EFFECTS OF DISPLAY TYPE, AGE, AND GENDER ON DRIVING PERFORMANCE AND SIMULATOR-INDUCED SICKNESS IN A MEDIUM-FIDELITY DRIVING SIMULATOR

by

William James Penhallegon

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

MASTER OF SCIENCE
In
Industrial and Systems Engineering
(Human Factors Engineering option)

Dr. John G. Casali, Committee Chair
Dr. Gary S. Robinson
Dr. Tonya L. Smith-Jackson

June 30, 2003
Blacksburg, Virginia

Copyright © 2003, William J. Penhallegon

EFFECTS OF DISPLAY TYPE, AGE, AND GENDER ON DRIVING PERFORMANCE AND SIMULATOR-INDUCED SICKNESS IN A MEDIUM-FIDELITY DRIVING SIMULATOR

by
William James Penhallegon
Industrial and Systems Engineering

(ABSTRACT)

This study investigated the link between age and gender susceptibility to simulator-induced sickness in conjunction with display type. Simulator-induced sickness and ataxia were measured before and after exposure to a medium-fidelity driving simulator. Participants in four age and gender categories (older and younger males and females) operated the simulator with a consumer-grade head-mounted display (HMD), and then with a large screen, direct-view plasma display.

This study set out to recommend a particular display type that would be appropriate for use with particular age/gender groups in a general-purpose driving simulator. Unfortunately, practice effects affected the simulator-induced sickness and driving performance results for display type, which precludes making recommendations regarding the appropriate use of each display. Despite this, several important discoveries were made, including: 1) older participants did experience significantly increased simulator-induced sickness discomfort than the younger participants – regardless of display type; and 2) there was no significant difference found between genders in either simulator-induced sickness or driving performance; although females generally expressed a subjective preference for the direct-view display.

Display type was not found to affect the degree of ataxia experienced by participants; however, this study did find that although older participants exhibited significantly higher rates of simulator-induced sickness discomfort than the younger participants, they recovered their postural equilibrium significantly faster. This indicates that the older participants had greater difficulty adapting to the simulation environment than younger persons. It also suggests that younger persons are at greater risk during immediate post-simulation activities such as driving. Although it is likely that this effect would disappear over time, it has implications for agencies such as the Department of Motor Vehicles or drivers education schools that are considering the use of a driving simulator device before an on-road skills test.

Knowledge comes by the eyes,
Always open and working hand,
And there is no knowledge that is not power.
– Jeremy Taylor

ACKNOWLEDGEMENTS

Many people helped me throughout the course of this study and I would like to express my gratitude for their hard work and generosity. First, I would like to thank my graduate committee chairman, Dr. John Casali, for his guidance during the conduct of this research. His direction was extremely helpful - particularly in the experimental design and data analysis phases. Dr. Gary Robinson was invaluable for his direction in constructing the simulator and for his great patience in helping me plan my analyses and reviewing my work. I would also like to thank Dr. Tonya Smith-Jackson for her very worthwhile advice and for serving on my committee. Although not on my committee, I am indebted to Dr. Norman Warner of the Naval Air Systems Command for his guidance in the use of Statistica, and for taking time out of his very busy schedule to help me with the statistical analysis.

I would also like to extend my appreciation to Dr. Ron Mourant of Northeastern University for supplying the simulation software used in the experiment, and for accepting our recommendations for modifications in order to enhance its usefulness for research. Dave Smith of ATP, Inc. was very helpful in explaining the ins and outs of the simulator console, which greatly facilitated the hardware/software integration.

Many thanks must also be given to the personnel of the Virginia Tech Machine Lab for their invaluable assistance in getting the simulator up and running. Particular thanks go to Will Vest, Kelcie Bower, Jeff Snider, Phillip Ratcliff, Ronald Stables, and Randy Waldron for their support in mounting the display and hardware integration. They are a true asset to the university.

Special thanks to Chuck Perala for his great friendship over the last three years, for his contributions to the simulator, and for acting as a sanity check on my research design and execution. His worldview and generous nature continue to both challenge and inspire me. This process wouldn't have been nearly as successful – or nearly as fun – without him.

I would certainly be remiss without acknowledging the contributions and sacrifices my family made to help me get where I am today. To my mother, Ann, who taught me that love is what it's all about; to my father, Robert, who showed me the path; and to my brother, Douglas, who is the best friend and confidant a person could hope for – thank you. This work is for you.

And finally, I would like to thank my wife Sybille for her love and patience throughout the course of this study. She has always encouraged me to excel, and she greatly contributed to this work by keeping my spirits high and my mind focused. She makes me a better person in every way.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

INTRODUCTION

The simulation of flying and driving environments is becoming increasingly useful, not only the training of operators expected to perform these tasks, but also for researchers examining particular aspects of the tasks in a controlled manner. The advantages of simulation are many, including: increased safety, experimental control, ease of measuring and recording operator performance, reduced complexity and time for data collection and, in some cases, lower operational costs (Casali, 1980). Simulators allow operators to experience the consequences of their decisions in operating these vehicles in a controlled, relatively low-risk environment.

One of the primary goals of a simulator is the creation of a *convincing* level of similarity (presence) to a corresponding real world situation; however, creating this level of realism presents a major challenge. A simulator's degree of presence is dependent upon many factors, not the least of which is the visual display system used in the simulator. Head-mounted displays (HMDs) are becoming increasingly attractive to simulator designers as they have been shown to increase the sense of presence and improve operator task performance over traditional, desk-top displays. However, HMDs are also associated with greater stress on the visual system, which can induce motion sickness-like symptoms in some individuals in a syndrome often referred to as simulator-induced sickness.

Simulator-induced sickness is the tendency of many vehicular simulators, including both driving and flight devices, to induce acute, residual, and sometimes aftereffect motion sickness-like symptoms of discomfort in operators and passengers (Casali & Wierwille, 1986; Kennedy, Lane, Lilienthal, Berbaum, & Hettinger, 1992). Symptoms include, but are not limited to: general discomfort, apathy, drowsiness, headache, disorientation, fatigue, pallor, sweating, increased salivation, stomach awareness, nausea, and, in rare cases, vomiting (Kolasinski, 1995).

In addition, postural instability symptoms, or ataxia, have been known to last for several hours following simulator exposure (Baltzley, Kennedy, Berbaum, Lilienthal, & Gower, 1989).

The vast majority of simulator-induced sickness research to date in vehicle simulators has concentrated on the use of traditional direct-view displays. Very few studies have directly investigated simulator-induced sickness with HMD display systems in driving simulators and even fewer have evaluated the effect on driving performance. Although a few studies were found that directly compared HMD and direct-view displays, most investigated physical effects on the visual system and treated simulator-induced sickness as an afterthought. At least two studies compared the effect of display type on participant navigation in a virtual environment, but none directly compared the two display types in a driving simulator. Additionally, the literature is inconsistent regarding the effects of age and gender on susceptibility to motion and simulator-induced sickness and driving performance. Several studies have found that younger persons and females are more susceptible to motion sickness while others have called these conclusions into question. No studies were found that directly compared HMDs to direct-view displays in a diving simulator and measured the susceptibility of certain population groups to simulator-induced sickness – particularly with older persons.

Due to issues of safety, reliability, and practicality, the Virginia Department of Motor Vehicles (DMV) has expressed interest in supplementing or even replacing the on-road, driver license skills test with relatively inexpensive, portable, and small driving simulators to be located in DMV offices (see Appendix A). Virginia Tech has developed a low-cost, medium-fidelity, proof-of-concept, PC-based driving simulator for research purposes that could also potentially be used for driver education and even driver license testing. Northeastern University served as a subcontractor for roadway scene generation, and Carilion Biomedical Institute, through the

University of Virginia, provided funding to Virginia Tech's Auditory Systems Laboratory for the development project.

One design goal of the simulator was to have a small footprint to allow for easy transportability. Designers of modern-day driving simulators are increasingly turning their interest to HMDs as a way to increase the fidelity of the simulation and decrease the footprint size of the equipment. As senior citizens would be potential users of a DMV or driving education simulator, it is important to understand not only how their driving performance in a simulator would compare to that of a younger population, but also if the incorporation of an HMD would increase their susceptibility to simulator-induced sickness.

This research will compare the use of a head-mounted display with that of a direct-view display in a medium-fidelity, PC-based driving simulator to examine the differences in driving performance and simulator-induced sickness based on the age and gender of a population. Additionally, this study will examine ataxic effects (post-simulation postural disequilibrium) of HMD use in a driving simulator to determine if they are any more pronounced than those experienced with the use of a direct-view display. Few studies have had the benefit of a driving simulator that allows the interchangeable use of both a direct-view display and an HMD. The VT-UVA-Carilion simulator will allow drivers to perform the same driving tasks using both displays. The direct comparison of the two display technologies in the same driving simulator task makes this research unique.

The primary goal of this research is to determine how one display type affects the population expected to use the simulator as compared to the other display type. Each type, HMD or direct-view, may be usable for simulation applications in and of itself, but one may be more appropriate for certain population groups than the other. As a DMV or driver's education simulator would be intended for use by an entire population of automobile operators, the goal of this research will be to provide a general recommendation for the use of a particular display type, within certain parameters, for use with specific groups.

This research will be a first step in investigating any increased risk of HMD use in devices intended for use by the general public. The potential liability inherent in sickness and ataxic effects following exposure makes understanding and prediction of individual susceptibility very important. As HMD technology improves, the increasing use of immersive display systems in simulation as well as other professional and consumer applications require that operators understand any risks beyond those with traditional direct-view systems. The results of this research may help engineers and researchers build less discomforting, more effective (realistic), and lower-cost VR-based driving simulators for DMVs and other applications.

PRESENCE

A vehicle simulator places an operator into a virtual environment (VE). A VE consists of a computer generated model that allows the participant to interact intuitively in real time with the environment or objects within it and to some extent has a feeling of actually 'being there,' or a sense of *presence* (Wilson, 1997). The participant manipulates the controls of the simulator much as he or she would the actual vehicle and the corresponding perceptual feedback should closely resemble real-world cues. Simulators are often described in terms of their *degree* and *fidelity*. Degree of simulation refers to how much of a real-world environment is presented to the operator. Fidelity refers to the degree to which visual features in the virtual environment (VE) conform to visual features in the real environment (Rinalducci, 1996). This can be in terms of photorealism, spatial orientation, operator interaction, etc. Specific types of simulator fidelity are discussed in the following sections.

The effectiveness of a simulation, or the sense of presence the operator has in the simulated environment, can often depend on the fidelity of the simulator and the characteristics of the visual display system that is employed. Presence within a virtual environment is often described as the extent to which participants allow themselves to be convinced that they are in a world separate from where their physical bodies are located while experiencing the effects of a computer-generated simulation (Barfield, Zeltzer, Sheridan, & Slater, 1995; Slater & Usoh, 1993). A discussion of presence in simulation is important because, despite the limited systematic research available to substantiate the assumption, it is widely believed that the degree of presence experienced by an individual may influence his or her task performance in the VE (Barfield & Weghorst, 1993; Pausch, Shackelford, & Proffitt, 1993; Stanney, Mourant, & Kennedy, 1998; Zeltzer, 1992).

Barfield and Weghorst (1993) describe an organizational framework useful for examining virtual presence in terms of factors that are likely to influence the sense of presence in a virtual environment and indicators that may be used to determine the degree of presence a viewer is experiencing. This model is shown in Figure 1.

