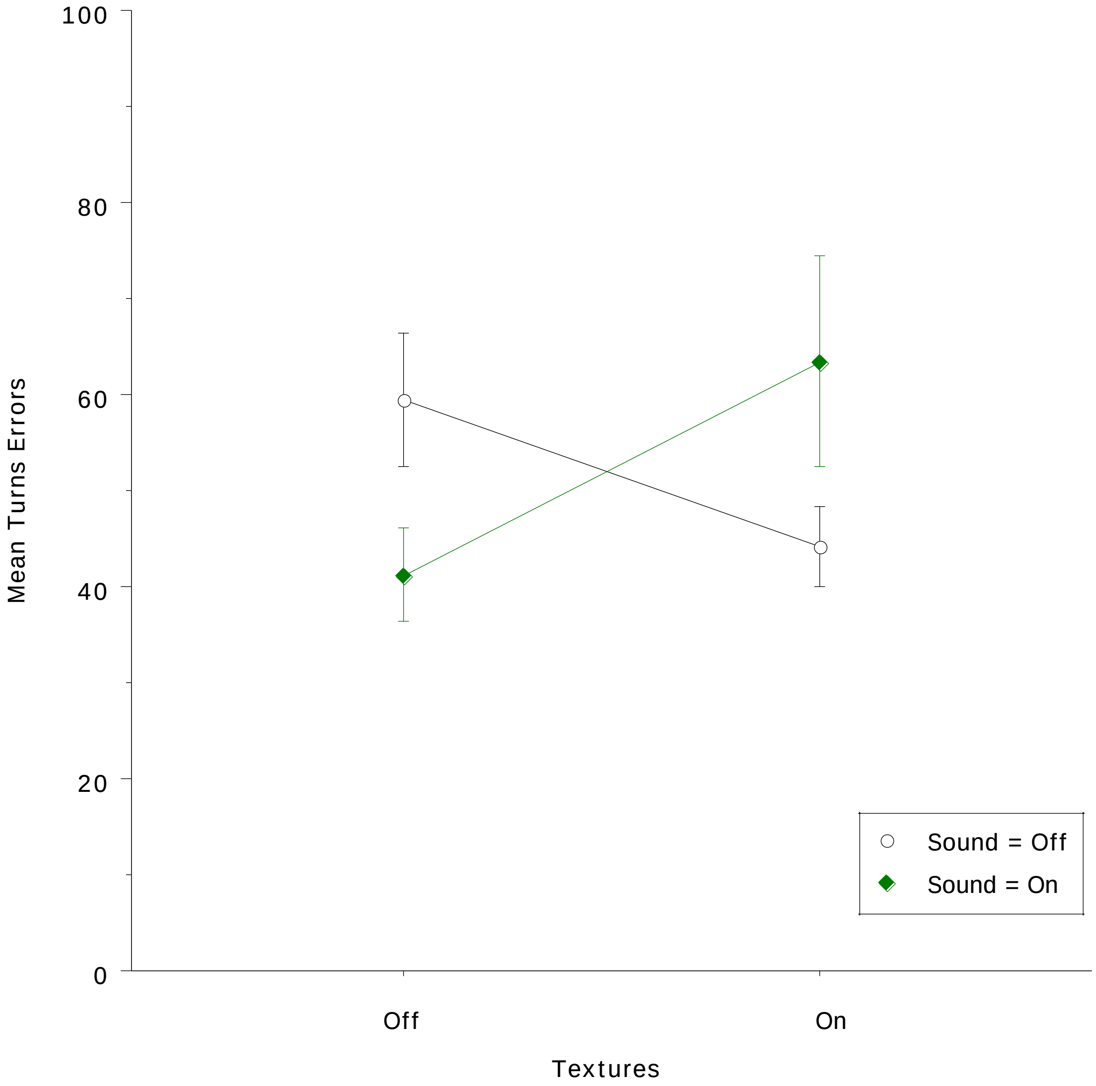


Figure 41. The effect of the interaction between sound and textures on errors made during the Turns task.

The ANOVA performed on time to complete the Turns task is summarized in Table 51. The only statistically significant effect with $\alpha \leq 0.05$ was that of head-tracking. The mean time to complete the Turns task with head-tracking off was 44.39 seconds with a standard error of 2.19 seconds. The mean time to complete the Turns task with head-tracking on was 37.35 seconds with a standard error of 1.28 seconds.

Table 52 shows the ANOVA summary table for the effects of sound, textures, head-tracking, stereopsis, and virtual personal risk on time to complete the Search task. Statistically significant effects (with $\alpha \leq 0.05$) were found for head-tracking and virtual personal risk and are shown in Table 21.

Table 21. Effect of head-tracking and virtual personal risk on time to complete the Search task in Experiment 2.

	Off		On	
Effect	Mean	Standard Error	Mean	Standard Error
Head-tracking	7.36	0.37	8.44	0.45
Virtual personal risk	7.38	0.36	8.42	0.46

The ANOVA summary table for the effects of the independent variables in Experiment 2 on time to complete the Choice task is shown in Table 53. No statistically significant effects ($\alpha \leq 0.05$) were found. No errors were made by any subjects in the completion of this task.

Second-order polynomial regressions were calculated using the independent variables in this experiment — plus time spent in the VE and perceived presence — as regressors to predict each of the dependent variables measuring performance. This was done in the same manner described for Experiment 1 and resulted in the equations shown in Tables 22 through 27.

Table 22. Regression equation predicting distance estimation in Experiment 2.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	*t* for H_0: $\beta = 0$	*p*
x_5	-0.02	0.01	-2.31	0.021
x_6	0.03	0.01	2.35	0.019
x_7	-0.03	0.01	-2.82	0.005
x_8	0.02	0.01	1.46	0.145
x_{14}	1.02	0.01	81.59	0.0001
x_{12}	-0.06	0.01	-5.68	0.0001
x_{13}	-0.06	0.01	-4.31	0.0001
x_{12}^2	0.03	0.01	2.97	0.003
x_{13}^2	0.04	0.01	6.10	0.0001
x_4x_5	-0.02	0.01	-1.71	0.088
x_4x_6	-0.01	0.01	-1.37	0.171
x_4x_7	-0.03	0.01	-3.07	0.002
x_5x_6	0.02	0.01	1.89	0.059
x_5x_7	-0.02	0.01	-1.93	0.054
x_6x_7	-0.02	0.01	-1.89	0.059
x_{14}^2	-0.09	0.01	-8.38	0.0001

$MS_E = 0.079$, PRESS = 61.85, $R^2 = 0.923$

$F_{16,\ 752}$ for H_0: No linear relationship = 560.86, $p < 0.0001$

Table 23. Regression equation predicting time to complete the Bins task in Experiment 2.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_6	-0.17	0.07	-2.44	0.016
x_8	-0.11	0.07	-1.58	0.116
x_{12}	-0.22	0.07	-3.25	0.001
x_4x_5	-0.11	0.07	-1.56	0.120
x_4x_7	-0.15	0.07	-2.16	0.032
x_5x_6	-0.13	0.07	-1.89	0.061
x_{13}^2	0.09	0.03	2.69	0.008

$MS_E = 0.870$, PRESS = 173.71, $R^2 = 0.158$

$F_{7,\ 185}$ for H_0: No linear relationship = 4.95, $p < 0.0001$

Table 24. Regression equation predicting time to complete the Turns task in Experiment 2.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_4	0.16	0.08	1.94	0.054
x_6	-0.15	0.07	-2.01	0.046
x_{12}	-0.25	0.07	-3.62	0.0004
x_{13}	-0.14	0.09	-1.60	0.112
x_4x_5	0.09	0.07	1.29	0.199
x_4x_6	-0.11	0.07	-1.57	0.119

$MS_E = 0.867$, PRESS = 171.96, $R^2 = 0.156$

$F_{6,\ 186}$ for H_0: No linear relationship = 5.73, $p < 0.0001$

Table 25. Regression equation predicting errors in the Turns task in Experiment

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_6	-0.14	0.07	-1.99	0.048
x_8	0.10	0.07	1.46	0.145
x_{12}	-0.22	0.07	-3.16	0.002
x_4x_5	0.16	0.07	2.39	0.018
x_4x_7	0.12	0.07	1.83	0.069
x_5x_7	0.10	0.07	1.54	0.126
x_{12}^2	0.10	0.05	2.03	0.044
x_{13}^2	-0.08	0.03	-2.30	0.022

$MS_E = 0.870$, PRESS = 174.69, $R^2 = 0.163$

$F_{8,\ 184}$ for H_0: No linear relationship = 4.47, $p < 0.0001$

Table 26. Regression equation predicting time to complete the Search task in Experiment 2.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_4	0.12	0.08	1.50	0.134
x_6	0.19	0.08	2.45	0.015
x_8	0.13	0.07	1.90	0.059
x_{13}	-0.13	0.09	-1.50	0.136
x_5x_6	0.09	0.07	1.29	0.199
x_5x_7	0.14	0.07	2.01	0.046
x_{12}^2	0.07	0.05	1.48	0.140

$MS_E = 0.942$, PRESS = 188.88, $R^2 = 0.087$

$F_{7,\ 185}$ for H_0: No linear relationship = 2.53, $p < 0.016$

Table 27. Regression equation predicting time to complete the Choice task in Experiment 2.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_5	-0.16	0.08	-1.92	0.057
x_6	-0.10	0.07	-1.44	0.152
x_7	-0.22	0.08	-2.75	0.007
x_{12}	-0.19	0.07	-2.59	0.011
x_{13}	0.22	0.11	1.94	0.054
x_5x_6	0.12	0.07	1.63	0.105
x_6x_7	0.09	0.07	1.24	0.216
x_6x_8	-0.10	0.07	-1.38	0.169
x_7x_8	0.10	0.07	1.39	0.170
x_{12}^2	0.08	0.05	1.43	0.154
x_{13}^2	-0.07	0.04	-1.47	0.144

$MS_E = 0.925$, PRESS = 186.91, $R^2 = 0.123$

$F_{11, 181}$ for H_0: No linear relationship = 2.31, $p < 0.011$

Presence

The ANOVA summary table for the effects of sound, textures, head-tracking, stereopsis, and virtual personal risk on presence is shown in Table 54. The main effects of sound, textures, head-tracking, and stereopsis were statistically significant ($\alpha \leq 0.05$). These effects are shown in Table 28. Virtual personal risk did not have a statistically significant effect, nor did any of the two-way interactions.

Table 28. Effects of sound, textures, head-tracking, and stereopsis on presence in Experiment 2.

	Off		On	
Effect	Mean	Standard Error	Mean	Standard Error
Sound	7.47	0.42	13.73	0.74
Textures	9.79	0.61	11.42	0.73
Head-tracking	8.12	0.54	13.08	0.71
Stereopsis	9.65	0.65	11.56	0.70

A second-order polynomial regression was calculated with presence as the predicted variable and the independent variables in Experiment 2 (plus minutes in the VE) as regressors. The method and criteria used in calculating the regression were the same as reported for Experiment 1. The equation shown in Table 29 was the result.

Table 29. Regression equation predicting perceived presence in Experiment 2.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_4	0.47	0.06	8.40	0.0001
x_5	0.12	0.06	2.18	0.031
x_6	0.37	0.06	6.65	0.0001
x_7	0.14	0.06	2.56	0.011
x_4x_6	0.09	0.06	1.64	0.103
$MS_E = 0.605$, PRESS = 119.27, $R^2 = 0.408$				
$F_{5,\ 1871}$ for H_0: No linear relationship = 25.74, $p < 0.0001$				

Experiment 3

Performance

As for the previous two experiments, analyses of variance (ANOVAs) were conducted for each dependent and independent variable. The ANOVA summary tables for the effects of interactions, presence of a second user, and environmental detail on all performance measures are shown in Tables 55 through 60. As shown in these tables, there were no statistically significant effects of these three variables or their interactions on any performance measure. No incorrect choices were made during performance of the Choice task in Experiment 3.

Second-order polynomial regressions were calculated using the independent variables in this experiment — again including time spent in the VE and perceived presence — as regressors to predict each of the dependent variables that measured performance. The method and criteria used in calculating the regression were again the same as reported for Experiment 1. This resulted in the equations shown in Tables 30 through 35.

Table 30. Regression equation predicting distance estimation in Experiment 3.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_9	0.02	0.01	1.91	0.057
x_{10}	0.02	0.01	2.79	0.005
x_{14}	1.06	0.01	94.46	0.0001
x_{13}	-0.06	0.01	-6.76	0.0001
x_9^2	0.04	0.01	3.79	0.0002
x_{11}^2	0.05	0.01	5.05	0.0001
x_{12}^2	0.02	0.004	4.39	0.0001
x_{13}^2	0.03	0.005	5.10	0.0001
x_{14}^2	-0.16	0.01	-14.51	0.0001

$MS_E = 0.065$, PRESS = 56.99, $R^2 = 0.935$

$F_{9,\,855}$ for H_0: No linear relationship = 1375.06, $p < 0.0001$

Table 31. Regression equation predicting time to complete the Bins task in Experiment 3.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_{12}	-0.19	0.08	-2.28	0.024
x_{13}	-0.13	0.07	-1.75	0.082
x_{12}^2	0.07	0.04	2.14	0.033

$MS_E = 0.943$, PRESS = 205.99, $R^2 = 0.066$

$F_{3,\,313}$ for H_0: No linear relationship = 4.98, $p < 0.002$

Table 32. Regression equation predicting time to complete the Turns task in Experiment 3.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_{11}	-0.11	0.07	-1.63	0.105
x_{13}	-0.25	0.07	-3.76	0.0002
x_{11}^2	-0.10	0.06	-1.68	0.094
x_{12}^2	0.08	0.03	2.59	0.010
x_{13}^2	0.08	0.04	1.99	0.048

$MS_E = 0.896$, PRESS = 197.67, $R^2 = 0.121$

$F_{5, 211}$ for H_0: No linear relationship = 5.79, $p < 0.0001$

Table 33. Regression equation predicting errors in completion of the Turns task in Experiment 3.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_{12}	-0.32	0.08	-3.77	0.0002
x_{13}	-0.10	0.07	-1.43	0.153
x_{12}^2	0.09	0.04	2.50	0.013
x_{13}^2	-0.06	0.04	-1.74	0.083

$MS_E = 0.898$, PRESS = 196.21, $R^2 = 0.115$

$F_{4, 212}$ for H_0: No linear relationship = 6.86, $p < 0.0001$

Table 34. Regression equation predicting time to complete the Search task in Experiment 3.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_{13}	-0.19	0.07	-2.81	0.005
$x_{10}x_{11}$	-0.09	0.07	-1.37	0.173

$MS_E = 0.960$, PRESS = 210.16, $R^2 = 0.044$

$F_{2,\ 214}$ for H_0: No linear relationship = 4.98, $p < 0.008$

Table 35. Regression equation predicting time to complete the Choice task in Experiment 3.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_{11}	0.09	0.07	1.35	0.179
x_{12}	-0.15	0.07	-2.24	0.026

$MS_E = 0.976$, PRESS = 212.46, $R^2 = 0.028$

$F_{2,\ 214}$ for H_0: No linear relationship = 3.12, $p < 0.046$

Presence

The ANOVA summary table for the effects of interactions, presence of a second user, and environmental detail on presence is shown in Table 61. The presence of a second user and the interaction between environmental detail and number of interactions were both statistically significant with $\alpha \leq 0.05$. With the second user absent, mean presence was 13.29 with a standard error of 0.39. With the second user present, mean presence was 14.58 with a standard error of 0.34. The effect of the interaction between environmental detail and number of interactions is displayed in Figure 42.

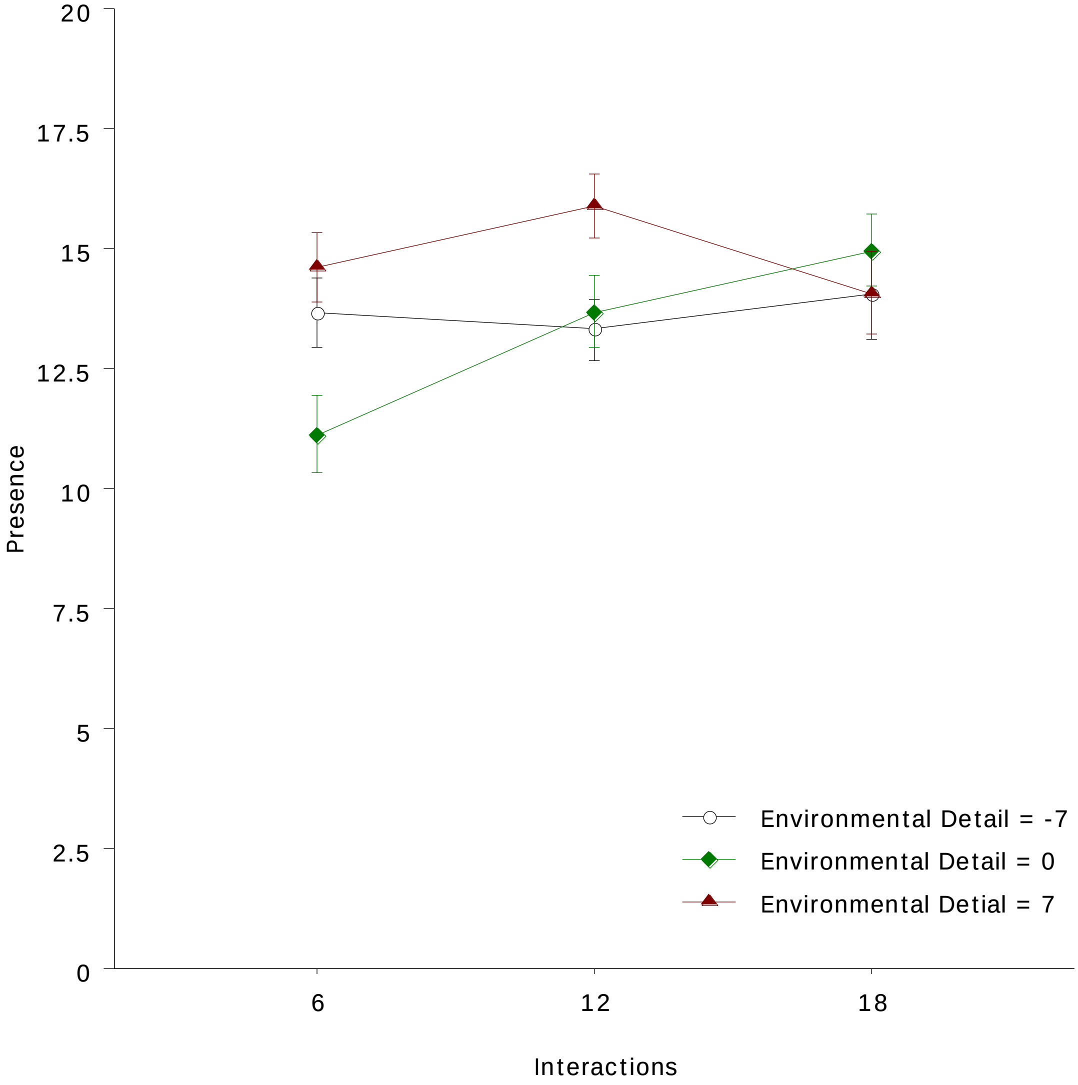


Figure 42. Effect of the interaction between environmental detail and number of interactions on presence.

Again, a second-order polynomial regression was calculated with presence as the predicted variable and with the independent variables in Experiment 3 (plus minutes in the VE) as regressors. The method and criteria used in calculating the regression were the same as reported for Experiment 1. This resulted in the model contained in Table 36.

Table 36. Regression equation predicting perceived presence in Experiment 3.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_9	0.13	0.06	2.10	0.037
x_{10}	0.20	0.06	3.27	0.001
x_{12}	0.51	0.07	7.36	0.0001
x_{11}^2	0.11	0.05	2.03	0.044
x_{12}^2	-0.09	0.03	-2.76	0.006

$MS_E = 0.767$, PRESS = 169.24, $R^2 = 0.247$

$F_{5,\,211}$ for H_0: No linear relationship = 13.86, $p < 0.0001$

Data Bridging

The experimental series was conducted under the constraints of a sequential experimentation paradigm. These constraints are:

- All independent variables must be defined prior to data collection.
- Experimental procedures and dependent measures must be constant across all experiments in the sequence.
- Information must be available at the outset concerning the levels of all factors manipulated and held constant.
- To the extent possible, all main and pure quadratic effects of each independent variable must be tested.

In addition, it has been recommended that the experiments in the sequence have at least one data point in common to provide a direct estimate of the comparability of the studies. This recommendation was made by Williges, Williges, and Han in 1992, but the current series of experiments was the first to implement it. Satisfaction of these constraints allows for the possibility of bridging the data set across experiments for the purpose of constructing second-order empirical models of the effects of the variables manipulated. It has been argued that a second-order polynomial approximation is adequate to account for most human performance (Williges, 1981). This assumption allows much greater economy of data collection than experimental designs that test third- and higher-order interactions.

Unfortunately, even if the above constraints are satisfied, there is no guarantee at the outset of data collection that data bridging of the resulting data set will be possible. This was perhaps especially true in the present case given the subjective nature of the primary dependent variable. As it turns out, data bridging across all three experiments in the current series was not feasible. It was originally intended that a fourth experiment be conducted to collect data needed to model two-way interactions not examined in the first three experiments. Specifically, it was hoped that a fourth experiment would provide the missing data needed to model the effects on perceived presence of all possible two-way interactions among the eleven independent variables manipulated. An examination of the experimental designs of the first three experiments reveals that, to model many of these interactions, a critical fourth data point is needed. For example, in the second experiment, sound and textures are completely crossed. However, sound is never crossed with any of the variables in the first or third experiments: sound was always on during these experiments. To examine the interaction between sound and visual display resolution, for instance, a data point is needed in which sound is off and visual display resolution is low. Such a data point does not exist in the first three experiments.

Detailed procedures for collecting such data points and bridging data across experiments in a sequential experimentation paradigm have been described in a number of papers by Williges, Williges, and Han (Han, 1991a; Han, 1991b; Han, Williges, and Williges, 1990; Williges and Williges, 1989; Williges, Williges, and Han, 1992; Williges, Williges, and Han, 1993). However, prior to implementing these procedures, it must first be determined — if possible — that the data gathered in all experiments in such a paradigm can be treated as one data set. Han (1991b) proposes two ways of doing this. In the first, each experiment is entered into the empirical modeling process as an indicator variable and the significance of the interactions between this indicator variable and the variables under study is tested. A significant interaction indicates that the data cannot be treated as one dataset and that data bridging procedures are not feasible. In the second, the experimental series is designed so that one or more data points are common to all experiments. If these data points show differences across experiments, data bridging then becomes problematic.

While Han’s first suggestion is technically sound, it seems useful only when data bridging is done in a post hoc fashion on experiments not preplanned as part of a sequential experimentation effort. In a designed series of studies, it is much more efficient to design in a common data point than it is to include the multiple levels of each variable in all experiments that would be needed to test experiment by variable interactions. In the three experiments reported above, each independent variable was manipulated in one and only one experiment. It is therefore impossible to test an interaction between any manipulated variable and the experiment indicator variable. Han’s second suggestion, though, is well-suited to designed sequential experimentation and was used as the basis for the data bridging decisions made in the current research.

Given that perceived presence was the primary focus of the research and empirical modeling effort conducted, this variable was also the primary basis for decisions regarding data bridging. Figure 43 shows the effect of experiment on perceived presence. This figure shows only the data from the common data point included in each experiment (i.e., one observation from each subject). It should be noted that — in this figure and for data bridging purposes only — the raw presence data from the three experiments were adjusted to a modulus common to the three experiments (the value of this modulus was 0.95).

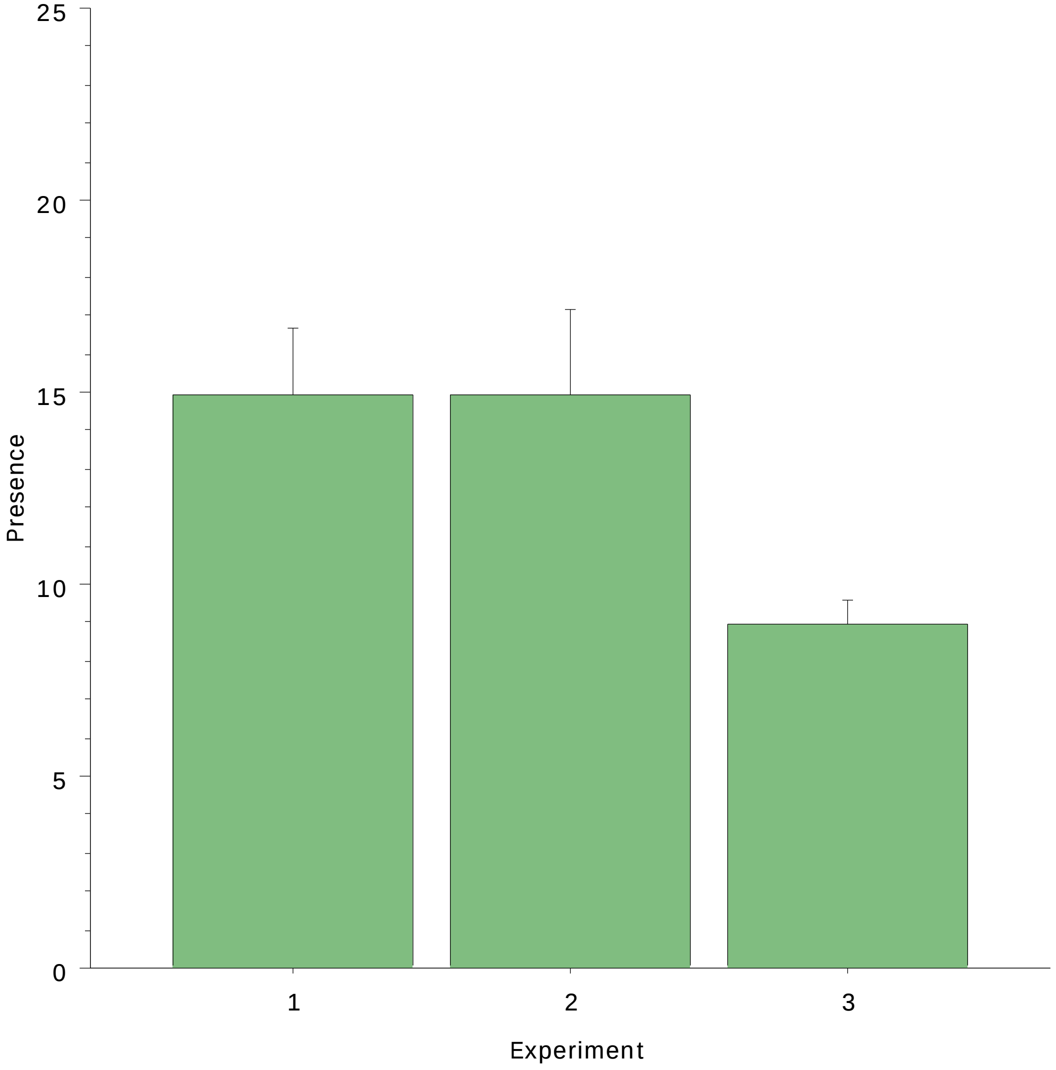


Figure 43. Effect of experiment on presence, common data point only.