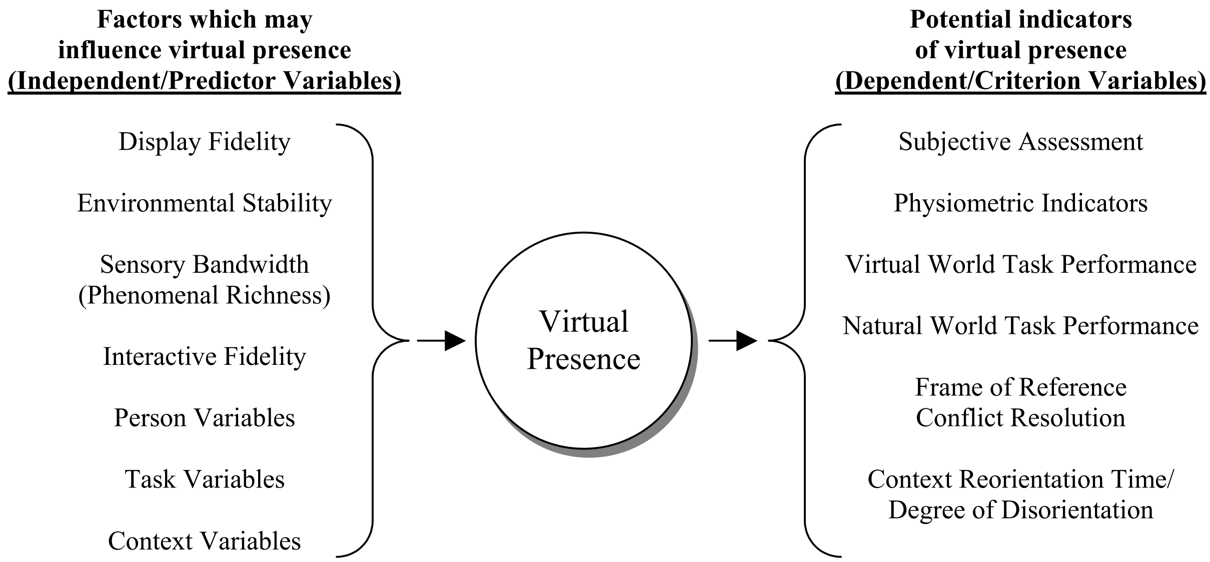

Figure 1. Barfield and Weghorst (1993) framework for exploring virtual presence

As explained in Barfield and Weghorst (1993), *display fidelity* refers to aspects of the simulation display system including, among others, spatial and contrast resolution, field of view, optical distortion, stereopsis, and other distance cues, lighting, and shading models. *Environmental stability* includes object attribute constancy and consistency. *Sensory bandwidth* describes the variety and dynamic range of the output displays and the kinesthetic and proprioceptive correlates of the participant's interaction behaviors. *Interactive fidelity* reflects the intuitiveness and ease of interaction within the virtual environment and includes items such as range and naturalness of control behaviors, display update rate, lag, and temporal and spatial predictability of the system response. *Person variables* are attributes unique to the participant: age, gender, adaptation, etc. Factors and indicators of particular importance to a driving simulator are discussed in the following sections.

Driving Simulator Display Technology

Modern driving simulators are generally comprised of the physical interfaces such as the seat and vehicle controls, the controlling computer system for scene generation and audio output, and the display system; a primary contributor to the sense of presence in a simulator. Early driving simulators employed various kinds of display technologies including a point-light source thorough transparency projection, film-based motion picture projection, model board objective with closed-circuit television, and infinity optics (Casali & Wierwille, 1986). Although many modern, high-fidelity driving simulators employ display techniques such as front or rear-projection, most medium to low-fidelity simulators present computer generated imagery (CGI) using direct-view displays such as cathode ray tubes (CRTs), liquid crystal displays (LCDs), or plasma displays.

Although there is overlap in the classification of driving simulators, there are three generally accepted categories based on cost and fidelity. At the low end are PC-based driving simulators that use standard, off-the-shelf hardware and software and can range in price from $2,000 to $10,000. Midrange driving simulators such as cab-based systems (Olsen, 1995; Weir & Clark, 1995) include mock-ups of vehicle interiors and physical controls (e.g. steering wheel, pedals, etc.). They can also include non-cab, highly specialized PC-based systems with custom built hardware and software. Depending on the degree of customization, these systems can cost in the tens of thousands of dollars. High-end systems such the National Advanced Driving Simulator (NADS) (Diewald, 1995) can cost millions of dollars. The display systems for any of these classifications can include commercially available desktop or head-mounted displays (HMD). Higher-end systems tend to incorporate high-fidelity HMDs or fully-immersive, 360° projections of the virtual driving environment and may also incorporate motion bases or other motion-cueing systems to increase the level of immersion an operator experiences in the virtual environment. See Appendix A for a brief history of driving simulators and concurrent display technology.

HMD-based driving simulators are not as prolific as direct-view-based systems due to several factors. These include: discomfort due to weight and balance of the display on the head (Cobb, Nichols, Ramsey, & Wilson, 1999), lags and delays with display update, the awkwardness of the cables attaching the HMD to the simulator (Nichols, 1999), and problems involved in representing physical controls and instrumentation in a VE. Despite this, HMDs are becoming increasingly attractive for use in simulators due to the reduced footprint, physical weight, and size of the overall simulator, and research continues in the application of fully-immersive HMDs in driving and flight simulation.

Direct-view displays. Direct-view displays encompass a wide range of display technologies including CRTs, LCDs, and plasma displays. Each has its own inherent advantages and disadvantages which will be discussed in the following sections.

<u>CRT displays</u>. Cathode ray tubes are a very common display technology and can be found in most televisions and computer monitors. Stuart (1996, pp. 131-132) describes their function:

> *A positive high voltage (commonly 15,000 – 20,000 volts) accelerates a beam of electrons emitted by an electron gun towards a phosphor-coated screen, and the point on the screen… hit by the beam emits visible light. The electron beam can be controlled and directed precisely through a focus system and a set of deflection coils that produce a magnetic field. When the phosphor that coats the screen is struck by the electron beam, the light it produces decays at a given exponential rate, so the phosphor must be struck by the electron beam very frequently in order for it to appear to be constantly illuminated. The frequency with which the entire screen image is redrawn is known as the refresh rate.*

CRTs are known for having a very bright, sharp, and stable picture at high refresh rates. CRTs can also achieve black levels that other types of direct-view display technologies find difficult to match. CRTs display black by the absence of a signal; those areas of the picture that appear black result from an absence of light coming through the black-tinted glass screen. The normally high luminance level of CRTs allow the picture to remain strong even when projected through the black-tinted glass (Brinkley, 2000).

CRT-based, direct-view displays large enough to be considered for use in low-to-medium-fidelity simulation have many other advantages. They are relatively inexpensive, easily obtained, and their effects on the visual system have been extensively studied and described in the literature (Bruno, 1993; Yeow & Taylor, 1989). However, they also tend to be bulky, heavy, and when used for simulation purposes, have limited field-of-view unless several displays are used simultaneously. Researchers often turn to CRT-based front projection devices to create large fields of view in high-fidelity simulations, but at least three are required to provide imagery in the periphery. In addition, front projection display systems tend to be large, expensive, difficult to mount, lower in luminance, difficult to calibrate, and noisy. They also tend to have high power requirements and heat output.

LCD displays. Instead of emitting light in the style of CRTs, liquid crystal displays control the light passing through them as emitted from a backlight. Stuart (1996, p. 133) describes their function:

> *LCDs consist of multiple thin layers, which include vertical and horizontal grid wires and polarizers, as well as the liquid crystal layer itself. The vertical and horizontal grid wires sandwich the thin liquid crystal layer and apply voltages that change the alignment of the molecules in the liquid crystal layer and hence their polarizing effect and the light that passes through them. The molecules maintain their alignment briefly after the voltage is withdrawn (analogous to the CRT's refresh rate), and voltage is applied to the grids in a raster-scan manner.*

Driven by the consumer electronics market, LCD development has exploded in recent years. In the mid-1990's LCDs suffered in comparison to CRTs due to lower resolution and brightness (Stuart, 1996). Now, high resolution, large-screen LCD panels are widely available as television sets and computer monitors, although they still tend to be more expensive than CRTs of comparable size. Despite their cost, LCDs offer many advantages over CRTs including lighter weight, lower power consumption, less heat generation, less variability, and good pixel registration (Stuart, 1996). It is expected that as prices decrease, LCDs will become more prominent than CRTs in a variety of different applications.

Plasma displays. A third major direct-view display technology that is becoming more prevalent is the plasma display. Brinkley (1998, p. 82) describes their function:

> *A plasma display is like a honeycomb sandwiched between two sheets of glass. Each cell in the honeycomb hold a neon-xenon gas mixture with a red, green, or blue phosphor coating on the front surface; a group of adjacent red, green and blue cells forms one picture element or pixel. A tiny, transparent electrode is attached to the front of the cell, and another electrode is attached to the back. When a voltage is applied to the electrodes, the gas molecules lose electrons; these electron-deficient molecules are called ions and the resulting gas is called plasma. In this process, the ions emit ultraviolet light, which excites the phosphor to glow in its characteristic color (red, green, or blue). …Unlike a CRT display which sweeps three electron beams across the back of the picture tube's faceplate to form each line of the image, a plasma display activates* all *of its pixels* simultaneously *60 times a second.*

Since becoming available only a few years ago, plasma displays have been slow to gain acceptance due to their relatively high cost, relatively low luminance, fragility, and limitations in displaying true black (Brinkley, 1998, 2000). The inefficiency of the electrical process is responsible for low luminance levels which significantly dims the picture as viewed through black-tinted glass (Brinkley, 2000). Consequently, the darkest areas of a picture that should be black are often seen as gray, which creates an image that appears to be washed-out. Recent plasma sets, however, cost between $8,000 and $16,000 and have increased the efficiency of the electrical process. They are much less dim than their predecessors and much improved in rendering black levels although they are still limited in black-level detail (Brinkley, 2000). Although subtle, the absence of detail in black areas of the image still creates an inferior picture as compared to a CRT. Despite the lack of black-level detail, plasma displays are inherently light and thin relative to their size, and are becoming increasingly accepted. As prices continue to fall, plasma displays are beginning to present themselves as an attractive option for simulator designers.

Head-mounted displays. Even with the availability of direct-view displays, simulator designers are increasingly turning to Head-Mounted Displays (HMDs) as a powerful new way to create an enveloping sense of presence within the simulation without the bulk and complexity of multiple CRT-based devices. The practical benefits of using HMDs in driving simulators are many. Because the only visual information that reaches the driver is that which is displayed on the HMD display screens, little effort is required to create a physical representation of the vehicle. The simulator does not require a console, doors, gauges, etc. that create the visual illusion of actually being inside an automobile. Vehicle physical characteristics are modeled within the software and the only contact the user has with the simulator hardware is tactile and auditory.

As a result, the physical simulator hardware will require a smaller footprint and may be less expensive to build. As the only physical contact the participant has with the hardware is tactile, basic controls such as the steering wheel, foot pedals, and turn-signal stalk must be provided. Nothing is required for the driver’s visual benefit. This will allow the simulator to be relatively lightweight and portable, and also provides an advantage in terms of flexibility for research use. Vehicle type can be altered in HMD-based simulators by modifying the software to present new imagery to the operator. Thus, they are not restricted to one type of vehicle as simulators with physical vehicle interiors tend to be. Depending on the HMD used for the simulator, hardware reduction may lead to a significant cost savings, however, these savings may be eliminated if a high-fidelity HMD is chosen.

A typical HMD consists of a helmet or mounting band, a display system (either miniaturized cathode-ray tubes or liquid crystal displays; each normally not larger than one inch square), and small optical elements which magnify, collimate, and project imagery via a mirror combiner into the eyes such that the original image appears at optical infinity (Barfield & Weghorst, 1993). A stereoscopic scene can be created with a display source for each eye and appropriate projection optics. The HMD can also provide the viewer with panoramic views covering a wide viewing angle which is difficult to achieve with conventional fixed (desk-top or panel mounted) displays (Ma, Hollerbach, & Hunter, 1993). Important characteristics of HMD devices that influence presence include, among other things, field-of-view, resolution, stereopsis, and headtracking.

Field of view. HMDs typically use active-matrix LCD display technology with a specific field of view (FOV). Field of view is the visual angle of the scene subtended at the viewer's eye (Barfield & Kim, 1991). Geometric Field of View (GFOV) is a subset of the participant's FOV and is defined as the FOV pertaining to the vertical and horizontal angle from the computer's virtual eye to the viewport (the part of the monitor screen displaying the image) which can be varied independently (Hendrix & Barfield, 1996; McGreevy & Ellis, 1986). Hendrix and Barfield (1996) describe the horizontal GFOV as the horizontal angle subtended at the center of the projection to the side edges of the viewpoint and the vertical GFOV as the vertical angle at the center of the projection to the top and bottom of the viewpoint. The displayed image can either be reduced or magnified by changing the horizontal and vertical GFOV while holding the size of the viewpoint constant (Hendrix & Barfield, 1996). Although the ideal instantaneous field of view for HMDs is the normal FOV of the unaided human eye (approximately ±100° horizontally and ±60° vertically), it is difficult to achieve with a large eye relief, a reasonable flat image field, and limited lens diameter (Ma et al., 1993). Hendrix and Barfield (1996) studied the effect of varying HMD horizontal GFOVs (10°, 50°, and 90°) on presence and concluded that participants' sense of presence, realism, and "correctness of proportionality" of the virtual objects was significantly greater with the 50° and 90° cases than with the 10°. Stanney et al. (1998) note that many VE tasks may require FOVs of 100° or more to achieve a feeling of immersion. Driving simulators, in particular, may require FOVs of greater than 100° due to the importance of peripheral cues (Levine & Mourant, 1995). This comes, however, at the expense of resolution.

Resolution. Display visual resolution can be measured in different ways, all of which address the spacing of individual picture elements (pixels). One measure is the number of pixels per degree of angle subtended by the display; another is the amount of visual angle taken up by a pixel or between adjacent pixels (Stuart, 1996). When discussing color HMD resolution, it is important to distinguish between individual pixels and pixel triads (red, green, and blue) used to produce a single, full color pixel. Roehl (1996) reviews three low-cost, consumer HMDs, each containing a pair of 180,000 pixel, 0.7-inch active matrix color displays. However, as "pixel" in the marketing refers to a pixel triad, the displays really only contain 60,000 pixels each for an approximate resolution of 300 x 200.

Roehl (1996) describes one of the most basic trade-offs in HMD design: field-of-view vs. resolution. Two HMDs with identical display screens have the same number of pixels but can produce different fields of view. Designers can force the HMD optics to produce a large field of view but this comes at the expense of resolution as a fixed amount of pixels are spread over a larger area. This results in an effect known as pixellation in which the pixels are magnified to the point at which they become visible. Wide field-of-view HMDs attempt to resolve this problem with the use of "depixellation" filters that soften the edges of individual pixels. The filters make the pixels appear to blend together which creates the appearance of a continuous image. The end result is a resolution that the wearer perceives to not be as sharp as an HMD with a narrower FOV. The HMD with a wider FOV has a softer perceived image due to the magnification and blending of the pixels but depending on the application, this may be an acceptable trade-off. Arthur, Hancock, and Chrysler (1993) refer to research that suggests that field of view is more important than image resolution for situations in which spatial orientation and cognition are paramount.

Stereopsis. Stereopsis describes the perceptual transformation of differences between the two monocular images seen by the eyes and is a functional component of depth perception (Stanney et al., 1998). Depth perception has been reported to affect virtual world performance (Ellis & Bucher, 1994; Ellis & Menges, 1995; McDowall, 1994) and is dependent upon whether a visual scene is static or dynamic (Russell & Miles, 1993). HMDs typically contain two displays; each presenting the same imagery to each eye (biocular display) or overlapping imagery in the binocular field of view (the portion of the visual field shared by both eyes (Stuart, 1996)). Partial binocular overlap (stereoscopic or binocular displays) can be used to achieve depth perception by horizontally displacing a monocular image; however, the degree of overlap required for different applications is currently not well understood (Stanney et al., 1998). A third possible viewing condition presents slightly offset views to each eye, but dynamically updates the computer graphic imagery to reflect the change in perspective normally perceived as a result of vergence movements when fixating on objects not at optical infinity (Ehrlich, 2000). This "vergence" viewing condition accounts for the tendency of the right eye to point leftward and the left eye to point rightward when viewing nearby objects and updates the graphic scene to present this perspective. Depth perception of dynamic scenes such as those in a VE of a driving simulator can be extremely complex; however, and are currently not well understood (Stanney et al., 1998).

Headtracking. Headtracking can significantly increase the sense of presence a person may experience in a virtual environment (Barfield & Hendrix, 1995). Headtracking provides input to the computer system that allows for the interactive update of visual (and auditory – when appropriate) displays in response to the position and orientation of the user's head (Stuart, 1996). Headtracking re-couples the kinesthetic senses and visual system and increases the perceptual fidelity of the system (Gigante, 1994). This provides for greater immersiveness in a VE without a wider FOV as the display dynamically updates to reflect the user's viewpoint. In addition, headtracking allows for motion parallax (when a user sees around the side of objects as the head is moved from side to side) which provides an extra visual cue that improves depth perception (Gigante, 1994). In a driving simulator, headtracking allows the operator to see peripheral imagery external to the vehicle when turning the head from side-to-side as well as the car interior.

As described in Stuart (1996), important headtracking issues include responsiveness, robustness, accuracy, resolution, and registration. Responsiveness refers to the degree of delay or lag in the system. Robustness is the capability of the tracker to function error-free in spite of false signals it might receive from the environment. Resolution refers to the smallest movement that can be detected by the system while accuracy describes how close the reported position is to the actual position. Registration is the relationship between real and reported position and orientation. These issues vary in importance depending on the task; however, for an HMD-based driving simulator, responsiveness, accuracy, and registration are particularly critical. These are discussed in more detail in the following sections.

Challenges of HMDs in Virtual Environments. Despite the advantages offered by an immersive display, there are numerous issues inherent to HMD technology that complicates the creation of a convincing sense of presence in a virtual environment. Of particular importance in a driving simulator are issues of responsiveness, accuracy, and registration. In addition, ergonomic issues and potentially harmful side effects force researchers to examine compatibility with the sensory systems in great detail. Research has examined problems with HMDs from a physical ergonomics standpoint (Nichols, 1999) to compatibility with the visual system (Barfield, Hendrix, Bjorneseth, Kaczmarek, & Lotens, 1995; Ehrlich, 2000; Mon-Williams, Wann, & Rushton, 1993; Neveu, Blackmon, & Stark, 1998; Stanney et al., 1998) to practical considerations of HMD use (Pierce, Pausch, Sturgill, & Christiansen, 1999). Wilson (1996) summarizes discussion in the scientific literature as well as the popular press of potential VE-related harmful side effects. Some, such as possible harmful visual effects, sickness and nausea, and disorientation, are significant and have been widely studied. Others, such as harmful musculoskeletal effects and behavioral changes have not been as widely researched and there is little hard evidence to suggest these areas pose significant problems (Wilson, 1996). A more detailed review of the significant, potential physiological problems of HMDs is discussed in later sections.

<u>Responsiveness</u>. As described in Nichols (1999), there are three main sources of lag in a VR system: system lag (related to the speed of the central processing unit of the computer); input device tracking lag (the time taken for the tracker to determine the input device position and relay the information to the computer); and display lag (the time taken to either update a new

image on the display, to refresh an existing image, or to change the image to reflect the movement of the user within the VR). A tradeoff exists between a fast update rate and the complexity of the computer graphics. The more polygons the computer is forced to render, the slower the update rate. In a VR-based driving simulator, system and display lag are as Nichols (1999) describes them; however, input device tracking lag is more appropriately expressed as headtracker lag. As the position of the user input devices (vehicle controls) remain fixed, the scene changes as the user moves his head. Delay occurs when the headtracker determines the new position and relays the new positional information to the computer. Specifically, headtracker delay stems from the time required to check sensors for data (sample rate), the setting of the number of computed positions per second (data rate), the number of new position coordinates reported by the tracking system to the host computer per second (update rate), and the delay between the movement of a remotely sensed object and the report of its new position (latency or lag) (Stuart, 1996). The cumulative effect of these delays constitutes *transport delay*, which this thesis defines as the time between the control input and the corresponding visual feedback to the operator.

Systems with transport delays of less than 100 milliseconds are generally considered interactive, or real-time (Pimentel & Teixeira, 1995). Wilson (1997) notes that the standard video update rate is 30 frames per second (fps), but rates of between 10 and 20 fps are common for relatively detailed and complex VEs. A 100 ms delay equates to 10 frames per second so any greater delay becomes noticeable as flicker and adversely affects users' performance in VR related tasks (Pimentel & Teixeira, 1995). One study found that subjective reports of presence in a virtual environment were significantly lower with update rates of five and 10 Hz as compared to rates of 20 and 25 Hz (Barfield & Hendrix, 1995). A reduction in the degree of presence can adversely affect the performance of tasks such as VE navigation. Wilson (1997) reports that 12 fps is probably the absolute minimum that will achieve recognizability in the world dynamics crucial to a sense of presence.

Accuracy and registration. A driving simulator coupled with an HMD can be included in the "mixed environment" class of virtual environment technology. A mixed environment simulator provides real controls to give tactile, force, and spatial feedback, however, these controls must be located via the HMD (Wilson, 1997). Mixed environments provide significant flexibility when aspects of the simulator need to be changed such as the interior of the vehicle, the sizes of the vehicle windows, instrument locations, etc.; however, care must be taken such that the sizes, proportions, and spatial locations of the controls as represented in the HMD are congruent with the same characteristics of the actual control hardware (Milgram & Kishino, 1994). Because actual body and vehicle parts including hands, steering wheel, turn-signal stalk, etc. cannot be viewed through the HMD, the operator must know where to reach for the control based on visual cues such as depth and spatial orientation as provided by the HMD. Mismatches can lead to a loss of the sense of presence, an increase in the mental burden and fatigue of the operator, and a possible decrease in performance (Sato, Kimura, & Abe, 1992).

Display and headtracker accuracy are important because of their effect on registration. Registration is critical for an HMD-based driving simulator task because even a small discrepancy between the position at which objects are displayed with respect to the real world can undermine the value of the system (Stuart, 1996). As the operator reaches for the steering wheel, for example, the approximate spatial location as displayed in the HMD should correlate to the real-world position of the control. If there is a mismatch, the sense of presence the operator has in the vehicle environment is diminished, and performance may decrease. Barfield and Hendrix (1995) suggest that spatial orientation is an even more important factor than photorealism in influencing the sense of presence in a simulation.

Accuracy of optical flow patterns is also extremely important for a virtual environment-based driving simulator (Stanney et al., 1998). The human visual system is very sensitive to any incongruities in perceived imagery and the smallest, almost imperceptible anomaly becomes very noticeable when visual flow field cues appear unnatural (Larijani, 1994 and Kalawsky, 1993 as cited in Stanney et al., 1998). The driving simulator operator becomes very aware that the experience is not “real” when motion in the virtual scene does not conform to expectations. Despite this, the human visual system can, over time, adapt to slight perceptual imperfections so that an adequate illusion is maintained for many users of current systems (Kalawsky, 1994). It is only when mismatches are beyond the tolerances of the human perceptual system that the illusion collapses.

Visual system stressors. Poor display compatibility with the visual system can lead to problems in users such as eyestrain, headaches, and blurred vision – symptoms collectively known as *asthenopia* (Hettinger & Riccio, 1992). Ehrlich (2000) describes several stresses that HMDs place on the visual system including: immovable screens that remain perpendicular to the user at all times, a relatively bright screen contrasted with the blacked out surround of the HMD housing, image distortion due to imperfections in the optics, limited color palette as compared to real-world imagery, lack of normal depth perception, and accommodation/convergence dissociation.

Accommodation/convergence dissociation is perhaps the most widely discussed VE issue that results from HMD-related stresses on the human visual system (Wilson, 1996). In most situations, the eyes tend to turn inwards together to focus on near objects (convergence) and focus (accommodate) to see the object clearly. “Accommodation and convergence are intrinsically linked, and in everyday vision the eyes will both accommodate and converge for the same distance” (Wilson, 1996, p. 45). When viewing a scene with an HMD, the cues used by the eyes to select an object to focus upon are not well understood. The eyes may converge to the distance of the screen or to the depth of a fixated object in the visual presentation as determined by its linear perspective (Ehrlich, 2000). Ehrlich (2000) presents evidence to suggest that the eyes do not converge to the distance implied by the linear perspective which suggests that virtual environments can lead to improper vergence in the visual system. Accommodation/convergence dissociation, in addition to the various other stressors, serves to create oculomotor discomfort in the wearer and may mitigate against the use of HMDs by various members of the population.

A fundamental study in the area of HMD effects on accommodation and vergence gave conventional ophthalmic tests of binocular function to 20 participants following VE immersion of 10 minutes with an early generation binocular HMD (Mon-Williams et al., 1993). The authors reported changes in visual function for more than 50% of the participants, which they ascribed to induced binocular stress. Among the noted effects were changes in distance heterophoria (the tendency for one or both eyes to wander away from the position where both eyes are looking together in the same direction (Okhravi, 1997)) for which the authors ascribe to demands on the accommodation/vergence system. Participants also noted short-lived effects including blurred vision, headaches, and nausea. The authors then performed a similar experiment with a newer generation biocular HMD that allowed for IPD adjustments, independent eye focus, lesser temporal lags, and higher screen resolution (Rushton, Mon-Williams, & Wann, 1994). They reported reduced visual performance decrements and symptomatic effects and noted that a crucial, although not the only, factor is the biocular rather than the binocular display.