It is apparent from this figure that the three experiments were *not* equivalent in terms of perceived presence. This conclusion is supported by a statistical test (an ANOVA) rejecting the hypothesis that experiment had no effect on perceived presence in the condition common to the three experiments ($F_{2, 33} = 4.46$, $p < 0.019$). Figure 43 shows that, not only was the mean perceived presence in Experiment 3 significantly lower than in Experiments 1 and 2, the variability of presence estimates in Experiment 3 (indicated by the standard error bars in this figure) was substantially smaller in this experiment than the other two. Scheffe’s multiple contrast procedure was used to verify the statistical significance of the difference apparent in Figure 43. The result, indicating that this difference was significant, is shown in Table 37.

Table 37. Test for differences among experiments using Scheffe's multiple contrast procedure and the perceived presence common data point.

Experiments	Mean Difference	Critical Difference	*p*
1 and 2	-0.034	5.914	0.999
1 and 3	5.951	5.914	0.048
2 and 3	5.985	5.914	0.047

It was felt that, even if a constant were added to the observations in Experiment 3, it still could not be said with confidence that all data collected represented a uniform data set sampled from the same population. Given this conclusion, the question remained of whether or not a fourth experiment should be conducted to collect data needed to examine two-way interactions among variables manipulated in the first two experiments only. Given the notable dearth of significant two-way interactions already found in these two experiments, and given that these experiments were (at least in part) designed to examine those two-way interactions most likely to be significant and important to the VE design community (e.g., visual display resolution and field of view), it was decided that further time, effort, and money spent in data collection would be counterproductive.

For these reasons, a polynomial model including all eleven manipulated variables and time spent in the VE was not constructed. However, the data from the first two experiments were combined and used to perform a polynomial regression predicting perceived presence from the eight variables manipulated in these experiments plus minutes spent in the VE. The method and criteria used in calculating the regression were the same as reported for Experiment 1. The following empirical model (seen in Table 38) was the result:

Table 38. Regression equation predicting perceived presence from the data collected in Experiments 1 and 2.

Variable	$\hat{\beta}$	Standard Error of $\hat{\beta}$	t for H_0: $\beta = 0$	p
x_2	0.13	0.04	3.09	0.002
x_3	0.52	0.05	11.30	0.0001
x_4	0.47	0.05	9.34	0.0001
x_5	0.15	0.05	3.15	0.002
x_6	0.40	0.05	8.25	0.0001
x_7	0.12	0.04	2.63	0.009
x_{12}	0.15	0.04	3.96	0.0001
x_{12}^2	-0.08	0.03	-2.72	0.007
x_1^2	0.06	0.03	2.17	0.031
x_4x_5	0.06	0.04	1.53	0.127
x_4x_6	0.09	0.04	2.32	0.021

$MS_E = 0.576$, PRESS = 241.40, $R^2 = 0.438$

$F_{11,\ 397}$ for H_0: No linear relationship = 28.13, $p < 0.0001$

Again, a fourth, data-bridging experiment was not conducted and so this model does not contain terms representing interactions among variables not manipulated in the same experiment. For purposes of this regression, the presence estimates from Experiments 1 and 2 were adjusted to a modulus common to these two experiments. The value of this modulus was 0.86.

Results Summary

Given the number of independent and dependent variables involved in this research, it was felt that a summary of the results would be helpful with respect to the two main focuses of the research: 1) the impact of independent variables on perceived presence, and, 2) the relationship between presence and performance. Table 39 summarizes the effects of the independent variables manipulated and their two-way interactions on perceived presence based on the results of the ANOVAs conducted. The second order effects and minutes spent in the VE are added based upon the regression analyses performed. Where possible (i.e., for significant main effects) the percent increase in perceived presence is indicated. This is the percent change associated with the difference between the low and high conditions for each variable. Where a percentage is not given, a ‘+’ indicates that the effect was significant (with $\alpha \leq 0.05$), while a ‘—’ indicates that it was not. Significance was tested by ANOVAs in the case of main effects and two-way interactions, and by t-tests on the parameter estimates in the empirical models in the case of second-order effects and minutes spent in the VE.

Table 39. Summary of results with respect to perceived presence.

Variable / Effect	Impact on Perceived Presence
Experiment 1	
Scene update rate	14%
Visual display resolution	23%
Field of view	148%
Scene update rate ^{2}	—
Field of view ^{2}	—
Scene update rate X Visual display resolution	—
Scene update X Field of view	—
Visual display resolution X Field of view	—
Minutes in the VE	+
Minutes in the VE ^{2}	+
Experiment 2	
Sound	83%
Textures	17%
Head-tracking	61%
Stereopsis	20%
Virtual personal Risk	—
Sound X Textures	—
Sound X Head-tracking	—
Sound X Stereopsis	—
Sound X Virtual personal risk	—
Textures X Head-tracking	—

Textures X Stereopsis	—
Textures X Virtual personal risk	—
Head-tracking X Stereopsis	—
Head-tracking X Virtual personal risk	—
Stereopsis X Virtual personal risk	—
Sound X Textures	—
Minutes in the VE	—
Minutes in the VE ^{2}	—
Experiment 3	
Interactions	—
Second user	10%
Environmental detail	—
Interactions ^{2}	—
Environmental detail ^{2}	+
Interactions X Second user	—
Interactions X Environmental detail	+
Second user X Environmental detail	—
Minutes in the VE	+
Minutes in the VE ^{2}	+

Table 40 contains a summary of the relationship between presence and performance. This table contains correlations between perceived presence and performance for each task in each experiment. The statistical significance of these correlations was tested using Fisher's r to Z transformation (Hays, 1981).

Table 40. Correlations between perceived presence and performance.

Dependent Variable	*r*	*z*	*p*
Experiment 1			
Time to complete the Bins task	0.055	0.803	0.422
Time to complete the Turns task	-0.192	-2.837	0.005
Errors made in completing the Turns task	-0.180	-2.652	0.008
Time to complete the Search task	-0.292	-4.381	0.0001
Time to complete the Choice task	-0.083	-1.207	0.227
Experiment 2			
Time to complete the Bins task	0.009	0.122	0.903
Time to complete the Turns task	-0.124	-1.712	0.087
Errors made in completing the Turns task	-0.162	-2.241	0.025
Time to complete the Search task	-0.003	-0.046	0.963
Time to complete the Choice task	-0.029	-0.400	0.689
Experiment 3			
Time to complete the Bins task	-0.191	-2.824	0.005
Time to complete the Turns task	-0.239	-3.560	0.0004
Errors made in completing the Turns task	-0.230	-3.417	0.001
Time to complete the Search task	-0.190	-2.807	0.005
Time to complete the Choice task	-0.116	-1.696	0.090

DISCUSSION

The following discussion is subdivided into sections discussing the results with respect to performance, presence, and sequential experimentation and data bridging. Finally, implications for future results are discussed and a summary is given of the conclusions that may be drawn from this research.

Performance

In general, the results of this research with respect to performance measures are as expected and consistent with literature examining similar tasks. However, many combinations of variables and tasks had not been previously examined. Perhaps the most interesting finding is the lack of effect the manipulated variables had on some tasks. This could be due to the simplicity of the tasks and these results may not generalize to more complex and/or demanding tasks or applications such as entertainment or flight simulation. Also, participants were not given extensive experience in the VE prior to performance measurement. This was so that the effects of time spent in the VE on perceived presence could be tested. However, possible practice effects should be taken into account in interpreting these findings. Finally, in discussing these results — especially with respect to regression and empirical modeling — it should be noted that they can only be discussed in the context of the parameter ranges tested. Extrapolation of these findings to variable ranges not tested is tenuous at best and could be, at worst, quite misleading. For example, given that the maximum field of view tested in these studies was 48°, one cannot make predictions from these data and analyses that would be valid for a full-field-of-view (i.e., 200°) display.

Distance Estimation

In performing the distance estimation task, participants seemed to invariably — even given extensive practice — overestimate distance at the middle distances. This effect can be seen in the concave shape of the data in all figures portraying distance estimation data. The results of this research indicate that this effect is increased by increasing the resolution of the visual display and decreasing its field of view. Conversely, the inclusion of texture-mapping and stereopsis seemed to cause participants to decrease their estimates at all distances, moreso for stereopsis than for texture-mapping. The statistically significant interactions found in the second experiment involving sound are puzzling: the only sounds present during this task were the ticking of the wall clock to participants' left and the feedback click the response buttons made as they were pressed. Replication of these results would seem advisable. It should be noted with caution that distance estimation was done from a stationary point of view: the fact that no effect of head-tracking was found in this task could well be due to the removal of motion parallax as a depth cue. The empirical models of distance estimation performance indicate that, while the variables manipulated in these studies did have significant effects on distance estimation, by far the most influential factor in predicting estimated distance is (as one perhaps would expect) actual distance.

Object Manipulation

One surprising result was the lack of a significant effect of scene update rate on performance of the Bins task, an object manipulation task. Based on the literature relating time delay to performance in object manipulation tasks (see Held and Durlach, 1991), one would expect a large, statistically and practically significant effect of scene update rate in this task. Some of the possible reasons that this effect was not found are that scene update rate is not strictly equivalent to lag as it has traditionally been investigated, that the minimum scene update rate (8 Hz) was not slow enough to generate significant performance decrements in this task, or that the variability of performance on this task was too great for a statistically significant effect to reveal itself. Indeed, no significant effects on performance of this task were found for any of the variables (or their interactions) manipulated in Experiments 1 and 3. The regression equations predicting performance in this task in these two experiments show a similar lack of predictive value ($R^2 < 0.07$ in both cases). What little predictive capability is contained in these equations seems to stem from experience with the task (represented by the parameter estimate associated with the term for minutes spent in the VE). While significant effects were found in Experiment 2 for this task, the pattern of these effects and possible reasons for them seem unclear. Again, the predictive value of the empirical model predicting performance in this task was low ($R^2 = 0.158$).

Locomotion

Locomotion in the VE, as measured by performance on the Turns task, was affected by scene update rate, field of view, head-tracking, virtual personal risk, and the interaction between sound and textures. The pattern of the results for this task with regard to errors and completion time indicates that these effects were not the result of a speed-accuracy tradeoff. Not surprisingly, performance on the task improved with increased scene update rate and field of view. Head-tracking also had a beneficial effect on participants' ability to move in the VE. The statistical significance of the effect of virtual personal risk on errors made in completion of the Turns task is difficult to interpret. Participants did not have contact with the manipulation of virtual personal risk (the absence of the back doors in the elevator) until well after completion of this task. The absolute magnitude of this effect is relatively small and replication would seem in order before VE design decisions are made based upon this result. The effect on errors made during the Turns task of the interaction between sound and textures is interesting: it seems to indicate that one only had a positive effect on performance in the absence of the other. This could be an effect of the amount of feedback participants received during this task. With sound on, they could hear themselves bumping into the walls. With textures on, they could see changes in optical flow due to contact with walls, especially when very close to them. With neither of these features active, feedback, especially with regard to errors (i.e., contacts with walls), was reduced. The predictive value of the regressions modeling time and errors in this task was generally low ($R^2 < 0.25$). Again, judging from the parameter estimates found in these equations, this predictive value was primarily the result of practice, or minutes spent in the VE.

Visual Search

Much literature exists showing that the time needed to perform a visual search is negatively related to the field of view of the person performing the search. The results of Experiment 1 essentially replicate this finding. Interestingly, visual display resolution also had an effect on performance of this task and interacted with field of view. These effects appear to be caused in part by the large increase in search time associated with the low-field-of-view, low-resolution condition. Sheer lack of information or bandwidth might have been an issue. The visual display in this condition consisted of only 160 x 100 pixels. The presence of head-tracking increased search time, probably due to the fact that participants were had to physically shift or rotate their real bodies (often several times) to view the entire room. With head-tracking turned off, the entire room could be searched simply by twisting the 3D controller. Again the statistically significant effect of virtual personal risk on performance of this task is puzzling: while they were told at the start of each trial (if they asked) whether or not virtual personal risk would be present, participants did not experience it until after the Search task was complete. None of the variables manipulated in Experiment 3 affected search time. This is somewhat unexpected. The second user, when present, attempted to assist participants by informing them of the location of the ball (e.g., "It's on the floor to your left.”). However, participants had often already found the ball or had already begun an established search pattern before assistance could be given. Except for the first experiment, in which field of view and visual display resolution played a key role, the variables manipulated in these studies seemed to have little effect on predicted search time in the models calculated ($R^2 < 0.10$). Even with the inclusion of visual display resolution and field of view in the first experiment, the resulting empirical model accounted for less than 20% of the variance in the search time data.