However, at least two studies found that visual effects were not significantly different between an HMD and a traditional desk-top display. Neveu et al. (1998) reported no significant ocular effects between a biocular HMD and a large-screen television after 13 participants viewed a two-hour motion picture. The authors note, however, that the task was a passive viewing exercise and results may change with a more active environment, a binocular HMD, and the addition of headtracking.

In a second study that directly compared two display types using the same participants and software, Peli (1998) measured functional changes in binocular vision, accommodation, and resolution following 30 minutes of immersion in both stereoscopic (binocular) and non-stereoscopic (biocular) modes and compared them to changes following the same computer gaming task performed on a desk-top CRT display. The study did find that, unlike Mon-Williams et al. (1993) and Rushton, et al. (1994), there was no significant difference in visual effects between the HMD in either biocular or binocular mode. The author notes, however, that a stereo display that presents larger disparities and a task requiring repeated oscillations between large distances may induce larger effects. The study also found that the HMD in binocular mode scored significantly lower in subjective participant comfort than the CRT display. Overall, however, the study supported the results of Neveu et al. (1998), and concluded that neither harmful nor statistically significant changes to the visual system are associated with the HMD in either biocular or binocular mode, although participants subjectively reported that the binocular condition was less comfortable relative to the use of a desk-top CRT display.

Physical ergonomics. Nichols (1999) discusses several of the physical ergonomics problems involved with HMD use in any application. Because of weight, bulk, and mounting constraints on the head, the displays must be limited in size. This creates a trade-off between display quality and weight, as displays with higher resolutions tend to be heavier. Other considerations are weight distribution (some headsets tend to be front-heavy due to the weight of the optic display equipment) and HMD contact with the head. There is also the problem of enclosing the user within some HMD designs. The headset can become hot which can cause the participant to perspire (Bauer et al., 1996 as cited in Nichols, 1999). This has implications for the sanitary state of the headsets; particularly if they are to be used in a situation in which the general population is expected to wear them, as in a DMV driver-licensing situation.

Stuart (1996) also describes how the physical design of the HMD must take into account the head geometry of the user. In order to maintain its placement, the HMD must be adjusted to fixed points on the head such as the cranium, nose bridge, etc. The HMD must also provide a 10 mm clearance for the user's eyelashes and as well as a five to ten mm additional eye relief to accommodate facial geometry in the "forward" direction. Even more clearance must be provided for eyeglasses wearers.

Presence as A Function of Display Type

Because presence is only at a very early stage in terms of analytic and conceptual models, it can be difficult to measure and quantify. A large number of interacting factors might influence presence including: display fidelity, environmental stability, sensory bandwidth, interactive fidelity, and characteristics of the individual, task, and context (See Figure 1). Difficulty in measuring presence is likely due to the multiplicity of issues involved, the difficulty in manipulating some of the relevant variables, and the absence of agreed methodology and measures (Wilson, 1997). Despite this, several studies have attempted to evaluate presence both objectively and subjectively (Barfield & Weghorst, 1993; Hendrix & Barfield, 1996; Robinett, 1992; Slater & Usoh, 1993; Slater, Usoh, & Steed, 1994; Witmer & Singer, 1994, 1998; Zeltzer, 1992). As adapted by Hendrix and Barfield (1996), Jex (1988) outlined types of objective measures to include task demands (task performed, performance, and priorities) task results (performance measures, errors, incidents, etc.), and correlated measures (gross motor activity, psychophysiological measures, socially conditioned responses postinteraction effects). The most common measures of presence are subjective and include on-line reports, post-test evaluations (questionnaires), and explanations of "high-presence" events.

Some studies have suggested human response to unexpected or threatening stimuli as a measure of presence. These could include "startle responses" such as flinching when a virtual object, such as a swinging baseball bat, is aimed at the participant (Held & Durlach, 1987, 1992). In the case of a driving simulator, presence could be indicated by the urgency with which an operator acts to avoid an imminent collision. Other measures might include socially conditioned responses such as unpremeditated grasping for an object, shaking hands, or responding to other beings in the virtual environment (Sheridan, 1992).

The degree to which different display methods vary in creating a sense of presence is unclear. Although several studies have examined presence in terms of a single display type, relatively few have directly compared the sense of presence and operator performance with the use of traditional desk-top (direct-view) displays vs. head-mounted displays. This could be due to the difficulty in controlling for the large number of differences between the two display types stemming from disparities in resolution, field of view, and input device. To date, interface comparisons usually involve a more qualitative analysis and are dependent on the particular application. Barfield and Weghorst (1993) propose that the organizational framework for exploring virtual presence, as shown in Figure 1, can provide a set of trade-off criteria to compare interface approaches for a specific application. For example, the authors suggest that immersive (HMD-type) interaction can provide greater sensory bandwidth and interactive fidelity, but desk-top displays have better display fidelity and environmental stability. Weighting of these factors is ultimately dependent on person, task, and context variables.

Despite the limited amount of research that directly compares HMDs to direct-view displays, Held and Durlach (1992) have suggested that HMDs provide a greater sense of presence in a virtual environment and thus, enhance performance of some tasks over traditional direct-view displays. By creating a strong illusion of presence inside a virtual space, HMDs increase task performance by allowing users to interact more naturally and directly with the simulated environment and by minimizing outside distractions which increases user focus and concentration on a task (Witmer, Bailey, & Knerr, 1996). One study (Ruddle, Payne, & Jones, 1999) found that participants developed a significantly more accurate sense of relative straight distance in the VE, faster environment navigation, and increased saliency of landmarks. They also discovered that participants looked around more when using an HMD than the CRT indicating that the HMD increased the sense of immersion in that environment and people tended to perform actions that came more naturally to them.

Hendrix and Barfield (1996) examined presence in an HMD as a function of visual display parameters which included presence or absence of headtracking, presence or absence of stereoscopic cues, and the geometric FOV used to create the visual image projected on the visual display. With the use of a questionnaire, the authors found that presence was increased with the use of stereoscopic cues, a headtracker, and a wider geometric field-of-view. The authors concluded that the one component absolutely critical to the operator's sense of presence within an environment is the ability to perceive spatial information about that environment. They suggest that presence is not necessarily defined by the photorealism of the representation of the objects in virtual environment per se, but is more dependent on the realism with which the user is able to interact with the virtual environment.

A sophisticated driving simulation is composed of both tracking and search tasks. The driver must track the dynamically changing position of the road relative to the vehicle and adjust the controls accordingly. In some cases, however, the driver must also search for objects within the virtual environment such as speed limit and road signs, alternate routes, destinations, etc. Pausch et al. (1993) found that users performing a generic search task decreased task time by 42 percent when they changed from a stationary display to a head-tracked HMD with identical properties (resolution, field of view, etc.). They also noted a transfer effect in that users who practiced with the HMD reduced task completion time by 23 percent in later trials with the stationary display. The authors speculate that the search time reductions were due to the use of headtracking that allowed the participants to more readily determine where they had and had not already searched resulting from a better internal representation of the environment. Although the study used a stationary HMD instead of a traditional desk-top display, their results demonstrate the importance of headtracking on task performance in a generic search task.

VR Research With Older Persons

Graphical computer interface, virtual reality, and simulation technology have, to date, only seen very limited application with an older population (Liu, Watson, & Miyazaki, 1999). With personal computers becoming increasingly affordable and useful in helping to perform everyday tasks such as shopping and banking, more and more senior citizens are interacting with the technology. Due to age-related deterioration in cognitive, psychomotor, and perceptual abilities (Guerrier, Manivannan, Pacheco, & Wilkie, 1995), older persons are finding greater difficulty interacting with technology than younger persons. An examination of the different requirements for user interfaces for the older population has only recently begun, but recent studies have been examining how older persons use technology and what design criteria are important to assist them (Czaja, 1996; Mead, Spaulding, Sit, Meyer, & Walker, 1997; Worden, Walker, Bharat, & Hudson, 1997).

Limitations within HMD technology may impede the use of virtual environments with older persons. Morgan (1986) describes how diminution of sensory input from the eyes experienced by these individuals may impede visual perception. Older persons tend to show impaired visual performance at low levels of ambient luminance, primarily resulting in a diminution of contrast sensitivity at intermediate and high frequencies, as well as a decreased sensitivity in color perception (Morgan, 1986). Reduced contrast sensitivity and lower visual acuity limits sensory input from a virtual environment which can present difficulties to older users in navigation and manipulation tasks in virtual worlds (Stanney et al., 1998).

Ehrlich (2000) notes that task demands in a VR environment may require the eyes to shift from their normal resting states leading to fatigue in the accommodative and vergence muscles. Ocular changes due to age include the thickening and stiffening of the lens making it less dense and less refractive (Koretz, Handelman, & Brown, 1984; Morgan, 1986; Neveu et al., 1998). As refractive power determines the range of accommodation, accommodation decreases with age until it nearly disappears (Neveu et al., 1998). The added accommodative and vergence demands on the eye create significant obstacles to the use of HMDs in an older population.

SIMULATOR-INDUCED SICKNESS

One major problem of simulators that provide a compelling sense of presence and self-motion is the potential occurrence of 'simulator-induced sickness' which is the tendency of many vehicular simulators, including both driving and flight devices, to induce acute, residual, and sometimes aftereffect symptoms of discomfort in operators and passengers (Casali & Wierwille, 1986; Kennedy et al., 1992). Operator discomfort or sickness occurring in a simulator and not in the actual vehicle that it is designed to replicate is an indication that the sickness is inappropriate and that the simulator is inadequate (Frank & Casali, 1986). Simulator-induced sickness symptomology varies widely among individuals and the simulators that cause it, but the most common symptoms resemble those of motion sickness: general discomfort, apathy, drowsiness, headache, disorientation, fatigue, pallor, sweating, salivation, stomach awareness, nausea, and vomiting (Kolasinski, 1995). In addition, sudden post-exposure effects such as flashbacks (sudden recurrence of symptoms), visual illusions, spinning sensations, disorientation, and ataxia (postural instability) have also been known to occur (Casali, 1986).

Simulator-induced sickness is called a syndrome because of the complex signs and symptoms associated with it (Kennedy & Fowlkes, 1992). Because of the associated diversity of symptoms, Kennedy and Fowlkes (1992) also referred to it as being polysymptomatic. They noted that people often show different degrees of the sickness, from showing nearly all the signs and symptoms to very few to none, and that no single symptom dominates. An often used tool for subjectively evaluating simulator-induced sickness is the Simulator Sickness Questionnaire (SSQ) which divides symptoms into three distinct clusters: oculomotor (eyestrain, difficulty focusing, blurred vision, headache), disorientation (dizziness, vertigo), and nausea (nausea, stomach awareness, increased salivation, burping) (Kennedy, Lane, Berbaum, & Lilienthal, 1993).

As discussed in Frank and Casali (1986), the occurrence of simulator-induced sickness presents many problems to the researcher attempting to gain valid data from a simulator-based experiment as well as to the trainee using it to learn skills for operating a specific vehicle. First, simulator-induced sickness constitutes an extraneous source of variance because it does not correspond to responses observed in the actual system. This threatens the validity of the simulation and the generalizability of the experimental data. Second, simulator-induced sickness may reduce transfer of training for persons attempting to learn to effectively operate a specific vehicle. To reduce discomfort, trainees may adopt certain behaviors in the simulator that when transferred to the actual vehicle, would be inappropriate or even dangerous. Third, experimental participants or trainees may be reluctant to take part in a simulator session that they understand might make them ill. There is also the danger of ataxia: post-simulation effects where the participant or trainee experiences postural disequilibrium, nausea, or disorientation in the hours following the session (Baltzley et al., 1989). This can create a serious hazard when operating a vehicle in the real world such as driving a car or flying an airplane. Finally, there are the ethics of making experimental participants ill for research purposes. The fact that a simulator may induce sickness should be made evident to participants; however, this may bias the results of the experiment. The experiment should be carefully constructed to reduce or eliminate bias resulting from participant expectations.

Simulator-Induced Sickness vs. Motion Sickness

It is believed that the first documented report of simulator-induced sickness was the 1957 Havron and Butler study which noted that 78% of questionnaire respondents indicated that they had experienced some degree of "motion sickness" while operating the 2-FH-2 Bell helicopter hover trainer (Casali, 1980; Frank & Casali, 1986). There is considerable debate regarding whether or not motion sickness and simulator-induced sickness are different phenomena.