Choice Reaction Time

It is remarkable that not a single error was committed on the Choice task in all three experiments. Instructions given to participants were not designed to bias them in favor of either speed or accuracy. None of the variables manipulated had a statistically significant effect on performance of this task except scene update rate, which caused choice reaction time to decrease as it was increased. This effect could be a simple artifact of the fact that participants’ responses in this task were polled twice as often and twice as fast in the high scene update rate condition as in the low scene update rate condition. Scene update rate was defined as the rate at which the entire VE updated. This included all code managing interaction with the VE, which included the interaction represented by participants’ responses in the Choice task. The empirical models of performance of this task, again, generally predicted a relatively small percentage of the variance in the data.

Comparison to Lampton, Knerr, et al.

The concave functions found in the distance estimation task were strikingly similar to the results reported by Lampton, Knerr, et al. (1994). Other than this, there seem to be few comparisons between the results of this research and those of Lampton, Knerr, et al. (1994). The absolute values of time to complete the Bins task (in seconds) were comparable. The focus of these authors was primarily on the differences between input devices. Based on the results of the three experiments conducted in the course of this research, this set of tasks also seems adequate for the purpose of testing the effects of other VE parameters with respect to both perceived presence and task performance. In some cases, though, more demanding or complex tasks might be better suited to this purpose.

Presence

The results with respect to perceived presence seem to vindicate the use of magnitude estimation as a dependent variable useful in measuring this phenomenon. Adjusted to a common modulus, participants’ free-modulus magnitude estimates show clear and systematic effects of the variables manipulated in this research. This is borne out by the regression analyses performed which, in most cases, account for over 40% of the variance in the presence data. The results of this research show that field of view, sound, and head-tracking had the largest impacts on perceived presence. Smaller, but still significant effects — in both a statistical and a practical sense — were found for visual display resolution, texture-mapping, stereopsis, and the presence of a second user. Finally, small, but statistically significant effects were found for scene update rate, and the interaction between environmental detail and the number of interactions possible in the VE. These trends are supported by the results of the regression analyses conducted: specifically, a comparison of the standardized parameter estimates associated with all of these effects in the empirical models reported, especially the model for Experiments 1 and 2 combined. These regressions show that time spent in the VE was also a major determinant of perceived presence.

Of the theories and schema proposed to account for presence in VEs and reviewed in the introduction, perhaps the one that best fits these data is Steuer's (1992). His ideas and predictions concerning vividness and interactivity generally match the findings of this research. The large effects of field of view and sound nicely fit the concepts of sensory depth and sensory breadth encompassed within vividness, as do the effects of visual display resolution, texture-mapping, and stereopsis. The effects of head-tracking, scene update rate, and presence of a second user seem to match his conception of interactivity. The one finding of this research that seems inconsistent with his theory (and others) is that perceived presence was not affected by the number of interactions possible in the VE in the third experiment. This could be explained by some participants focusing strictly on task performance and not exploring possible interactions, by the unnaturalness of interacting with the VE through the use of a 2D mouse, or it may be that the number of possible interactions in a VE really has little or no effect on perceived presence. One common theme among many theories is that time spent in the VE or adaptation to the VE causes an increase in the experience of presence. This seems consistent with the results of this research.

The relationship between perceived presence and performance seems real, but weak. A number of observations may be made regarding the role of presence and its second-order effect in the empirical models that included these terms as regressors. While one or both of these terms are usually found in the final empirical model predicting performance, the standardized parameter estimates associated with them are sometimes small relative to the other included parameters (e.g., the models for errors made in completion of the Turns task and time to complete the Choice task in Experiment 1). Sometimes the empirical model itself had little predictive value (e.g., the Bins task models for Experiments 1 and 3 and the Search task model for Experiment 3). There were a few cases in which the parameter estimate associated with presence or its second-order effect indicated a dominant or significant influence of these terms on task performance (e.g., performance of the Search task in Experiment 1), but this dominance seems infrequent and not pervasive. With regard to correlations between presence and performance, the results indicate that there is a positive relationship between presence and performance. Correlations between performance improvement and perceived presence were positive (when they were statistically significant). However, while the majority of these correlations were statistically significant, their absolute value was low — usually in the range of 0.05 to 0.30. This implies that, in all cases, over 90% of the variance in the performance data was accounted for by something other than perceived presence.

Sequential Experimentation and Data Bridging

This research seemed to prove the value of a common data point in sequential experimentation designs. Without this common data point, there would have been no means of testing the assumption that the separate data sets resulting from each experiment could be combined into one. Given the outcome of this research, it seems especially important to design in a common data point in experimental sequences in which a subjective measure is the primary dependent variable.

The reasons for the difference between perceived presence in the third experiment and perceived presence in the other two experiments remains somewhat unclear. The experimental procedure, apparatus, and the experimenter himself were all identical. Post-hoc explorations of the sample populations with respect to age and gender revealed no discernible differences. It was hypothesized that, in spite of the random ordering of trials, the common data point trial in the third experiment might have happened to fall, on average, earlier in the experimental session than in the other two experiments. However, such was not the case. Perhaps the likeliest explanation for this difference is the difference between the range of perceived presence in the experiments. In the third experiment, the two effects that showed a significant impact upon perceived presence were relatively small. It may be that participants simply did not notice much difference from trial to trial in this experiment. This is contrasted with the other two experiments in which, given the effects found, it may be surmised that participants typically experienced significant increases or decreases in presence from trial to trial.

Future Research Implications

One of the foremost implications for future research is the possibility of adding the results of further studies to the data set resulting from this research. However, given that all three of the experiments conducted could not be bridged into a single data set, it would seem imperative that such studies include a data point containing the same conditions as those set by the common data point in these experiments. Some variables and design points that might be examined, but could not be in this research because of hardware and software limitations, are presence of a virtual body, natural locomotion (i.e., walking) through the VE, expanded field of view, enhanced visual display resolution, 3D localization and externalization of auditory stimuli, and tactile and force feedback.

The results of Lampton, Knerr, et al. (1994) concerning performance on the distance estimation task have been replicated. It is apparent that participants performing this task overestimate at the middle distances. This effect was not removed by any variable manipulated in these studies. Research is needed to discover why this effect occurs. One important piece of information that needs to be gathered is whether this overestimation is due to the nature of this particular task or whether it occurs in all tasks in which distance estimation is a component. A finding that participants in a VE generally tend to overestimate distances to objects moving toward them would be important indeed.

Given the apparent common wisdom in the VR community that more interactions increase presence, further study and replication of the non-effect of this variable found in Experiment 3 are needed. It seems possible and even likely that, when interactions are not integral to the task performed in the VE, they do not affect participants' sense of presence. It may be that the number of possible interactions in a VE only increases presence in entertainment applications. Alternatively, this variable may have an effect only when highly interactive tasks are performed. Further research is needed to establish under what conditions, if any, the number of interactions possible in a VE affects perceived presence.

Similarly, it has been proposed by a number of authors that the presence of virtual personal risk increases presence. However, no such effect was found in this research. This could mean either that this variable really has no effect or simply that the manipulation used in this research had no effect. A study is needed in which perceived virtual personal risk is directly measured (perhaps using magnitude estimation) together with perceived presence. This seems the only way to reach a definitive conclusion concerning the contribution of this variable to perceived presence.

Individual differences would seem the next frontier in presence research. Differences due to gender, age, computer experience, and other variables may be quite large and should be investigated. The empirical models of perceived presence reported in the present research accounted for less than half the variance in the presence data. While this is a good start, the question remains as to what accounts for the rest: individual differences seem a likely candidate.

Finally, it should be noted that all of the findings reported in this document may apply only to a relatively simple set of laboratory tasks. Similar research investigating more elaborate tasks, more demanding tasks, entertainment applications, and learning environments needs to be conducted.

Summary

With regard to the research hypotheses stated in the introduction, the conclusions one may draw from this research are:

1) That the VR system parameters manipulated and analyzed in this research affected users' subjective feeling of presence in a VR system. Field of view, sound, and head-tracking showed the largest effects. Other significant effects found were those of visual display resolution, texture-mapping, stereopsis, and the presence of a second user.

2) Free-modulus magnitude estimation can be used to measure subjective presence in a VE. Furthermore, this measure can be used to generalize and compare results across users, across experiments, across tasks, and across differently-configured VR systems.

3) A positive relationship exists between a user's subjective feeling of presence in a VE and the user's performance of tasks within that environment, at least for the simple tasks used in this research. While this relationship is positive, it does not seem very strong.

4) Empirical models of the effects of the parameters manipulated were constructed based on the results of this research that predict both presence and performance in a VE based upon the parameters of the VR system in question. The models of presence accounted for over 40% of the variance in the observations. However, in most cases, the predictive value of the models with regard to performance variables was relatively weak . Given that these models are the first of their kind, it is hoped that even the performance models will aid the VE design community to a small extent.

REFERENCES

Allen, J. A., Hays, R. T., and Buffardi, L. C. (1986). Maintenance training simulator fidelity and individual differences in transfer of training. *Human Factors*, *28*(5), 497-509.

Ansley, D. A. (1991). *Effects of color CRT misconvergence and display luminance on reading performance and perceived image quality judgments*. Unpublished masters thesis, Virginia Polytechnic Institute and State University.

Arthur, K. W., and Booth, K. S. (1993). Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems, 11*(3), 239-265.

Atherton, P. R., and Caporael, L. R. (1985). A subjective judgment study of polygon based curved surface imagery. In *CHI '85 Proceedings*, (pp. 27-34). New York: ACM.

Azuma, R. (1993). Tracking requirements for augmented reality. *Communications of the ACM, 36*(7), 50-52.

Bailey, R. E., and Knotts, L. H. (1987). Effect of time delay on manual flight control and flying qualities during in-flight and ground-based simulation. In *AIAA Flight Simulation Technologies Conference*, (pp. 30-38). New York, NY: American Institute for Aeronautics and Astronautics.

Barfield, W., and Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. In G. Salvendy and M. Smith (Eds.), *Human-computer interaction: Software and hardware interfaces* (pp. 699-704). Amsterdam: Elsevier.

Bolanowski, S. J., Jr., Zwislocki, J. J., and Gescheider, G. A. (1991). Intersensory generality of psychological units. In S. J. Bolanowski, Jr. and G. A. Gescheider (Eds.), *Ratio scaling of psychological magnitude: In honor of the memory of S. S. Stevens* (pp. 277-293). Hillsdale, NJ: Lawrence Erlbaum.

Bradley, W. E. (1967). Telefactor control of space operations. *Astronautics & Aeronautics*, May, 32-38.

Carr, K. T., and England, R. D. (1993). The role of realism in virtual reality. In *Virtual Reality International 93: Proceedings of the Third Annual Conference on Virtual Reality*, (pp. 24-33). London, UK: Meckler.

Cha, K., Horch, K. W., and Normann, R. A. (1992). Mobility performance with a pixelized vision system. *Vision Research, 32*(7), 1367-1372.

Chung, J. C. (1992). A comparison of head-tracked and non-head-tracked steering modes in the targeting of radiotherapy treatment beams. In *Proceedings of the 1992 ACM Symposium on Interactive 3D Graphics*, (pp. 193-196).

Collins, A. A., and Gescheider, G. A. (1989). The measurement of loudness in individual children and adults by absolute magnitude estimation and cross-modality matching. *Journal of the Acoustical Society of America, 85,* 2012-2021.

Cross, D. V. (1974). Some technical notes on psychophysical scaling. In H. R. Moskowitz, B. Scharf, and J. C. Stevens (Eds.), *Sensation and measurement: Papers in honor of S. S. Stevens*. Dordrecht, Holland: D. Reidel Publishing.

Draper, J. V., Fujita, Y., and Herndon, J. N. (1987). *Evaluation of High-definition Television for Remote Task Performance* (ORNL/TM-10303). Oak Ridge, TN: Oak Ridge National Laboratory.

Ehrlich, J. A., and Singer, M. J. (1994). Are stereoscopic displays necessary for virtual environments? In *Proceedings of the 38th Annual Meeting of the Human Factors and Ergonomics Society,* (pp. 952). Santa Monica, CA: HFES.

Ellis, S. R. (1991). Nature and origins of virtual environments: A bibliographical essay. *Computing Systems in Engineering, 2*(4), 321-347.

Ellis, S. R. (1993). What are virtual environments? In *ICAT '93,* (pp. 81-89). Tokyo, Japan: ICAT.

Ellis, S. R. (1996). Presence of mind: A reaction to Thomas Sheridan's "Further musings on the psychophysics of presence". *Presence, 5*(2), 247-259.

Engen, T. (1972). Psychophysics: II. scaling methods. In J. W. Kling and L. A. Riggs (Eds.), *Woodworth & Schlosberg's Experimental Psychology*. (pp. 47-86). New York, NY: Holt, Rinehart, and Winston.

Frank, L. H. (1986). *Effects of visual display and motion system delays on operator performance and uneasiness in a driving simulator*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University.

Frank, L. H., Casali, J. G., and Wierwille, W. W. (1988). Effects of visual display and motion system delays on operator performance and uneasiness in a driving simulator. *Human Factors, 30*(2), 201-218.

Gescheider, G. A. (1985). *Psychophysics: Method, theory, and application*. Hillsdale, NJ: Lawrence Erlbaum.

Gescheider, G. A., and Bolanowski, S. J., Jr. (1991a). Final comments on ratio scaling of psychological magnitudes. In S. J. Bolanowski, Jr. and G. A. Gescheider (Eds.), *Ratio scaling of psychological magnitude: In honor of the memory of S. S. Stevens* (pp. 295-311). Hillsdale, NJ: Lawrence Erlbaum.

Gescheider, G. A., and Bolanowski, S. J., Jr. (1991b). Introduction to conference on ratio scaling of psychological magnitudes. In S. J. Bolanowski, Jr. and G. A. Gescheider (Eds.), *Ratio scaling of psychological magnitude: In honor of the memory of S. S. Stevens* (pp. 1-7). Hillsdale, NJ: Lawrence Erlbaum.

Gilkey, R. H., and Weisenberger, J. M. (1995). The sense of presence in the suddenly deafened adult: Implications for virtual environments. *Presence, 4*(4), 357-363.

Green, C. A., and Williges, R. C. (1995). Evaluation of alternative media used with a groupware editor in a simulated telecommunications environment. *Human Factors, 37*(2), 283-289.

Han, S. H. (1991a). Data bridging: Statistical methods for integrating results across studies. In *Proceedings of the Human Factors Society 34th Annual Meeting*, (pp. 1265-1269). Santa Monica, CA: Human Factors Society.

Han, S. H. (1991b). *Integrated empirical models based on a sequential research strategy*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University.

Han, S. H., Williges, R. C., and Williges, B. H. (1990). Empirical models based on data from several experiments. In *Proceedings of the Human Factors Society 34th Annual Meeting*, (pp. 1224-1228). Santa Monica, CA: Human Factors Society.

Harris, C. S. (1965). Perceptual adaptation to inverted, reversed, and displaced vision. *Psychological Review*, 72(72), 419-444.

Hatada, T., Sakata, H., and Kusaka, H. (1980). Psychophysical analysis of the "sensation of reality" induced by a visual wide-field display. *SMPTE Journal, 89*, 560-569.

Hays (1981). *Statistics*. New York: Holt, Rinehart and Winston.

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence, 1*(2), 262-271.

Held, R. and Durlach, N. (1991). Telepresence, time delay and adaptation. In S. R. Ellis, M. K. Kaiser, and A. J. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (pp. 232-246). New York: Taylor & Francis.

Held, R. M. and Durlach, N. I. (1992). Telepresence. *Presence, 1*(1), 109-112.

Hix, D. and Hartson, H. R. (1993). *Developing user interfaces: Ensuring usability through product and process*. New York: John Wiley & Sons.

Hsu, J., Pizlo, Z., Babbs, C. F., Chelberg, D. M., and Delp, E. J. (1994). Design of studies to test the effectiveness of stereo imaging. Truth or dare: Is stereo viewing really better? In *Proceedings of the SPIE - The International Society for Optical Engineering vol. 2177*, (pp. 211-222). Bellingham, WA: SPIE.

Jones, E. R., Hennessy, R. T., and Deutsch, S. (Ed.) (1985). *Human factors aspects of simulation*. Washington, DC: National Academy Press.

Kalawsky, R. S. (1993). *The science of virtual reality and virtual environments*. Reading, MA: Addison-Wesley.

Kawahara, H., Kaoru, W., Watanabe, A., and Funabiki, K. (1996). Pilot evaluation results of a half-dome type simulator system. In *AIAA Flight Simulation Technologies Conference*, (pp. 716-725). New York, NY: American Institute for Aeronautics and Astronautics.

Kinkade, R. G., Snyder, H. L., and Greening, C. P. (1963). Simulation of a star field. *Human Factors, 5*(3), 335-338.

Kollin, J. (1993). A retinal display for virtual-environment applications. In *Society for Information Display International Symposium*, (pp. 827-828). Playa del Rey, CA: Society for Information Display.

Lampton, D. R., Knerr, B. W., Goldberg, S. L., Bliss, J. P., Moshell, J. M., and Blau, B. S. (1994). The virtual environment performance assessment battery (VEPAB): Development and evaluation. *Presence, 3*(2), 145-157.

Lampton, D. R., Kolasinski, E. M., Knerr, B. W., Bliss, J. P., Bailey, J. H., and Witmer, B. G. (1994). Side effects and aftereffects of immersion in virtual environments. In *Proceedings of the 38th Annual Meeting of the Human Factors and Ergonomics Society*, (pp. 1154-1157). Santa Monica, CA: HFES.

Lane, H. L., Catania, A. C., and Stevens, S. S. (1961). Voice level: autophonic scale, perceived loudness, and effects of sidetone. The *Journal of the Acoustical Society of America, 33*(2), 160-167.

Loomis, J. M. (1992a). Distal attribution and presence. *Presence, 1*(1), 113-119.

Loomis, J. M. (1992b). Presence and distal attribution: phenomenology, determinants, and assessment. In *Human Vision, Visual Processing and Digital Display III*, (pp. 590-595). Bellingham, WA: SPIE.

Luce, R. D., and Krumhansl, C. (1988). Measurement, scaling, and psychophysics. In R. C. Atkinson, R. J. Hernstein, G. Lindzey, and R. D. Luce (Eds.), *Stevens' handbook of experimental psychology*. (pp. 3-74). New York: Wiley.

Mansfield, R. J. (1974). Measurement, invariance, and psychophysics. In H. R. Moskowitz, B. Scharf, and J. C. Stevens (Eds.), *Sensation and measurement: Papers in honor of S. S. Stevens*. Dordrecht, Holland: D. Reidel Publishing.

McGreevy, M. W. (1993). The presence of field geologists in a Mars-like terrain. *Presence, 1*(4), 375-403.

McGreevy, M. W. (1994). *Ethnographic Object-Oriented Analysis of Explorer Presence in a Volcanic Terrain Environment: Claims and Evidence* (NAS 1.15:108823). Moffett Field, CA: National Aeronautics and Space Administration.

Merriken, M. S., Riccio, G. E., and Johnson, W. V. (1987). Temporal fidelity in aircraft simulator visual systems. In *AIAA Flight Simulation Technologies Conference*, (pp. 50-54). New York, NY: American Institute for Aeronautics and Astronautics.

Miller, M. E. (1993). The effect of display characteristics on perception of motion from remotely sensed imagery. In *Society for Information Display International Symposium*, (pp. 118-121). Seattle, WA: Society for Information Display.

Montgomery, D. C. and Peck, E. A. (1992). *Introduction to linear regression analysis*. New York: John Wiley & Sons.

Morrison, H. B. (1995). *The influence of depth and stutter on consumer preference for static three-dimensional lenticular-sheet images*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University.

Myers, R. H. (1990). *Classical and modern regression with applications* (2nd ed.). Boston, MA: PWS-Kent.

Padmos, P. and Milders, M. V. (1992). Quality criteria for simulator images: A literature review. *Human Factors, 34*(6), 727-748.

Patterson, R. and Martin, W. R. (1992). Human stereopsis. *Human Factors, 34*(6), 669-692.

Piantanida, T. P., Boman, D., Larimer, J., Gille, J., and Reed, C. (1992). Studies of the field-of-view/resolution tradeoff in virtual-reality systems. In *Human Vision, Visual Processing and Digital Display III*, (pp. 448-456). Bellingham, WA: SPIE.

Post, D. L., Sarma, K. R., Heinze, W. C., Ellis, K. R., Larson, B. D., Franklin, H., Trimmer, J., and Rodgers, J. (1994). A new color display for head-mounted displays. *Journal of the SID, 2*(4), 155-164.

Psotka, J., Davison, S. A., and Lewis, S. A. (1993). Exploring immersion in virtual space. *Virtual Reality Systems, 1*(2), 70-84.

Ray, P. A. (1995). Federal Aviation Administration flight simulation standards: A regulatory update. In *AIAA Flight Simulation Technologies Conference*, (pp. 37-47). San Diego, CA: American Institute for Aeronautics and Astronautics.

Ray, P. A. (1996). Quality flight simulation cueing: Why? In *AIAA Flight Simulation Technologies Conference*, (pp. 138-147). San Diego, CA: American Institute for Aeronautics and Astronautics.

Regan, E. C. (1994). Some side-effects of immersion virtual reality. In *AGARD, Virtual Interfaces: Research and Applications*. Neuilly sur Seine, France: AGARD.

Regan, E. C. and Price, K. R. (1994). The frequency of occurrence and severity of side-effects of immersion virtual reality. *Aviation, Space, & Environmental Medicine, 65*(6), 527-530.

Reid, L. D. and Nahon, M. A. (1987). The response of airline pilots to flight simulator motion. In *AIAA Flight Simulation Technologies Conference*, (pp. 77-85). New York, NY: American Institute for Aeronautics and Astronautics.

Reinhart, W. F., Beaton, R. J., and Snyder, H. L. (1990). Comparison of depth cues for relative depth judgments. In *Proceedings of the SPIE - The International Society for Optical Engineering vol. 1256*, (pp. 12-21). Bellingham, WA: SPIE.

Roberts, F. S. (1979). *Measurement theory*. New York: Addison-Wesley.

Rolfe, J. M. and Staples, K. J. (Ed.) (1986). *Flight simulation*. New York, NY: Cambridge University Press.

Sanders, M. S. and McCormick, E. J. (1987). *Human factors in engineering and design* (6th ed.). New York, NY: McGraw-Hill.

Schloerb, D. W. (1995). A quantitative measure of telepresence. *Presence, 4*(1), 64-80.

Sebok, A. L. (1991). *Effects of color CRT misconvergence and font type on text readability and subjective preference*. Unpublished masters thesis, Virginia Polytechnic Institute and State University.

Shaw, J. (1993, August). Virtual reality resource guide. *AI Expert*, p. 34-43.

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence, 1*(1), 120-126.

Sheridan, T. B. (1996). Further musings on the psychophysics of presence. *Presence, 5*(2), 241-246.

Singer, M. J. and Witmer, B. G. (1995). *Presence measures for virtual environments: Background and scoring instructions* (FEB 1995). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Singer, M. J., Witmer, B. G., and Bailey, J. H. (1994). Development of "presence" measures for virtual environments. In *Proceedings of the 38th Annual Meeting of the Human Factors and Ergonomics Society*, (pp. 983). Santa Monica, CA: Human Factors and Ergonomics Society.

Slater, M. and Usoh, M. (1993a). *An experimental exploration of presence in virtual environments*. London, UK: Department of Computer Science, QMW University.

Slater, M. and Usoh, M. (1993b). The influence of a virtual body on presence in immersive virtual environments. In *Virtual Reality International 93. Proceedings of the Third Annual Conference on Virtual Reality*, (pp. 34-42). London, UK: Meckler.

Slater, M. and Usoh, M. (1993c). Presence in immersive virtual environments. In *1993 IEEE Annual Virtual Reality International Symposium*, (pp. 90-96). Piscataway, NJ: IEEE.

Slater, M. and Usoh, M. (1993d). Representation systems, perceptual position and presence in virtual environments. *Presence, 2*(3), 221-233.

Slater, M., Usoh, M., and Steed, A. (1994). Depth of presence in virtual environments. *Presence, 3*(2), 130-144.

Snyder, H. L. (1988). Toward the determination of electronic display image quality. In *Advances in Man-Machine Systems Research, Vol. 4* , (pp. 1-68). Greenwich, CN: JAI Press.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4), 73-93.

Stevens, S. S. (1953). On the brightness of lights and loudness of sounds. *Science, 118*, 576.

Stevens, S. S. (1955). The measurement of loudness. *Journal of the Acoustical Society of America, 27*, 815-829.

Stevens, S. S. (1971). Issues in psychophysical measurement. *Psychological Review, 78*, 426-450.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York, NY: John Wiley & Sons.

Sutherland, I. E. (1965). The ultimate display. *International federation of information processing, 2*, 506.

Travis, D., Watson, T., and Atyeo, M. (1994). Human psychology in virtual environments. In L. MacDonald and J. Vince (Eds.), *Interacting with virtual environments*, (pp. 43-59). New York: John Wiley & Sons

Verillo, R. T. (1991). Measurement of vibrotactile sensation magnitude. In S. J. Bolanowski, Jr. and G. A. Gescheider (Eds.), *Ratio scaling of psychological magnitude: In honor of the memory of S. S. Stevens* (pp. 260-275). Hillsdale, NJ: Lawrence Erlbaum

Waldern, J. D. (1993). Virtual reality — the next generation. In *Virtual Reality International 93. Proceedings of the Third Annual Conference on Virtual Reality*, (pp. 87-100). London, UK: Meckler.

Ware, C. and Balakrishnan, R. (1994). Reaching for objects in VR displays: lag and frame rate. *ACM Transactions on Computer-Human Interaction, 1*(4), 331-356.

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

Wenzel, E. M. (1992). Localization in virtual acoustic displays. *Presence, 1*(1), 80-107.

Wexelblat, A. (1993). The reality of cooperation: Virtual reality and CSCW. In A. Wexelblat (Ed.), *Virtual reality: Applications and explorations*. (pp. 23-44). New York: Academic Press.

Williges, R. C. (1981). Development and use of research methodologies for complex system/simulation experimentation. In M. Morral and K. Kraiss (Eds.), *Manned system design* (pp. 59-87). New York, NY: Plenum.

Williges, R. C. and Williges, B. H. (1989). Integrated research paradigm for complex experimentation. In *Proceedings of the Human Factors Society 33rd Annual Meeting*, (pp. 606-610). Santa Monica, CA: Human Factors Society.