Despite the similar symptomology, many researchers have argued that the two should not be used synonymously (Casali, 1980; Kolasinski, 1995; Pausch, Crea, & Conway, 1992). They suggest that the term “motion” in “motion sickness” implies that physical movement of the participant is necessary and solely responsible for any sickness produced. But because sickness occurs in people operating fixed-base simulators with no translational or rotational movement of the participant, it must be precipitated by factors other than motion (Casali, 1980). Kolasinski (1995) notes that simulator-induced sickness is more likely a result of the compounding of the visual and motion cuing and not due to merely the motion alone.

Others feel, however, that simulator-induced sickness is merely a subset of motion sickness (Benson, 1987; Ebenholtz, 1992). Benson (1987) notes that, apart from similar signs and symptoms, the adaptation effect as well as severity of symptoms linked to exposure indicates that simulator-induced sickness is another form of motion sickness. Ebenholtz (1992) notes that direct inertial stimulation of the vestibular apparatus is not always necessary to produce motion sickness. He argues that motion sickness in virtual environments results from “any condition yielding error in eye movement control along with the ensuing feedback and error-correcting signal” (Ebenholtz, 1992, p. 303).

It is the position of this thesis that this is a largely a matter of definition, however; the two phenomena should be considered different. Despite the similarity of symptoms, simulator-induced sickness shows greater oculomotor-related symptoms and less actual vomiting than other forms of motion sickness (Kennedy et al., 1992). This paper agrees that “motion sickness” does imply motion of some sort, and that vestibular stimulation is an inherent component. While simulator-induced sickness has been known to occur in motion base simulator devices, it is unlikely that motion cues play more than a minute role in causing the syndrome (Casali, 1980). One study found that helicopter pilot operators of both motion-based and fixed base simulators became equally sick (McCauley & Sharkey, 1992). As simulator-induced sickness often occurs in fixed base devices, it must be tied to visual stimuli and oculomotor control.

Causes / Theories of Simulator-Induced Sickness

*Structure of the vestibular system***.** A discussion of the theories of simulator-induced sickness must begin with a brief review of the vestibular system. Stuart (1996) and Yoo (1999) provide a description of the structure and functioning of the vestibular system which is paraphrased as follows. The vestibular system senses movements of the head such as rotation, tilting, and linear acceleration relative to the force of gravity and humans are generally only conscious of the vestibular system when normal functioning is disturbed (Stuart, 1996). Head rotation is sensed by three pairs of vestibular canals: non-auditory, labyrinthine organs embedded in the bone on each side of the head. The canals are filled with endolymph fluid which, during head rotations, lags behind on the canal walls in the direction opposite that of head rotation. Members of a pair are approximately but not exactly parallel, and pairs are approximately but not exactly at right angles, thus no plane of head rotation stimulates only one of the pairs. The canals are sensitive only to a change in acceleration or deceleration rate; not to body movement at a constant velocity (Yoo, 1999).

The otolith organs, connected to the vestibular canals and also filled with endolymph fluid, consist of two small, sack-shaped organs (the utricle and the saccule), which sense body posture relative to the vertical. Cilia on the membranes (macula) of both otolith organs protrude into a gelatinous structure (statoconia) that contains calcite crystals. The statoconia has a higher specific gravity than the surrounding fluid, and is displaced by tilting or linear acceleration of the head (Stuart, 1996). Although the primary function of the otolith organs is to sense postural conditions of the body, their sensitivity to acceleration or deceleration likely assist the vestibular canals in sensing these changes (Yoo, 1999).

Cue conflict theory. In the absence of a comprehensive model of simulator-induced sickness, the primary theory of why the phenomena occurs is the cue conflict theory, also known as the theories of perceptual conflict, sensory conflict, sensory rearrangement, perceptual decorrelation, and neural mismatch (Casali & Frank, 1986; Kolasinski, 1995). The underlying premise of cue conflict theory is borrowed from Reason and Brand's (1975) motion sickness research and is based on a lack of correlation between appearance and reality (Kennedy & Frank, 1985). The theory posits that humans have a neural store of expected sensations that reflect past experiences which are built up over time and become more salient with continuing motion experience (Casali & Frank, 1986). Under ordinary circumstances, there is a correspondence between what is sensed by the vestibular apparatus, what is viewed by the eyes, and what is expected by neural system wiring. Conflict occurs when a person is placed into an environment in which motion information, as signaled by the eyes, is at odds with the gravito-inertial force as sensed by the vestibular apparatus and what is expected by the neural store.

Cue conflict occurs in a fixed-base driving simulator when, for example, the neural store of a vehicle operator leads him to expect acceleration forces pushing him into the seat as relayed by the vestibular and haptic sensory systems. However, in a fixed-base simulator, the only acceleration cue perceived by the operator is through the visual system. In this case, the lack of stimulus input to the vestibular system creates a visual/vestibular intermodality conflict and discomfort can result. Visual/vestibular conflicts can also occur when inputs have different degrees of delay. In the case of a motion-base driving simulator, depending on the simulator design and computing power, motion can lead the visual stimulus or vice-versa. Simulator-induced sickness symptoms can occur depending on the order and severity of the delay (Frank, Casali, & Wierwille, 1988). Additionally, these conflicts may also lead to a disruption in balance and coordination, resulting in ataxia (Fregly, 1974). Figure 2, slightly modified from Yoo (1999), graphically presents the visual/vestibular conflict in fixed-base simulators.

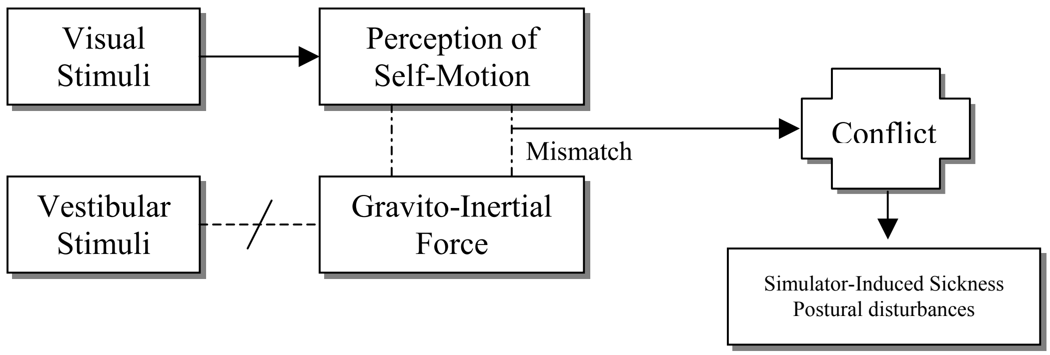

Figure 2. Perceptual Conflict Model of simulator-induced sickness in fixed-base simulators

Cue conflict theory originally tended to concentrate on the lack of intermodality correlation such as between visual and vestibular inputs (Casali & Frank, 1986). However, conflict can occur either between sensory modalities or within a sensory modality. In attempting to develop an explanatory principle for space sickness, Guedry (1970) described situations that potentially lead to vestibular/vestibular conflict within the canals and otoliths. Drawing on the work of Leibowitz and Post (1982), Frank and Casali (1986, p. 2) describe a visual/visual conflict between focal and ambient visual systems. They provide the example that, “a simulator operator’s focal visual system, which is concerned with recognition and identification, may receive stimulation from video images written in one direction, whereas, the ambient visual system, which is concerned with spatial orientation, may be receiving video information written in another direction.” Although visual/vestibular conflicts are usually the most obvious, the considerable amount of other potential conflict combinations that can occur makes prediction of simulator-induced sickness problematic.

Despite the power of the cue conflict theory to explain how simulator-induced sickness can arise, at least three constraints exist limiting its usefulness. First, no good method exists within the model to determine the magnitude of the conflict for specific conflict combinations (Frank, Kennedy, McCauley, & Kellogg, 1983). Second, as discussed above, research is limited with respect to conflicts within a sensory modality. And third, the model has very little predictive power with respect to simulator-induced sickness. This is not necessarily a fault of the model, but is more related to the paucity of systematic research on the numerous relevant variables (Frank & Casali, 1986).

Stoffregen and Riccio (1991) suggest other weaknesses of the cue conflict theory. They argue that the theory provides no principled basis to distinguish between nauseogenic and nonnauseogenic situations and traditional theories of perception are based on static environments, which ignore action as a critical role as dynamic stimulus for perception. They also assert that expectation of redundancy cannot plausibly result from interaction with the environment and question the assumption that redundancy among/within the visual, vestibular, and proprioceptive senses is expected. As this redundancy is not necessarily expected, cue conflict cannot provide a theoretical explanation for the existence of motion sickness (Stoffregen & Riccio, 1991).

Ecological theory. Posture is maintained by the tight coupling of the vestibular and visual systems, and the vestibulo-ocular reflex (the tendency of the eyes to move in the opposite direction of head rotation) causes observed objects to maintain their perceptual constancy during movement (Stuart, 1996). Posture is naturally unstable and effort is required for actions that minimize uncontrolled movements. Postural stability is the state in which uncontrolled movements of the perception and action systems are minimized (Riccio & Stoffregen, 1991). As proposed by Riccio and Stoffregen (1991), the Ecological Theory of Simulator Sickness suggests that sickness occurs in situations in which the individual does not possess or has not yet learned strategies for maintaining effective postural control. The authors argue that prolonged postural instability leads to a greater likelihood of and intensity of simulator-induced sickness symptoms. Figure 3, originally presented in Yoo (1999), presents a simplified model of the Ecological theory.

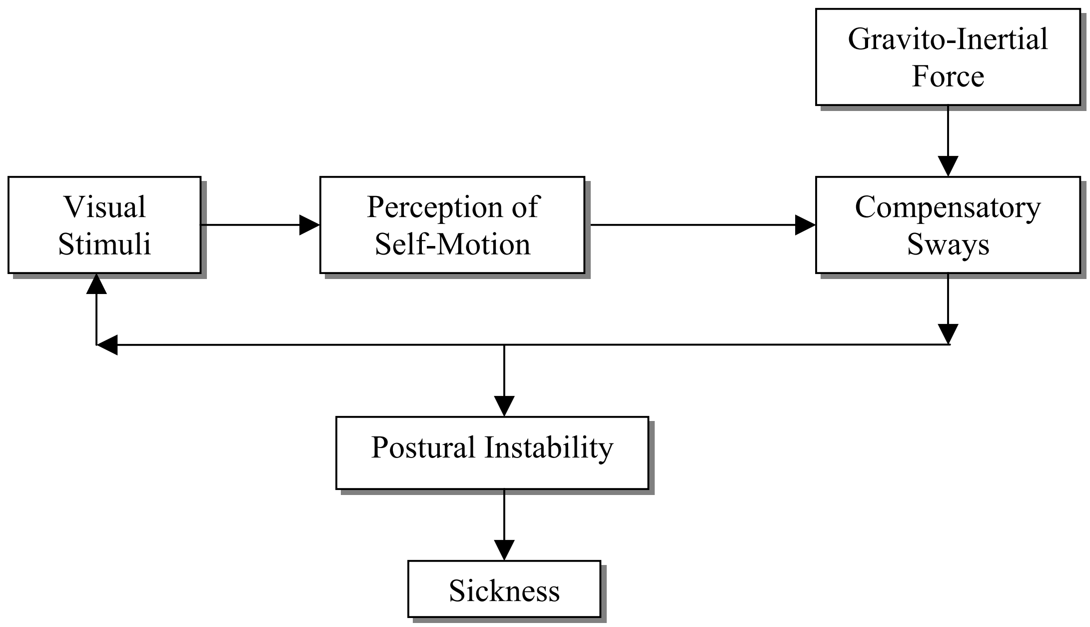

Figure 3. Ecological Model of simulator-induced sickness in fixed-base simulators

S. A. Jones (1998, p. 37) provides an explanation for how postural instability causes sickness in a fixed-base driving simulator:

> *The perception/action cycle is exemplified in a driving simulator. According to Riccio and Stoffregen (1991), visual perception of the gravito-inertial force vector implied via optokinetic stimulation (a perception) leads to compensatory postural adjustments (an action) in order to align posture with the perceived gravito-inertial upright. In fixed-base conditions such as the driving simulator used in the present study, the absence of somatosensory/kinesthetic motor feedback during attempts to maintain postural upright in accord with the visually-implied gravito-inertial force results in erroneous perception of orientation. A maladaptive feedback loop is initiated by leaning into an anticipated force that isn't there. The result is postural instability, which, if prolonged, plays a causal role in producing sickness (Riccio & Stoffregen, 1991).*

Based on this proposed relationship, S. A. Jones (1998) hypothesized that postural restraint should reduce or even eliminate sickness by eliminating the demands on postural control. She found, however, that the Ecological Theory of Motion Sickness and Postural Instability was not supported by the results and that there was no significant reduction in simulator-induced sickness when participants were restrained. S. A. Jones (1998) explains that the presence or absence of postural instability for any given participant is nearly impossible to ascertain and that in their experiment, it was not directly measured per se. The author admits that her results may have been due to lack of direct measures of postural instability, inadequacy of the restraint or lack of power in the experimental design and analysis.