Williges, R. C., Williges, B. H., and Han, S. H. (1992). Developing quantitative guidelines using integrated data from sequential experiments. *Human Factors, 34*(4), 399-408.

Williges, R. C., Williges, B. H., and Han, S. H. (1993). Sequential experimentation in human-computer interface design. In H. R. Hartson and D. Hix (Eds.), *Advance in Human-Computer Interaction* (pp. 1-30). Norwood, NJ: Ablex

Zeltzer, D. (1992). Autonomy, interaction, and presence. *Presence, 1*(1), 127-132.

Zenyuh, J., Reising, J. M., Walchli, S., and Biers, D. (1988). A comparison of a stereoscopic 3-D display versus a 2-D display using advanced air-to-air format. In *Proceedings of the Human Factors Society 32nd Annual Meeting*, (pp. 53-57). Santa Monica, CA: Human Factors Society.

Zwislocki, J. J. (1983). Group and individual relations between sensation magnitude and their numerical estimates. *Perception and Psychophysics, 33*, 460-468.

Zwislocki, J. J., and Goodman, D. A. (1980). Absolute scaling of sensory magnitudes: A validation. *Perception and Psychophysics, 28*, 28-38.

APPENDIX A: INTRODUCTION TO THE STUDY

The purpose of this study is to investigate the perception of presence in a virtual environment and the relationship between presence and performance. Presence is the degree to which a user feels as if he/she is actually present in a virtual environment. The study is being conducted in the Human-Computer Interface Laboratory (HCIL), Department of Industrial and Systems Engineering (ISE) at Virginia Tech (530 Whittemore Hall). The principal investigators are Michael P. Snow, a graduate student in ISE, and Dr. R. C. Williges, director of the HCIL.

In this study you will be asked to perform a variety of tasks in a virtual environment. These tasks involve navigating, estimating distances, manipulating an object, looking for an object, and reacting to the appearance of an object. The design of the virtual environment is being evaluated, not you. Please do not be nervous about your performance on any of the tasks, just follow instructions and proceed in a manner that is comfortable for you.

You are being asked to spend two to three hours in the HCIL participating in this experiment. After reading this introduction, you will be asked to fill out an informed consent form. If you agree to participate, your vision will be tested to ensure that you have normal visual acuity, color vision, and ability to perceive objects in depth. A measurement will also be taken to establish your virtual height.

Before the actual experimental session begins, you will be given further instruction about the tasks and what is expected of you as a participant. You will be given practice on a rating task that you will perform and shown the tasks to be performed in the virtual environment.

Once you are familiar with the tasks, you will be asked to perform them several times. The virtual environment will be altered in some way each time.

If you pass the vision screening and participate in the study, you will be paid $5 per hour for your participation. During the experiment, if for any reason you decide not to continue, you will be paid for the time that you have participated. Similarly, if the experiment is interrupted and must be terminated because of an equipment failure, you will be paid for the time spent up to the point of termination.

If you are still interested in participating in the study, please read and sign the informed consent form. Thank you for your participation.

APPENDIX B: INFORMED CONSENT FORM

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Informed Consent for Participants of Investigative Projects

Title of Project: Presence and performance in virtual environments

Principal Investigator: Michael P. Snow

I. THE PURPOSE OF THIS RESEARCH AND ITS PROCEDURES

- You are invited to participate in this study whose purpose, description, and procedures are contained in the Introduction to the Study document, which you have already read. This study involves 48 participants, including yourself. There is a small chance that you will experience some temporary side effects from wearing the head-mounted display. Examples reported in the past by subjects in other virtual reality studies at other institutions include eyestrain, disorientation, dizziness, and symptoms similar to motion sickness. It is unlikely that you will experience these effects; however, if you experience any adverse symptoms, please notify the experimenter immediately. Remember that you are free to withdraw from the study at any time.

II. BENEFITS OF THIS PROJECT

- While there are no direct benefits to you from participating in this study (other than payment), you may find the experiment interesting and even entertaining.

- No guarantee of benefits has been made to encourage you to participate.

- You may receive a synopsis or summary of this research when it is completed. Please leave a self-addressed envelope or your electronic mailing address with the experimenter if this is what you wish.

III. EXTENT OF ANONYMITY AND CONFIDENTIALITY

- Your anonymity will be strictly preserved. The information you provide will have your name removed and only a subject number will identify you during analyses and any written reports of the research.

IV. COMPENSATION

- For participation in the project you will receive $5 per hour. Payment will be made immediately following participation in the experiment.

V. FREEDOM TO WITHDRAW

- You are free to withdraw from this study at any time without penalty. If you choose to withdraw, compensation will be prorated and you will be paid for the time you spent participating in the study. This will also be the case if the investigator terminates the experiment because of equipment failure.

VI. APPROVAL OF RESEARCH

- This research project has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University.

VII. SUBJECT'S RESPONSIBILITIES AND PERMISSION

- I know of no reason I cannot participate in this study.

- I have read and understand the informed consent and conditions of this project. All of my questions have been answered. I hereby acknowledge the above and give my voluntary consent to participate in this project.

- If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

Signature ____

Date ____

VIII. PARTICIPANT'S CONTACTS

Should I have any questions about this research or its conduct, I will contact:

Michael P. Snow	1-3193
Investigator	Phone
Dr. R. C. Williges	1-6270
Faculty Advisor	Phone
Ernest R. Stout	1-9359
Chair, IRB Research Division	Phone

APPENDIX C: PARTICIPANT INSTRUCTIONS

In this experiment, you will be asked to perform a set of tasks in a virtual environment. This set of tasks will be repeated a number of times and the virtual environment will be altered in some way each time. You will be asked to assign a number to your feeling of "presence" in the virtual environment each time you perform the set of tasks. By presence, what is meant is how much you feel as if you are actually present in the virtual environment each time you perform the tasks. Assign any number that seems appropriate to you to the first set of tasks. Then assign successive numbers in such a way that they reflect your subjective impression of presence in the virtual environment. For example, if your sense of presence during a set of tasks seems 10 times greater than on the first set of tasks, assign it a number 10 times larger than the first. If your sense of presence during a set of tasks seems one fifth what it was on the first set of tasks, assign it a number one fifth as great as the first. There is no limit to the range of numbers that you may use. You may use any positive numbers that seem appropriate to you — whole numbers, decimals, or fractions. Try to make each number proportional to your feeling of presence in the virtual environment as you perceive it.

To give you some practice at this rating procedure, you will first be asked to rate the length of twenty lines. These lines will be presented to you one at a time in random order. Assign a number to each line that matches your impression of its length. Assign any number to the first line that seems appropriate to you and assign successive numbers so that they reflect your subjective impression of how long each line is. There is no limit to the range of numbers you may use and you may use any positive numbers that seem appropriate to you — whole numbers, decimals, or fractions. Try to make each number proportional to your impression of the length of each line. Let the experimenter know when you are ready to begin.

*****

Next, you will be shown the set of tasks you will be performing in the virtual environment. You will be given control of one 3D controller and a mouse. The 3D controller controls your viewpoint (except during the *Bins* task; this will be explained later) and the mouse controls the cursor (the arrow). You will interact with the world by moving around in it using the 3D controller and clicking on objects with the mouse. The experimenter will show you which objects may be activated by clicking on them. You will also be able to look around the virtual environment (in most conditions) by moving your head. There are five tasks in each set. The tasks have the following names: *Distance Estimation*, *Bins*, *Turns*, *Search*, and *Choice*. You should attempt to perform all tasks as quickly and accurately as possible. All tasks are timed except for the *Distance Estimation* task. During the timed tasks (except for the *Bins* task) interactions and manipulations in the environment will be turned off. You will be given feedback at the end of each set of tasks on how quickly you performed the tasks and any errors you made. The tasks are described below.

Distance Estimation. A human figure will appear at the end of a forty-foot corridor. This figure is six feet tall. You will also see a response panel to your right. The figure will begin to move toward you when you contact the green start block at the entry to the corridor. Use the mouse to click on the appropriate button when the figure appears to be 20 feet away, 10 feet away, 5 feet away, and 2.5 feet away ("arm's length").

Bins. You will face a three by three stack of open-ended, box-like compartments (bins) much like Tic-Tac-Toe squares. To begin the task, move into the red start block until you contact the green part. Face the bins as you do this. A red ball will appear in one of the bins and an X will appear in another. Use the 3D controller to move the blue 3D cursor in front of you until it contacts the ball. Then drag the ball out of its start bin and put it into the bin marked with an X.

Turns. You will navigate through a continuous narrow corridor that turns alternately to the left and right.

Search. To begin the task, move into the red start block until you contact the green part. Face the door across from the one you entered as you do this. A red ball will appear near the floor, ceiling, or walls (outside your view), and move slowly around the room. Search for the ball by making head movements and/or by controlling your viewpoint using the 3D controller. Click on the ball with the mouse when you see it.

Choice. To begin the task, move into the red start block until you contact the green part. Face the four white blocks on the wall as you do this. An X will appear on one of the blocks. Use the mouse to click on either the X or the block on which it appears.

After the Choice task, move onto the red end block in the corner of the room to end the trial and get feedback on your performance. You are free to click on interactive objects at any time except during timed tasks. Let the experimenter know when you are ready for a demonstration.

APPENDIX D: ANOVA SUMMARY TABLES

Table 41. ANOVA summary table for distance estimation performance in Experiment 1.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	3259.38	296.31		
Within-Subject					
Actual Distance (ACT)	3	39099.30	13033.10	1473.44	0.0001
ACT x S	33	291.90	8.85		
Scene Update Rate (SUR)	2	18.01	9.01	2.17	0.138
SUR x S	22	91.30	4.15		
Visual Display Resolution (VDR)	1	110.61	110.61	13.33	0.004
VDR x S	11	91.28	8.30		
Field of View (FOV)	2	67.07	33.53	4.12	0.030
FOV x S	22	179.11	8.14		
ACT x SUR	6	5.10	0.85	1.56	0.174
ACT x SUR x S	66	36.02	0.55		
ACT x VDR	3	21.16	7.05	6.14	0.002
ACT x VDR x S	33	37.91	1.15		
ACT x FOV	6	25.54	4.26	4.27	0.001
ACT x FOV x S	66	65.77	1.00		

SUR x VDR	2	10.88	5.44	1.40	0.268
SUR x VDR x S	22	85.40	3.88		
SUR x FOV	4	6.69	1.67	0.42	0.196
SUR x FOV x S	44	176.66	4.02		
VDR x FOV	2	38.04	19.02	3.31	0.054
VDR x FOV x S	22	126.39	5.74		
ACT x SUR x VDR	6	5.30	0.88	1.04	0.406
ACT x SUR x VDR x S	66	55.95	0.85		
ACT x SUR x FOV	12	10.79	0.90	1.22	0.278
ACT x SUR x FOV x S	132	97.55	0.74		
ACT x VDR x FOV	6	24.07	4.01	3.15	0.009
ACT x VDR x FOV x S	66	83.98	1.27		
SUR x VDR x FOV	4	49.52	12.38	2.52	0.057
SUR x VDR x FOV x S	44	216.37	4.92		
ACT x SUR x VDR x FOV	12	13.71	1.14	1.12	0.350
ACT x SUR x VDR x FOV x S	132	134.69	1.02		
Total	852	41276.07	13298.43		

Table 42. ANOVA summary table for time to complete the Bins task in Experiment 1.

Source	***df***	***SS***	***MS***	***F***	***p***
<u>Between-Subjects</u>					
Subjects (S)	11	19550.64	1777.33		
<u>Within-Subject</u>					
Scene Update Rate (SUR)	2	953.22	476.61	2.36	0.118
SUR x S	22	4439.69	201.80		
Visual Display Resolution (VDR)	1	448.56	448.56	2.97	0.113
VDR x S	11	1661.53	151.05		
Field of View (FOV)	2	89.42	44.71	0.39	0.685
FOV x S	22	2551.88	115.99		
SUR x VDR	2	10.86	5.43	1.01	0.987
SUR x VDR x S	22	9045.13	411.14		
SUR x FOV	4	491.80	122.95	0.52	0.720
SUR x FOV x S	44	10370.60	235.69		
VDR x FOV	2	198.73	99.37	0.66	0.527
VDR x FOV x S	22	3313.19	150.60		
SUR x VDR x FOV	4	299.30	74.83	0.60	0.662
SUR x VDR x FOV x S	44	5453.36	123.94		
Total	204	39327.27	2662.67		

Table 43. ANOVA summary table for errors made during the Turns task in Experiment 1.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	3259.38	296.31		
Within-Subject					
Scene Update Rate (SUR)	2	40.56	20.28	1.18	0.327
SUR x S	22	379.66	17.26		
Visual Display Resolution (VDR)	1	9.38	9.38	0.19	0.675
VDR x S	11	554.35	50.40		
Field of View (FOV)	2	182.95	91.48	1.81	0.187
FOV x S	22	1110.94	50.50		
SUR x VDR	2	115.58	57.79	0.91	0.416
SUR x VDR x S	22	1392.19	63.28		
SUR x FOV	4	56.38	14.09	0.49	0.745
SUR x FOV x S	44	1272.40	28.92		
VDR x FOV	2	5.03	2.51	0.08	0.923
VDR x FOV x S	22	685.75	31.17		
SUR x VDR x FOV	4	121.47	30.37	0.96	0.439
SUR x VDR x FOV x S	44	1391.75	31.63		
Total	204	7318.39	499.06		