Ataxia

Participants who exhibit severe sickness symptoms during simulator exposure are at greater risk for posteffects (Kennedy, Lilienthal, & Fowlkes, 1991). In their review of simulator-induced sickness occurrences following training sessions in Army and Navy flight simulators, Baltzley et al. (1989) divided observed symptoms into three categories: visuomotor (eye fatigue, strain, difficulty focusing, etc.), disorientation (dizziness, vertigo, and balance problems), and nausea (burping, stomach awareness, and nausea). Of 742 participants, 45% reported experiencing simulator-induced sickness symptoms; 25% of symptoms lasted more than one hour following exposure, and 8% lasted more than six hours. Of particular significance were the reports of post-simulation disorientation, which have the greatest potential for causing personal injury in participant activities (such as driving) following the simulation. 5% of pilots indicating simulator-induced sickness symptoms reported disorientation for more than six hours following simulator exposure. Another study using Air Force pilots found that 60.4% of participants reported ataxia immediately following simulator exposure and in 14.6% of the cases, disequilibrium persisted as long as 30 minutes to 10 hours (Kellogg & Gillingham, 1986). Kennedy and Stanney (1996) note that there is a rapid initial readaptation period for ataxia of roughly three hours in which approximately 50% of normal functioning is achieved. Complete readaptation normally occurs within 100 hours.

Thomley, Kennedy, and Bittner (1986) have suggested that ataxia is due to a disruption in balance and coordination resulting from perceptual cue conflict during simulator exposure. The posture system is highly flexible which enables individuals to adjust to a variety of environmental conditions (Fowlkes, Kennedy, & Lilienthal, 1987). Adaptation of the visual and vestibular systems to the simulator environment disrupts balance and coordination after return to a "normal" environment (Thomley et al., 1986) and there is evidence to suggest that intensity and duration of ataxia increases with greater simulator exposure (Crosby & Kennedy, 1982; McCauley & Sharkey, 1992). Participants who have adapted to simulator environments have reported a decrease in simulator-induced sickness symptoms and a concurrent worsening of post-exposure ataxia (Kennedy, Lanham, Drexler, & Lilienthal, 1995). This has important implications for simulators in which the general public is only expected to operate for a limited amount of time. A brief discussion of validated and sensitive dependent measures can be found in the Experimental Methodology section.

Simulator-Induced Sickness in HMD-Based Virtual Environments

Any environment that produces a compelling sense of presence and self motion in the absence of physical cues in an individual is vulnerable to eliciting simulator-induced sickness-type symptoms (Kennedy et al., 1992). Virtual environments that use head-mounted displays produce similar discomforts as other simulation technologies that immerse the user in a wide-view, illusory experience (Biocca, 1992). Although similar in many ways to motion and simulator-induced sickness, some researchers feel that *cybersickness*, or Virtual Reality Induced Symptoms and Effects (VRISE) has a sufficient enough etiology to be considered a new, although related, phenomenon. Wilson (1997) reports that across the range of VR studies completed up to that time, (12 experiments, 223 participants), approximately 80% of participants reported some increase in symptoms. For most, the symptoms were mild and short-lived but for 5% they were so severe that the participants had to end their participation.

One study differentiated “cybersickness” from simulator-induced sickness based on the predominance of particular symptom clusters (Stanney, Kennedy, & Drexler, 1997). After a comparison of simulator-induced sickness self-reports of operators of military flight simulators to simulator-induced sickness self-reports of users of VE-based simulators, the authors found greater oculomotor symptoms in the former, and greater disorientation symptoms in the latter. They conclude that this profile difference is sufficient enough to differentiate the two maladies are and classify them as different *strains* of motion sickness.

Although there is little other quantitative research that differentiates VRISE from simulator-induced sickness, experiments that manipulated type of visual feedback provided along with certain vestibular cues indicates that like simulator-induced sickness, sensory conflict may be responsible for some VRISE (Wilson, 1997). However, as with simulator-induced sickness, the sheer numbers of potential factors and their interactions precludes a full examination of all combinations. Despite this limitation, research using fractional-factorial designs has been suggested as a way of identifying the most relevant variables to the examination of VRISE (Stanney et al., 1998).

In contrast to the advantages offered by HMDs in terms of increased presence and lower simulator hardware complexity, it is possible that HMD use may increase the incidence of user discomfort over that of direct-view displays – particularly as a result of the additional weight on the head during head movements and stress on the visual system. DiZio and Lackner (1992) observe that since vestibular function appears to be a causal factor in motion sickness, space motion sickness investigations have focused on changes in otolith function resulting from null gravitoinertial force background, otolith-mediated alterations in other orientation subsystems, and rearrangements of the normal correlations among signals from the otolith organs, semicircular canals, and vision. They note that the gravitoinertial force that affects the inner ear also determines the effective weight on the head and altered head weight rearranges the relationships among motor commands and the cervical joint, tendon, and muscle afferent feedback contingent on head movements. DiZio and Lackner (1992) conclude that HMDs weighing over 2.5 pounds increase the effective weight of the head by 20% or more and, due to uneven mass distribution of the display system, HMDs can shift the effective center of mass of the head which alters static and dynamic balance. This affects perceived movement of the wearer and elicited symptoms of motion sickness.

HMD stressors on the visual system were discussed earlier. Several studies have suggested that ocular stress resulting from HMDs, such as discrepancies between accommodation and convergence, can lead to VRISE (Ehrlich, 2000; Kawara, Ohmi, & Yoshizawa, 1996; Mon-Williams et al., 1993; Rushton & Riddell, 1999) even though Stanney et al. (1997) found oculomotor symptoms to be the least prevalent symptom cluster. Although later-generation HMDs can reduce visual system compatibility problems (Ehrlich, 2000; Rushton et al., 1994), other HMD effects, particularly latency, vection, and stereopsis, remain contributing factors to VRISE. In sum, “current VEs generate nauseogenic sensorimotor rearrangements because of hardware and software limitations; future systems may be nauseogenic because improved technology will allow the accurate presentation of environments containing sensorimotor rearrangements as an integral part” (DiZio & Lackner, 1992, p. 322).

Simulator-Induced Sickness / VRISE Equipment and Task Factors

Kennedy and Fowlkes (1992) describe simulator-induced sickness as “polygenic” as no single factor has been identified as *the* cause; i.e. the cause of sickness in one simulator cannot be assumed to be the cause of sickness in another. They outline three "drivers" of simulator-induced sickness that include equipment factors, simulator usage, and pilot variables. Expanding on their work, Kolasinski (1995) grouped 40 factors that have been linked with simulator-induced sickness in virtual environments into three possible categories: factors associated with the individual, the simulator, and the task. Individual factors include, among others: age, experience with real-world task, gender, mental rotation ability, perceptual style, and postural stability. Simulator factors include calibration, field-of-view, binocular viewing, lag, position-tracking error, refresh rate, etc. Task factors include degree of control, duration, head movements, luminance level, unusual movements, vection, etc. An individual's susceptibility to simulator-induced sickness can be due to one of these factors or several factors in combination making the syndrome very difficult to predict. However, symptoms in virtual environments have been known to occur most frequently in two general situations: 1) when perceivable and excessive lags are present between head movements and updating of the HMD, and 2) when participants view compelling visual representations of self-motion in the absence of physical cues (vection) (Hettinger & Riccio, 1992). Factors relevant to the current study will be discussed in the following sections.

Lags and latency. To date, an important goal in the creation and presentation of virtual environments is pictorial realism. However, realism is only a necessary feature of a VE when it has a positive effect on the behavior goals of the device (Hettinger & Riccio, 1992). As discussed earlier, striving for a photorealistic pictorial scene increases the demand on the computational system as the subsequent frames of the moving scene are generated and rendered into the frame buffer for display (Pausch et al., 1992). Increased scene complexity increases the update rate latency of the system and can lead to visual lag.

The degree to which lags in the visual system of a simulator affect simulator-induced sickness is unclear. Casali and Wierwille (1980) and Frank et al. (1988) found that visual lags of 170, 300, and 340 ms produced mild discomforting effects in participants in motion-base simulators and were more disruptive to a user's performance and comfort than motion lag. However, in fixed-base simulators, Uliano, Kennedy, and Lambert (1986) found that despite causing disruptions in performance, lags of 108 – 285 ms had no effect on simulator-induced sickness. Although participants in this study performed only two tasks, Uliano et al. (1986) concluded that visual lag is not likely to be a significant factor in simulator-induced sickness.

Other studies, however, have reported that transport delay is a potential factor in simulator-induced sickness (Piantanida et al., 1993 as cited in Nichols, 1999). Transport delay can present problems for operators predicting how their actions affect the system. As described in Pausch et al. (1992), operators could not predict the length of larger transport delays in one flight simulator study of transport delay and attempts to "guess and lead" the system resulted in overcompensation. Overcompensation can create abnormal accelerations, which can contribute to sickness, although the authors are not clear if the accelerations were perceived only in the visual scene or whether the simulator was on a motion base that moved erratically. In either case, however, it is important to minimize transport delay in a driving simulator intended for general public use to reduce at least one potential contributing factor of simulator-induced sickness.

Vection. As discussed earlier, a major goal of many types of virtual environments is to create a compelling sense of presence in the user. A significant component of presence, particularly in vehicular simulators, is the illusion of self-motion or vection. The provocativeness of the simulator's realism or felt presence depend upon the observer's impressions of self-motion in that environment (Kennedy, Berbaum, & Smith, 1993). Vection, however, has been suggested to be a major contributor to simulator-induced sickness and VRISE.

Hettinger and Riccio (1992) state that it is reasonable to hypothesize that any visual display factors that enhance the experience of self-motion in a stationary observer will increase the probability of sickness in VEs. Additionally, displays which produce strong vestibular effects are most likely to elicit simulator-induced sickness symptoms (Kennedy, Hettinger, & Lilienthal, 1990). Based on data obtained in flight simulation studies, Hettinger, Berbaum, Kennedy, Dunlap, and Nolan (1990) conclude that visual displays that produce illusory self-motion are more likely to produce symptoms of simulator-induced sickness. Additionally, individuals who experience the illusion of vection appear to be more at risk for the adverse effects of simulator-induced sickness (Hettinger & Riccio, 1992).

The degree of vection experienced is influenced by the field of view (FOV) of the visual display. Research in traditional driving simulators indicates that greater vection due to a wider FOV may lead to a greater likelihood of conflict with attenuated (or absent) vestibular cues in the simulator (Biocca, 1992; Casali & Wierwille, 1986; Kennedy et al., 1990; McCauley & Sharkey, 1992; Pausch et al., 1992). Anderson and Braunstein (1985), however, induced vection using only a small portion of the central visual field (7.5º) with stimuli which appeared to have depth. This led them to conclude that representation of motion and texture cues in the display may be more critical than the display field of view.

Determination of FOV in HMDs involves tradeoffs beyond that of perceived resolution. A smaller FOV in an HMD requires more head movement (Nichols, 1999) which can increase the likelihood of VRISE (DiZio & Lackner, 1992). Narrower FOVs may also hinder task performances such as maneuvering, grasping objects, and locating moving targets (Witmer et al., 1996). Wider FOVs may improve performance and also feelings of involvement and presence, but this comes at the expense of weight, size, and resolution (Wilson, 1997). Additionally, wider

FOVs exacerbate postural instability in HMD-based, VR systems than other types of simulators (Kennedy & Stanney, 1996).