Table 44. ANOVA summary table for time to complete the Turns task in Experiment 1.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	14145.72	1285.97		
Within-Subject					
Scene Update Rate (SUR)	2	3214.85	1607.43	11.27	0.001
SUR x S	22	3136.65	142.57		
Visual Display Resolution (VDR)	1	2.29	2.29	0.03	0.869
VDR x S	11	876.54	79.69		
Field of View (FOV)	2	670.40	335.20	3.20	0.061
FOV x S	22	2307.97	104.91		
SUR x VDR	2	145.27	72.63	0.84	0.446
SUR x VDR x S	22	1907.07	86.68		
SUR x FOV	4	259.24	64.81	0.46	0.764
SUR x FOV x S	44	6178.48	140.42		
VDR x FOV	2	113.17	56.59	0.26	0.776
VDR x FOV x S	22	4820.80	219.13		
SUR x VDR x FOV	4	551.65	137.91	1.52	0.212
SUR x VDR x FOV x S	44	3985.03	90.57		
Total	204	28169.41	3140.83		

Table 45. ANOVA summary table for the Search task in Experiment 1.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	1023.87	93.08		
Within-Subject					
Scene Update Rate (SUR)	2	128.45	64.22	1.20	0.321
SUR x S	22	1179.83	53.63		
Visual Display Resolution (VDR)	1	306.40	306.40	5.30	0.042
VDR x S	11	636.28	57.84		
Field of View (FOV)	2	986.80	493.40	11.33	0.001
FOV x S	22	958.38	43.56		
SUR x VDR	2	15.27	7.64	0.13	0.879
SUR x VDR x S	22	1295.59	58.89		
SUR x FOV	4	341.40	85.35	1.64	0.180
SUR x FOV x S	44	2283.58	51.90		
VDR x FOV	2	403.92	201.96	3.83	0.037
VDR x FOV x S	22	1160.18	52.74		
SUR x VDR x FOV	4	52.14	13.03	0.30	0.877
SUR x VDR x FOV x S	44	1918.39	43.60		
Total	204	11666.61	1534.16		

Table 46. ANOVA summary table for the Choice task in Experiment 1.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	2.18	0.20		
Within-Subject					
Scene Update Rate (SUR)	2	1.89	0.94	5.46	0.012
SUR x S	22	3.80	0.17		
Visual Display Resolution (VDR)	1	0.58	0.58	3.12	0.105
VDR x S	11	2.05	0.19		
Field of View (FOV)	2	0.35	0.17	1.41	0.265
FOV x S	22	2.70	0.12		
SUR x VDR	2	0.07	0.03	0.25	0.780
SUR x VDR x S	22	2.97	0.13		
SUR x FOV	4	0.87	0.22	1.87	0.133
SUR x FOV x S	44	5.12	0.12		
VDR x FOV	2	0.06	0.03	0.20	0.823
VDR x FOV x S	22	3.41	0.16		
SUR x VDR x FOV	4	0.40	0.10	1.04	0.398
SUR x VDR x FOV x S	44	4.25	0.10		
Total	204	28.52	3.06		

Table 47. ANOVA summary table for presence in Experiment 1.

Source	*df*	*SS*	*MS*	*F*	*p*
Between-Subjects					
Subjects (S)	11	233.94	21.27		
Within-Subject					
Scene Update Rate (SUR)	2	33.01	16.51	4.70	0.020
SUR x S	22	77.34	3.52		
Visual Display Resolution (VDR)	1	112.95	112.95	8.93	0.012
VDR x S	11	139.14	12.65		
Field of View (FOV)	2	1282.64	641.32	18.13	0.0001
FOV x S	22	778.13	35.37		
SUR x VDR	2	1.10	0.55	0.10	0.901
SUR x VDR x S	22	115.49	5.25		
SUR x FOV	4	15.26	3.82	0.60	0.668
SUR x FOV x S	44	281.80	6.40		
VDR x FOV	2	6.31	3.15	0.39	0.681
VDR x FOV x S	22	177.47	8.07		
SUR x VDR x FOV	4	11.28	2.82	0.64	0.636
SUR x VDR x FOV x S	44	193.47	4.40		
Total	204	3225.39	856.78		

Table 48. ANOVA summary table for distance estimation performance in Experiment 2.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	503.04	45.73		
Within-Subject					
Actual Distance (ACT)	3	30593.32	10197.77	678.75	0.0001
ACT x S	33	495.80	15.02		
Sound (SND)	1	0.06	0.06	0.01	0.940
SND x S	11	114.24	10.39		
Textures (TEX)	1	12.85	12.85	9.10	0.012
TEX x S	11	15.54	1.41		
Head-tracking (HT)	1	3.79	3.79	0.39	0.547
HT x S	11	107.79	9.80		
Stereopsis (STE)	1	53.34	53.34	4.99	0.047
STE x S	11	117.62	10.69		
Virtual Personal Risk (VPR)	1	3.31	3.31	0.49	0.498
VPR x S	11	74.02	6.73		
ACT x SND	3	0.35	0.12	0.12	0.947
ACT x SND x S	33	31.78	0.96		
ACT x TEX	3	0.93	0.31	0.44	0.723

ACT x TEX x S	33	22.91	0.69		
ACT x HT	3	5.01	1.67	1.23	0.316
ACT x HT x S	33	44.93	1.36		
ACT x STE	3	1.10	0.37	0.30	0.824
ACT x STE x S	33	39.92	1.21		
ACT x VPR	3	4.56	1.52	1.14	0.348
ACT x VPR x S	33	44.07	1.34		
SND x TEX	1	0.24	0.24	0.03	0.866
SND x TEX x S	11	88.44	8.04		
SND x HT	1	0.12	0.12	0.01	0.912
SND x HT x S	11	107.38	9.76		
SND x STE	1	39.82	39.82	9.61	0.010
SND x STE x S	11	45.59	4.14		
SND x VPR	1	0.89	0.89	0.18	0.680
SND x VPR x S	11	54.36	4.94		
TEX x HT	1	16.66	16.66	2.37	0.152
TEX x HT x S	11	77.48	7.04		
TEX x STE	1	5.20	5.20	1.20	0.296
TEX x STE x S	11	47.56	4.32		
TEX x VPR	1	13.87	13.87	3.53	0.087

TEX x VPR x S	11	43.19	3.93		
HT x STE	1	6.60	6.60	1.03	0.331
HT x STE x S	11	70.21	6.38		
HT x VPR	1	0.47	0.47	0.16	0.701
HT x VPR x S	11	33.36	3.03		
STE x VPR	1	1.90	1.90	0.37	0.553
STE x VPR x S	11	55.88	5.08		
ACT x SND x TEX	3	5.83	1.94	3.76	0.020
ACT x SND x TEX x S	33	17.06	0.52		
ACT x SND x HT	3	3.40	1.13	1.92	0.145
ACT x SND x HT x S	33	19.45	0.59		
ACT x SND x STE	3	2.16	0.72	0.74	0.534
ACT x SND x STE x S	33	32.04	0.97		
ACT x SND x VPR	3	1.42	0.47	0.33	0.800
ACT x SND x VPR x S	33	46.63	1.41		
ACT x TEX x HT	3	5.91	1.97	1.09	0.366
ACT x TEX x HT x S	33	59.47	1.80		
ACT x TEX x STE	3	1.53	0.51	0.58	0.634
ACT x TEX x STE x S	33	29.07	0.88		
ACT x TEX x VPR	3	3.77	1.26	1.37	0.269

ACT x TEX x VPR x S	33	30.29	0.92		
ACT x HT x STE	3	1.01	0.34	0.21	0.891
ACT x HT x STE x S	33	53.91	1.63		
ACT x HT x VPR	3	3.14	1.05	1.45	0.246
ACT x HT x VPR x S	33	23.76	0.72		
ACT x STE x VPR	3	3.80	1.27	1.38	0.266
ACT x STE x VPR x S	33	30.25	0.92		
Total	756	32870.36	10498.16		

Table 49. ANOVA summary table for time to complete the Bins task in Experiment 2.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	7299.47	663.59		
Within-Subject					
Sound (SND)	1	45.67	45.67	0.26	0.622
SND x S	11	1949.31	177.21		
Textures (TEX)	1	0.35	0.35	0.001	0.966
TEX x S	11	2000.97	181.91		
Head-tracking (HT)	1	611.83	611.83	4.26	0.063
HT x S	11	1578.24	143.48		
Stereopsis (STE)	1	101.41	101.41	1.29	0.280
STE x S	11	863.77	78.52		
Virtual personal risk (VPR)	1	379.91	379.91	4.51	0.057
VPR x S	11	926.94	84.27		
SND x TEX	1	109.26	109.26	0.78	0.396
SND x TEX x S	11	1538.18	139.83		
SND x HT	1	224.64	224.64	1.41	0.260
SND x HT x S	11	1748.07	158.92		
SND x STE	1	745.92	745.92	8.38	0.015

SND x STE x S	11	978.81	88.98		
SND x VPR	1	6.11	6.11	0.05	0.831
SND x VPR x S	11	1407.94	127.99		
TEX x HT	1	476.97	476.97	7.69	0.018
TEX x HT x S	11	682.28	62.03		
TEX x STE	1	14.05	14.05	0.14	0.717
TEX x STE x S	11	1120.35	101.85		
TEX x VPR	1	292.25	292.25	2.37	0.152
TEX x VPR x S	11	1355.84	123.26		
HT x STE	1	0.01	0.01	0.00009	0.993
HT x STE x S	11	1002.30	91.12		
HT x VPR	1	0.02	0.02	0.0002	0.989
HT x VPR x S	11	984.51	89.50		
STE x VPR	1	73.31	73.31	0.44	0.520
STE x VPR x S	11	1823.63	165.78		
Total	180	23042.85	4896.36		

Table 50. ANOVA summary table for errors made in completing the Turns task in Experiment 2.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	102979.68	9361.79		
Within-Subject					
Sound (SND)	1	15.76	15.76	0.01	0.924
SND x S	11	18316.56	1665.14		
Textures (TEX)	1	584.51	584.51	0.42	0.528
TEX x S	11	15144.31	1376.76		
Head-tracking (HT)	1	12336.05	12336.05	6.18	0.030
HT x S	11	21949.77	1995.43		
Stereopsis (STE)	1	2387.13	2387.13	1.42	0.258
STE x S	11	18437.18	1676.11		
Virtual personal risk (VPR)	1	4730.26	4730.26	5.09	0.045
VPR x S	11	10214.56	928.60		
SND x TEX	1	16818.80	16818.80	6.72	0.025
SND x TEX x S	11	27510.52	2500.96		
SND x HT	1	1931.67	1931.67	0.72	0.415
SND x HT x S	11	29592.64	2690.24		
SND x STE	1	4651.17	4651.17	3.60	0.084

SND x STE x S	11	14220.14	1292.74		
SND x VPR	1	2603.38	2603.38	1.09	0.318
SND x VPR x S	11	26194.43	2381.31		
TEX x HT	1	5094.38	5094.38	2.42	0.148
TEX x HT x S	11	23130.43	2102.77		
TEX x STE	1	5302.51	5302.51	3.63	0.083
TEX x STE x S	11	16060.81	1460.07		
TEX x VPR	1	967.51	967.51	0.42	0.530
TEX x VPR x S	11	25282.31	2298.39		
HT x STE	1	14.63	14.63	0.01	0.926
HT x STE x S	11	17697.18	1608.83		
HT x VPR	1	2220.88	2220.88	0.62	0.449
HT x VPR x S	11	39647.43	3604.31		
STE x VPR	1	1360.01	1360.01	0.66	0.433
STE x VPR x S	11	22540.31	2049.12		
Total	180	386957.23	90649.43		

Table 51. ANOVA summary table for time to complete the Turns task in Experiment 2.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	10014.65	910.42		
Within-Subject					
Sound (SND)	1	983.19	983.19	3.66	0.082
SND x S	11	2955.40	268.67		
Textures (TEX)	1	3.68	3.68	0.02	0.900
TEX x S	11	2436.55	221.50		
Head-tracking (HT)	1	2374.59	2374.59	10.59	0.008
HT x S	11	2466.48	224.23		
Stereopsis (STE)	1	256.32	256.32	0.98	0.344
STE x S	11	2881.80	261.98		
Virtual personal risk (VPR)	1	5.19	5.19	0.04	0.847
VPR x S	11	1457.24	132.48		
SND x TEX	1	789.43	789.43	1.74	0.214
SND x TEX x S	11	4980.63	452.78		
SND x HT	1	631.48	631.48	1.31	0.277
SND x HT x S	11	5305.02	482.27		
SND x STE	1	0.02	0.02	0.0002	0.988

SND x STE x S	11	773.39	70.31		
SND x VPR	1	140.22	140.22	0.32	0.581
SND x VPR x S	11	4758.49	432.59		
TEX x HT	1	612.26	612.26	1.60	0.232
TEX x HT x S	11	4198.53	381.68		
TEX x STE	1	61.83	61.83	0.27	0.614
TEX x STE x S	11	2527.46	229.77		
TEX x VPR	1	323.18	323.18	1.69	0.220
TEX x VPR x S	11	2104.60	191.33		
HT x STE	1	135.48	135.48	1.64	0.227
HT x STE x S	11	909.96	82.72		
HT x VPR	1	154.91	154.91	0.36	0.561
HT x VPR x S	11	4735.86	430.53		
STE x VPR	1	231.97	231.97	1.38	0.264
STE x VPR x S	11	1844.06	167.64		
Total	180	51039.22	10734.23		

Table 52. ANOVA summary table for time to complete the Search task in Experiment 2.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	545.00	49.55		
Within-Subject					
Sound (SND)	1	10.05	10.05	0.55	0.476
SND x S	11	202.76	18.43		
Textures (TEX)	1	0.05	0.05	0.002	0.962
TEX x S	11	256.49	23.32		
Head-tracking (HT)	1	56.78	56.78	4.74	0.052
HT x S	11	131.76	11.98		
Stereopsis (STE)	1	1.15	1.15	0.08	0.785
STE x S	11	162.00	14.73		
Virtual personal risk (VPR)	1	51.57	51.57	5.70	0.036
VPR x S	11	99.50	9.05		
SND x TEX	1	9.36	9.36	1.04	0.329
SND x TEX x S	11	98.57	8.96		
SND x HT	1	0.82	0.82	0.20	0.660
SND x HT x S	11	43.86	3.99		
SND x STE	1	10.65	10.65	0.48	0.503

SND x STE x S	11	244.57	22.23		
SND x VPR	1	24.62	24.62	1.75	0.213
SND x VPR x S	11	154.88	14.08		
TEX x HT	1	25.05	25.05	1.42	0.259
TEX x HT x S	11	194.60	17.69		
TEX x STE	1	62.94	62.94	3.93	0.073
TEX x STE x S	11	176.04	16.00		
TEX x VPR	1	1.90	1.90	0.14	0.712
TEX x VPR x S	11	146.32	13.30		
HT x STE	1	1.65	1.65	0.10	0.752
HT x STE x S	11	173.00	15.73		
HT x VPR	1	2.26	2.26	0.21	0.658
HT x VPR x S	11	120.17	10.92		
STE x VPR	1	4.24	4.24	0.34	0.569
STE x VPR x S	11	135.60	12.33		
Total	180	2603.21	475.83		

Table 53. ANOVA summary table for time to complete the Choice task in Experiment 2.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	4.00	0.36		
Within-Subject					
Sound (SND)	1	0.17	0.17	0.99	0.341
SND x S	11	1.84	0.17		
Textures (TEX)	1	0.25	0.25	2.25	0.162
TEX x S	11	1.21	0.11		
Head-tracking (HT)	1	1.07	1.07	2.32	0.156
HT x S	11	5.06	0.46		
Stereopsis (STE)	1	0.11	0.11	0.61	0.451
STE x S	11	1.93	0.18		
Virtual personal risk (VPR)	1	0.01	0.01	0.03	0.857
VPR x S	11	3.25	0.30		
SND x TEX	1	0.98	0.98	3.68	0.081
SND x TEX x S	11	2.94	0.27		
SND x HT	1	0.26	0.26	2.05	0.180
SND x HT x S	11	1.42	0.13		
SND x STE	1	0.49	0.49	3.27	0.098

SND x STE x S	11	1.64	0.15		
SND x VPR	1	0.21	0.21	0.85	0.377
SND x VPR x S	11	2.77	0.25		
TEX x HT	1	0.24	0.24	1.30	0.279
TEX x HT x S	11	2.07	0.19		
TEX x STE	1	0.41	0.41	1.22	0.294
TEX x STE x S	11	3.69	0.34		
TEX x VPR	1	0.10	0.10	0.55	0.473
TEX x VPR x S	11	2.02	0.18		
HT x STE	1	0.39	0.39	2.12	0.173
HT x STE x S	11	2.01	0.18		
HT x VPR	1	0.00006	0.00006	0.0002	0.989
HT x VPR x S	11	3.25	0.30		
STE x VPR	1	0.15	0.15	0.74	0.408
STE x VPR x S	11	2.23	0.20		
Total	180	42.17006	8.25006		

Table 54. ANOVA summary table for presence in Experiment 2.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	541.55	49.23		
Within-Subject					
Sound (SND)	1	1881.03	1881.03	11.76	0.006
SND x S	11	1759.27	159.93		
Textures (TEX)	1	126.84	126.84	12.87	0.004
TEX x S	11	108.38	9.85		
Head-tracking (HT)	1	1180.64	1180.64	14.40	0.003
HT x S	11	901.92	81.99		
Stereopsis (STE)	1	174.58	174.58	6.17	0.030
STE x S	11	311.32	28.30		
Virtual personal risk (VPR)	1	0.01	0.01	0.0006	0.980
VPR x S	11	206.33	18.76		
SND x TEX	1	27.63	27.63	4.21	0.065
SND x TEX x S	11	72.19	6.56		
SND x HT	1	71.48	71.48	3.94	0.072
SND x HT x S	11	199.32	18.12		
SND x STE	1	3.36	3.36	0.37	0.556

SND x STE x S	11	100.18	9.11		
SND x VPR	1	26.50	26.50	2.63	0.133
SND x VPR x S	11	110.70	10.06		
TEX x HT	1	3.44	3.44	0.76	0.403
TEX x HT x S	11	49.93	4.54		
TEX x STE	1	1.16	1.16	0.07	0.789
TEX x STE x S	11	170.25	15.48		
TEX x VPR	1	0.37	0.37	0.05	0.821
TEX x VPR x S	11	75.86	6.90		
HT x STE	1	0.06	0.06	0.01	0.923
HT x STE x S	11	70.38	6.40		
HT x VPR	1	0.09	0.09	0.01	0.926
HT x VPR x S	11	115.96	10.54		
STE x VPR	1	11.31	11.31	1.01	0.336
STE x VPR x S	11	122.95	11.18		
Total	180	7883.44	3906.22		

Table 55. ANOVA summary table for distance estimation performance in Experiment 3.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	398.11	36.19		
Within-Subject					
Actual Distance (ACT)	3	37228.73	12409.58	736.74	0.0001
ACT x S	33	555.85	16.84		
Interactions (INT)	2	14.79	7.39	1.31	0.291
INT x S	22	124.50	5.66		
Second User (SU)	1	4.03	4.03	0.64	0.442
SU x S	11	69.71	6.34		
Environmental Detail (ED)	2	1.24	0.62	0.11	0.894
ED x S	22	120.47	5.48		
ACT x INT	6	3.92	0.65	0.78	0.587
ACT x INT x S	66	55.08	0.83		
ACT x SU	3	1.36	0.45	0.28	0.836
ACT x SU x S	33	52.39	1.59		
ACT x ED	6	6.81	1.14	0.81	0.567
ACT x ED x S	66	92.73	1.41		
INT x SU	2	4.43	2.22	0.47	0.631

INT x SU x S	22	103.63	4.71		
INT x ED	4	33.76	8.44	1.82	0.142
INT x ED x S	44	203.90	4.63		
SU x ED	2	8.92	4.46	0.74	0.487
SU x ED x S	22	131.91	6.00		
ACT x INT x SU	6	8.67	1.45	1.35	0.248
ACT x INT x SU x S	66	70.73	1.07		
ACT x INT x ED	12	21.49	1.79	1.63	0.090
ACT x INT x ED x S	132	144.92	1.10		
ACT x SU x ED	6	5.78	0.96	0.82	0.552
ACT x SU x ED x S	66	77.60	1.18		
INT x SU x ED	4	19.34	4.83	0.97	0.432
INT x SU x ED x S	44	218.50	4.97		
ACT x INT x SU x ED	12	5.42	0.45	0.56	0.873
ACT x INT x SU x ED x S	132	107.17	0.81		
Total	852	39497.78	12511.08		

Table 56. ANOVA summary table for time to complete the Bins task in Experiment 3.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	124697.53	11336.14		
Within-Subject					
Interactions (INT)	2	715.86	357.93	0.40	0.677
INT x S	22	19799.08	899.96		
Second User (SU)	1	848.35	848.35	0.66	0.432
SU x S	11	14056.28	1277.84		
Environmental Detail (ED)	2	942.58	471.29	0.18	0.839
ED x S	22	58542.24	2661.01		
INT x SU	2	2937.45	1468.72	0.98	0.390
INT x SU x S	22	32875.51	1494.34		
INT x ED	4	2199.84	549.96	0.52	0.720
INT x ED x S	44	46367.84	1053.81		
SU x ED	2	2342.80	1171.40	0.86	0.438
SU x ED x S	22	30091.87	1367.81		
INT x SU x ED	4	7693.60	1923.40	1.12	0.361
INT x SU x ED x S	44	75849.70	1723.86		
Total	204	295263	17269.68		

Table 57. ANOVA summary table for errors made during the Turns task in Experiment 3.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	63417.81	5765.26		
Within-Subject					
Interactions (INT)	2	1344.18	672.09	1.02	0.377
INT x S	22	14498.16	659.01		
Second User (SU)	1	14.52	14.52	0.01	0.926
SU x S	11	17727.15	1611.56		
Environmental Detail (ED)	2	2522.81	1261.41	1.71	0.204
ED x S	22	16203.85	736.54		
INT x SU	2	1004.34	502.17	0.73	0.495
INT x SU x S	22	15226.66	692.12		
INT x ED	4	4271.82	1067.96	1.43	0.239
INT x ED x S	44	32784.51	745.10		
SU x ED	2	650.48	325.24	0.41	0.666
SU x ED x S	22	17259.52	784.52		
INT x SU x ED	4	3931.16	982.79	1.70	0.167
INT x SU x ED x S	44	25401.18	577.30		
Total	204	152840.34	10632.33		

Table 58. ANOVA summary table for time to complete the Turns task in Experiment 3.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	37600.15	3418.20		
Within-Subject					
Interactions (INT)	2	311.69	155.85	0.85	0.442
INT x S	22	4043.36	183.79		
Second User (SU)	1	68.32	68.32	0.14	0.715
SU x S	11	5368.30	488.03		
Environmental Detail (ED)	2	1498.36	749.18	3.36	0.053
ED x S	22	4907.47	223.07		
INT x SU	2	514.00	257.00	0.93	0.411
INT x SU x S	22	6108.05	277.64		
INT x ED	4	918.16	229.54	1.60	0.191
INT x ED x S	44	6308.90	143.38		
SU x ED	2	745.86	372.93	1.66	0.213
SU x ED x S	22	4937.02	224.41		
INT x SU x ED	4	694.55	173.64	1.02	0.408
INT x SU x ED x S	44	7492.56	170.29		
Total	204	43916.6	3717.07		

Table 59. ANOVA summary table for the Search task in Experiment 3.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	412.22	37.47		
Within-Subject					
Interactions (INT)	2	15.94	7.97	0.72	0.500
INT x S	22	244.95	11.13		
Second User (SU)	1	1.42	1.42	0.11	0.752
SU x S	11	148.98	13.54		
Environmental Detail (ED)	2	25.39	12.70	1.17	0.329
ED x S	22	238.80	10.85		
INT x SU	2	26.94	13.47	0.98	0.391
INT x SU x S	22	302.49	13.75		
INT x ED	4	21.93	5.48	0.66	0.626
INT x ED x S	44	367.60	8.35		
SU x ED	2	45.31	22.65	1.77	0.194
SU x ED x S	22	281.59	12.80		
INT x SU x ED	4	76.32	19.08	1.09	0.371
INT x SU x ED x S	44	766.93	17.43		
Total	204	2564.59	170.62		

Table 60. ANOVA summary table for the Choice task in Experiment 3.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	6.35	0.58		
Within-Subject					
Interactions (INT)	2	0.12	0.06	0.48	0.626
INT x S	22	2.73	0.12		
Second User (SU)	1	0.04	0.04	0.57	0.467
SU x S	11	0.68	0.06		
Environmental Detail (ED)	2	0.29	0.14	1.16	0.333
ED x S	22	2.72	0.12		
INT x SU	2	0.03	0.02	0.08	0.924
INT x SU x S	22	4.24	0.19		
INT x ED	4	0.64	0.16	1.13	0.355
INT x ED x S	44	6.25	0.14		
SU x ED	2	0.10	0.05	0.39	0.683
SU x ED x S	22	2.91	0.13		
INT x SU x ED	4	0.32	0.08	0.55	0.699
INT x SU x ED x S	44	6.45	0.15		
Total	204	27.52	1.46		

Table 61. ANOVA summary table for presence in Experiment 3.

Source	***df***	***SS***	***MS***	***F***	***p***
Between-Subjects					
Subjects (S)	11	52.86	4.81		
Within-Subject					
Interactions (INT)	2	69.32	34.66	1.77	0.193
INT x S	22	430.13	19.55		
Second User (SU)	1	88.88	88.88	9.51	0.010
SU x S	11	102.77	9.34		
Environmental Detail (ED)	2	97.83	48.92	2.74	0.087
ED x S	22	393.37	17.88		
INT x SU	2	32.56	16.28	2.44	0.110
INT x SU x S	22	146.87	6.68		
INT x ED	4	162.39	40.60	3.05	0.026
INT x ED x S	44	585.40	13.30		
SU x ED	2	4.10	2.05	0.13	0.882
SU x ED x S	22	356.03	16.18		
INT x SU x ED	4	88.81	22.20	1.63	0.183
INT x SU x ED x S	44	598.48	13.60		
Total	204	3156.94	350.12		

VITA

Michael P. Snow

Michael graduated from Washington State University in 1985 with a B.S. in Psychology (minor in Military Science). He received his M.S. in Psychology (Human Factors option) from Purdue University in 1988. His research interests include human factors in virtual environments, design of visual displays, human-computer interaction, and human factors in aviation.