A large GFOV coupled with slower refresh rates contributes to the likelihood that an operator will perceive display *flicker* – a factor which has been associated with simulator-induced sickness (Pausch et al., 1992). Flicker is defined as pulsed changes in luminance. A flicker rate just high enough to make a target appear steady (*fusion*) is known as the flicker fusion frequency or critical flicker frequency (Stuart, 1996). Stuart (1996) describes factors that influence sensitivity to flicker including location in the visual field (greatest sensitivity in the periphery), the observer's state of adaptation (greatest sensitivity when light-adapted), and the luminance of the target and background (sensitivity increases with contrast). In average room light with a typical CGI luminance level of approximately 20 cd/m^2, a refresh rate of 50 – 60 Hz will appear to be free of flicker. Larger and brighter display systems, however, may require refresh rates as high as 80 – 90 Hz depending on the rate of phosphor decay in the display (Padmos & Milders, 1992). As cited in Stuart (1996), Piantanida (1993) notes that the peripheral visual system is more sensitive to flicker than the foveal visual system which might require faster refresh rates to produce temporal fusion for wider FOV visual displays. The perception of flicker contributes to stress on the visual system which has been shown to be a contributing factor to simulator-induced sickness (Pausch et al., 1992).

Stereopsis. As discussed earlier, there are conflicting reports that the presence or absence of stereopsis can affect the degree of asthenopia an individual may experience. It has long been thought that a binocular viewing condition is more stressful to the visual system than a biocular condition and that greater VRISE can occur. Rushton et al. (1994) found an advantage for biocular over stereoscopic (binocular) viewing conditions in two different generations of HMDs; however, different viewing conditions using the same HMD, software, and stimuli in a within-subjects experimental design were not directly compared (Ehrlich, 2000). In two studies that did use the same controls when testing biocular vs. binocular viewing conditions (Ehrlich, 2000; Peli, 1998), no significant difference was found between the two conditions and Ehrlich (2000) suggested that there may even be a disadvantage for biocular compared to stereoscopic viewing in terms of greater nausea, oculomotor discomfort, disorientation, and severity of discomfort, although results were not conclusive. She did find, however, that a *vergence* viewing condition did appear to allow the oculomotor system to recover more quickly which may reduce the likelihood and severity of post-simulation ataxia.

Simulator-Induced Sickness / VRISE Individual Factors

The literature appears unequivocal in suggesting that susceptibility to simulator-induced sickness differs among individuals; however, the factors that differentiate one individual's susceptibility from another's are not well understood. As noted earlier, Kolasinski (1996) examined individual factors as predictors of susceptibility to VRISE that included, among others: age, concentration level, ethnicity, experience with real-world task, level of adaptation, flicker fusion frequency threshold, gender, illness, mental rotation ability, perceptual style, postural stability, and susceptibility to vection. It is important to understand an individual's susceptibility to simulator-induced sickness and VRISE so that all necessary design and use precautions be taken in a simulator to minimize the potential for simulator-induced sickness.

Age. As driver license testing is performed across a large population range of ages (16 and above), it is important to ensure that a major segment of the population is not excluded in a simulator designed for this testing. Up to this point, the majority of simulator-induced sickness studies have focused on military pilots in flight simulators. McCauley and Sharkey (1992) note that pilots are self-selected based on their resistance to motion sickness and are thus, as a group, less susceptible than the general population. If a driving simulator is to be placed in DMV offices for use by the general public, it is likely that simulator-induced sickness will be a more common occurrence than the literature would suggest. Additionally, unlike military pilots, commercial users may be under the influence of medications or alcohol and it is possible that such substances may increase susceptibility to sickness (Kolasinski, 1995; McCauley & Sharkey, 1992).

Reason and Brand (1975) reported that people of two to 12 years of age have the greatest susceptibility to motion sickness. Susceptibility decreases quickly from ages of 12 to 21 years and then slowly thereafter. They note that sickness is almost non-existent after age 50. The majority of simulator-induced sickness studies have focused on pilots in flight simulators or have used a participant pool of students in a research laboratory – all relatively young in age. Kolasinski (1996) developed a regression model that portrayed a relationship between sickness and age, gender, mental rotation ability, and pre-exposure postural stability. The oldest person included in her model, however, was 32 years old. Very few studies have reported on simulator-induced sickness susceptibility in older persons, perhaps due to the assertion of limited motion sickness susceptibility after 50 (Reason & Brand, 1975), and also the fact that most simulators are not designed for use by the general population.

There is evidence to suggest that the sensitivity of the vestibular system declines with age. A loss in the number of vestibular nerve cells has been found to begin at age 55 and although it increases in severity as a person ages, the brain can compensate for vestibular problems and functionality remains mostly intact in people up to 70 or 80 years of age (Shupert, 1992). This gradual loss in function may explain the limited findings of susceptibility to motion sickness in older persons.

Motion sickness and simulator-induced sickness, however, are not necessarily the same phenomenon. Liu et al. (1999) found that in contradiction with the literature, simulator-induced sickness actually increased with age in an HMD-based, medium-fidelity driving simulator. This is may be due to the relationship between age and real-world experience. Kolasinski (1995) reiterates that conflicts are thought to occur between an actual pattern of stimuli and an expected pattern of stimuli stemming from repeated, real-world experiences. In the case of flight simulators, Kennedy et al. (1990) suggested that the pilot's experience with the sensory aspects of actual flight might lead to a greater sensitivity to what he is experiencing in the simulator versus what he would expect from the real world task. Experienced pilots may be more inclined to expect kinesthetic and vestibular cues that may be missing in the simulator and may be more sensitive to display distortions such as blurring or inappropriate collimation that cause the visual scene to greatly diverge from the dynamic real scene (Casali, 1986). It is possible that users of a driving simulator might react the same way; that is, older persons with greater driving experience may be more sensitive to sensory conflicts and may thus show a greater incidence of sickness than those with less.

As there is scant and conflicting evidence to suggest that age may be a significant factor in simulator-induced sickness susceptibility in virtual environments, and as few studies have looked at this factor directly, it will be used as an independent variable in this research.

Gender. Reason and Brand (1975) state that women are more susceptible to all forms of motion sickness than men. However, as women do not differ from men in their sensory response to motion stimuli (Reason & Brand, 1975), and as this conclusion is based partially on self-reports of motion sickness, it is possible that males may have underreported their susceptibility (Biocca, 1992). In her study of individual factors as a means of predicting susceptibility to simulator-induced sickness, Kolasinski (1996) found that females did not significantly differ from males on any of the four sickness measures used (total severity, nausea, oculomotor discomfort, and disorientation), although gender was found to interact with mental rotation ability. As mental rotation ability improves, predicted sickness increases for females but decreases for males; however, only the *trends* were predicted to differ – not the actual sickness scores between males and females (Kolasinski, 1996). S. A. Jones (1998) found that although unrestrained females experienced greater vection sensations in a driving simulator experiment than unrestrained males, there was no significant difference in susceptibility to simulator-induced sickness based on gender. Another study found that motion sickness symptoms were actually greater in males than in females after participants were subjected to a motion drum (Kennedy, Hettinger, Harm, Ordy, & Dunlap, 1996).

Conversely, the majority of studies have reported that overall, females do report greater symptoms of simulator-induced sickness than males (Kennedy, Lanham, Massey, Drexler, & Lilienthal, 1994). Kennedy et al. (1994) specifically investigated gender differences between male and female Navy helicopter pilots and found that females scored significantly higher on all measures of sickness of the SSQ. In their study of sickness symptoms following HMD-based virtual environment exposure, Stanney, Kennedy, Drexler, and Harm (1999) found that female participants had significantly higher mean SSQ scores than males.

Yoo (1999) suggests three possible factors that may increase female susceptibility to simulator-induced sickness: hormonal influence, peripheral field of view, and vection latency. Anecdotal evidence does suggest that there may be a relationship between hormonal levels and susceptibility to motion sickness as women appear to be most susceptible during the menstrual cycle and pregnancy (Biocca, 1992; Kennedy & Frank, 1985; Reason & Brand, 1975). Additionally, Kennedy and Frank (1985) noted that from the standpoint of functional peripheral fields, women exhibit larger fields of view than men. As simulator-induced sickness appears to be more prevalent in simulators with wide FOV visual systems, women may be more sensitive to vection latency, which is implicated as a factor in simulator-induced sickness. Vection latency is a measure of the potency of a visual stimulus in inducing circular vection and depends on stimulus characteristics such as stimulus texture, area, and *location within the visual field* (Yoo, 1999). As the peripheral retina dominates visually induced vection (Dichgans & Brandt, 1978), a larger display increases the motion detection capability of the larger peripheral retina in females which leads to a difference in susceptibility to simulator-induced sickness between males and females (Yoo, 1999). The conflicting evidence of gender as a predictor of simulator-induced sickness leads this thesis to examine it as a possible factor of an individual's susceptibility.

Conclusion

The sense of presence an individual may experience in a virtual environment is dependent on many factors. Providing an immersive display system can significantly increase one's sense of presence; however, it is often at the expense of eliciting symptoms of discomfort. Simulator-induced sickness has been shown to be related to certain display characteristics and an absence of an expected, corresponding vestibular sensation; however, the exact causes are still unclear. Also unclear is the degree to which individual characteristics affect one's susceptibility to feelings of discomfort. Display type, age, and gender have all been shown to play roles in the potential for individual discomfort; however, there is little in the current research that establishes a link among them.

RESEARCH OBJECTIVES

As there is some Virginia Department of Motor Vehicle interest in placing simulators in regional offices to potentially supplement driver license testing as well as other potential licensing and training applications, it is important that potential side effects such as simulator-induced sickness in the entire range of potential users is well understood. This range of potential users includes: research subjects, driver licensees, and trainees; however, age and gender are factors in motion sickness that show inconsistent results in the literature. There is very little research overall on the effects of gender and age to susceptibility of *simulator-induced* sickness in a traditional simulator or VRISE in a virtual reality environment. As senior citizens would constitute a large pool of potential users of the DMV driving simulator, it is important to understand not only how their driving performance compares against a younger population, but also if they are more susceptible to simulator-induced sickness and VRISE. Although several studies have found that young persons and females are more susceptible to motion sickness, others have called these conclusions into question. However, none have apparently investigated age and gender susceptibility to simulator-induced sickness in conjunction with display type.

Numerous studies have examined factors relevant to the prediction of simulator-induced sickness. Most have found that, although a particular factor can be shown to be significant in the onset of simulator-induced sickness, it is extremely difficult to recommend the levels to which a factor should be set to predict greater susceptibility to simulator-induced sickness. The primary goal of this research, therefore, is determining how one display type affects the population expected to use the simulator in terms of driving performance and simulator-induced sickness susceptibility as compared to the other display type. Each display type, HMD or direct-view, may be usable for simulation applications in and of itself, but one may be more appropriate for certain population types than the other. As a general-purpose driving simulator would be intended for use by an entire population of potential users, the goal of this research will be to provide a general recommendation for the use of a particular display type, with certain parameters, for use with specific populations.

As driving simulator designers are becoming more and more interested in using head-mounted displays as a replacement for traditional direct-view displays, it is important to understand if susceptibility to simulator-induced sickness is decreased and driving performance is increased with the use of one or the other. This research will directly compare the use of a head-mounted display (HMD) with that of a direct-view display in a medium-fidelity driving simulator. This study will investigate the effects of a VR-based driving simulator with that of a direct-view, display-based simulator to see if there is a difference in driving performance and simulator-induced sickness. Additionally, this study will examine ataxic effects of HMD use in a driving simulator to determine if they are any more pronounced than those experienced with the use of a direct-view display.

Specific concerns of this research include:

1. What is the effect of display type on driving performance and susceptibility to simulator-induced in a fixed-base driving simulator in older and younger males and females?
2. What combinations of age and gender show a greater susceptibility to simulator-induced sickness in a fixed-base driving simulator?
3. What is the effect of display type on postural instability (ataxia) following the simulation?
4. What is the relationship between more severe simulator-induced sickness and task performance degradation?

EXPERIMENTAL METHODOLOGY

Experimental Design

The experiment consisted of a mixed-factor design directed at evaluating the effect of <u>*Display Type*</u> (2 levels), <u>*Age*</u> (2 levels), and <u>*Gender*</u> (2 levels) on driving performance and susceptibility to simulator-induced sickness. A complete factorial design (8 treatment combinations) resolved all main effects and interactions in an analysis of variance (ANOVA). Figure 4 portrays the experimental design graphically.

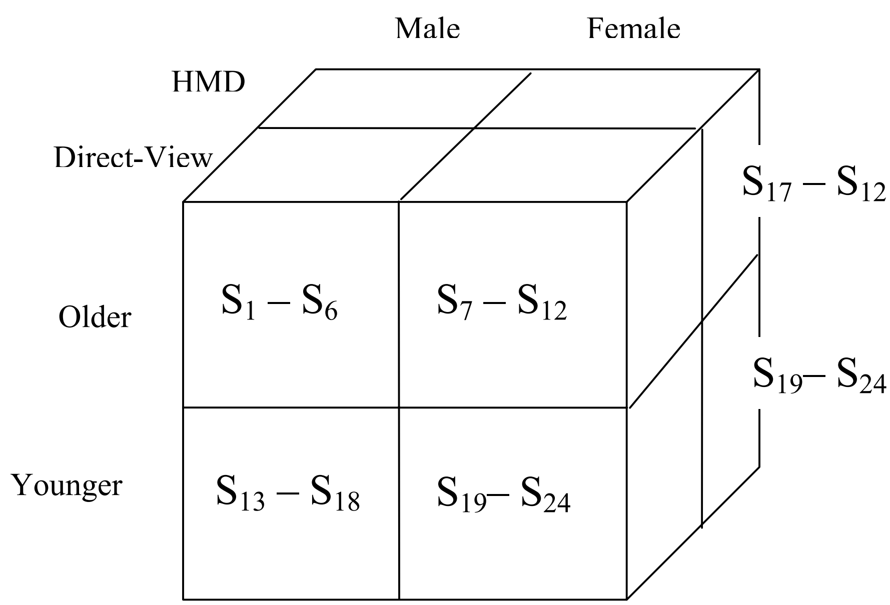

Figure 4. Graphical representation of experimental design

Within-subject designs are founded on the assumption that practice effects eventually reach an asymptote such that they can be disregarded (Keppel, 1991). Based on research that has shown adaptation to be a robust effect in motion and simulator-induced sickness (Gower, Lilienthal, Kennedy, & Fowlkes, 1988; Reason & Brand, 1975), Kennedy and Fowlkes (1992) strongly caution against the use of a within-subjects experimental methodology for evaluation of equipment features in simulator-induced sickness studies. They argue that repeated exposure to a simulator leads to participants experiencing reduced simulator-induced sickness effects. In other words, practice effects will eventually cause participants to adapt to the simulator environment, which increases the error in the experiment.

Adaptation to an environment that can induce simulator or motion sickness has been demonstrated in the literature. However, these studies show adaptation in a simulation or environment over a limited time period and with multiple simulator trials that occur with little time between them. Reason and Brand (1975) discussed adaptation to motion sickness over long exposures in relatively short time frames, and Gower et al. (1988) described adaptation based on nine to ten exposures over a two-week period.

This study assumed that adaptation would not significantly affect the results because: 1) the visual display system would be different for each session, effectively creating a new environment, and 2) participants would only be exposed to the simulation for 15 – 20 minutes with a week hiatus between sessions. This exposure duration would be long enough to induce simulator-induced sickness symptoms, but short enough so that adaptation to the simulator is minimal. The week period between sessions would ensure complete readaptation to the "real world" (Kennedy & Stanney, 1996) such that participants arrive with minimal contamination from the first session. Additionally, treatment order would be counterbalanced to minimize any contamination due to practice effects. Studies such as Howarth and Costello (1997) and Uliano et al. (1986) have successfully drawn conclusions from repeated-measures designs to examine simulator-induced sickness with short exposure times.

Independent Variables

Per Figure 4, three independent variables of two levels each were studied using the VT-UVA-Carilion driving simulator. The within-subject variable, *Display Type*, was either *HMD* or *Direct-View*. The literature is divided on whether the biocular or stereo (binocular) viewing condition places more stress on the visual system. However, for the binocular condition, adjustments for interpupillary distance (IPD) are critical and errors can cause problems during post-exposure readaptation to normative conditions (Stanney et al., 1998). Therefore, the HMD was used in biocular mode, which reduced the complexity of fitting and calibration and thus, between-subject variability.

Age and *Gender* were the between-subject variables. The levels of the *Age* variable were named *Younger* and *Older* and ages were obtained from the participant information forms (Appendix B). Younger participants were defined as those between 18 and 30 years of age and older participants were persons over 55 years of age. *Gender* was also two levels, *Male* and *Female.* As this study was only concerned with the effect of display type and was not attempting to make predictions of simulator-induced sickness based on individual differences, other factors such as perceptual style, mental rotation ability, etc. were not collected.

Dependent Measures

As previously discussed, this research was concerned with the effect of display type, age, and gender on driving performance and susceptibility to simulator-induced sickness or VRISE. Dependent measures were divided into three variable classes believed to be sensitive indexes of driving performance, postural instability, and simulator-induced sickness. DiZio and Lackner (1992) cite several studies that indicate that the monitoring of physiological correlates of motion sickness such as electrogastrograms, cardiopulmonary function, phasic skin conductance, and peripheral blood flow, although helpful for understand the physiology of motion sickness under laboratory conditions, is insensitive, nonspecific, unreliable, and/or restrictive under operational conditions. As this study was more concerned with the effects of simulator-induced sickness rather than its underlying causes, the SSQ was sensitive enough to detect simulator-induced sickness. As such, physiological measures were not taken.

1. <u>Simulator Sickness Questionnaire (SSQ)</u>

   In accordance with related research (S. A. Jones, 1998; Yoo, 1999), Total Severity Score (TSS) was used to assess pre- and post-simulator exposure symptomology by using the Symptom Checklist portion of the Simulator Sickness Questionnaire (SSQ; see Appendix C). The SSQ was originally developed by Essex Corporation (Kennedy, Lane et al., 1993) for the purpose of measuring side effects following military flight simulator training sessions. As much of the SSQ is specific to aviators, only the 16-symptom checklist portion was used.

   The self-report 16-symptom checklist of the SSQ is the most common means of measuring simulator-induced sickness (Yoo, 1999). Due to the polysymptomatic nature of simulator-induced sickness, it is necessary to assess several signs and symptoms as only one would not be sensitive enough (Kennedy & Fowlkes, 1992). The checklist consists of 16 items that participants either indicated as absent by marking "none," or present by indicating the degree of severity (slight, moderate, or severe). The SSQ measures consist of three symptom subscales: 1) Oculomotor (eyestrain, difficulty focusing, blurred vision, headache); 2) Disorientation (dizziness, vertigo); and 3) Nausea (nausea, stomach awareness, increased salivation, burping). Participants were also provided additional space to indicate the presence of any symptoms not specifically included on the checklist.

   Scoring was in accordance with the method developed by Kennedy, Lane et al. (1993). Symptoms indicated as absent were given a value of 0 and slight, moderate, and severe symptoms were given values of 1, 2, and 3, respectively. The overall index of sickness (TSS) score was derived by summing the individual symptom scores and then multiplying the total by a weighting factor. The total severity score can range from a low of 0 to the maximum possible value of 235.62 if all symptoms are marked "severe." Individual values have no particular interpretive meaning; their function is to produce scales with similar variabilities on which they can be compared (Kennedy, Lane et al., 1993). Despite this, a score of 15 or higher probably indicated that individuals experienced enough discomfort such that they would not voluntarily seek additional exposure (Stanney et al., 1997).

The SSQ checklist was administered before and after the driving task. Responses to the pre-exposure checklist were used as a screening tool to determine the fitness for participation in the study since pre-exposure symptoms can predispose susceptibility to motion sickness (Kennedy et al., 1990; Kennedy et al., 1992). In accordance with S. A. Jones (1998) and Yoo (1999), a potential participant scoring a TSS of 7.48 or higher on the pre-exposure checklist was not invited to take part in the experiment at that time. This allowed post-exposure differences to be directly attributable to simulator exposure and conclusions could then be drawn about post-exposure symptomology without interference from any predisposing factors.

2. <u>Postural disequilibrium tests</u>

Postural stability before and after exposure is a frequently used measure of vestibular disruption following exposure to altered sensory environments. Thomley et al. (1986) recommends the Stand-On-Nonpreferred Leg (SONPL) test as the method of first choice to measure highly transitory ataxic effects; however, the Two-Leg, Heel-To-Toe test (TLHTT) was employed here because it is just as reliable, nearly as sensitive, and safer for older participants (Kennedy & Stanney, 1996).

Participants were instructed to stand erect with their arms crossed and their eyes closed while on a hard, smooth floor. Leg preference was determined by asking participants to stand on one leg with no further instructions. The participants' choice became their "preferred" leg. Participants were then asked to stand with feet heel-to-toe with their preferred leg in the back as an anchor. Participants were asked to hold the position for a maximum of 30 seconds or until they moved a foot, uncrossed their arms, or opened their eyes. Each test consisted of three trials with a one-minute rest between each trial. With the older age group, the experimenter stood next to the participants in order to catch them if they come close to falling.

Scoring consisted of the total amount of time participants maintained their balance without moving a foot, opening their eyes, or unfolding their arms. To avoid ceiling effects, the mean of the trials was used (Thomley et al., 1986). The test was administered three times for each participant: before the simulator exposure, directly after simulator exposure, and 30 minutes following the simulator exposure.

3. Driving performance measures

Time to complete course (TCC). The total time required by the operator to complete the simulated course is an indication of the comfort and skill level of the participant. All participants were instructed to maintain the posted speed limits as closely as possible. Participants who took longer to complete the course were not as comfortable driving at the posted speed limits, and thus may have had greater difficulty operating the simulator.

Steering angle variance (SAV). Although steering wheel reversals have been shown to be a reliable measure of driver performance and vehicle controllability (McLane & Wierwille, 1975; McLean & Hoffman, 1975), the simulator's computer output datafile did not allow for extraction of this particular measure. Instead, the computer sampled steering wheel angle position (in radians) relative to a vertical centerline every 125 ms. A higher total steering wheel angle variance indicates that the operator either manipulated the steering wheel more often (frequency) or to a greater degree with each rotation (magnitude) than a driver with a lower total steering wheel angle variance. This implies that an operator with a higher variance made more steering corrections and did not steer the vehicle as "smoothly" as an operator with a lower variance. Although not identical to a discrete steering wheel reversal measure, it suggests a similar effect.

Lane deviation (LD). The center of the right lane and left lanes were set at zero with negatively increasing deviations from center to left and positively increasing values for deviation from center to right. The amount of lane deviation, in meters, was collected by the computer every 125 ms and written to a datafile. Flags in the datafile indicated when the vehicle was out of or changed lanes.

Participants and Treatment Order

This study recruited six people for each experimental condition for a combined total of 24 participants. The average age of the younger male participants was 26 (range: 22 to 29), and the average age of the younger female participants was 24 (range: 19 to 27). The average age of the older male participants was 71 (range: 58 to 79), and the average age of the older female participants was 65 (range: 57 to 77). Screening criteria and procedures are discussed in the experimental procedures section. All participants were from the local Blacksburg community.

Counterbalancing was used to reduce practice effects as it allowed the ordering of treatment effects such that each treatment was administered an equal number of times in the same order (Keppel, 1991). In the current experiment, counterbalancing was constructed so that three older females received the HMD session first and three received the DV display first. Three older males received the HMD session first and three received the DV display first and so on until all participants receive both treatment conditions. When possible, the first three participants of each group that arrived to Session 1 were randomized as to which display type they received first. Table 1 summarizes participant information and treatment order.

TABLE 1. Participant Information and Counterbalanced Order of Treatments

Participant Number	Age	Gender	Session 1 Display Type	Session 2 Display Type
28	22	M	HMD	DV
3	23	M	HMD	DV
13	26	M	HMD	DV
20	26	M	DV	HMD
6	28	M	DV	HMD
10	29	M	DV	HMD
8	19	F	DV	HMD
5	21	F	HMD	DV
17	22	F	HMD	DV
21	25	F	DV	HMD
2	27	F	DV	HMD
4	27	F	HMD	DV
32	58	M	DV	HMD
26	66	M	DV	HMD
23	69	M	HMD	DV
30	75	M	HMD	DV
22	76	M	DV	HMD
25	79	M	HMD	DV
9	57	F	DV	HMD
29	60	F	HMD	DV
11	63	F	HMD	DV
19	64	F	DV	HMD
31	68	F	DV	HMD
18	77	F	HMD	DV

Apparatus

Hardware. The simulator used in this study was the fixed-base driving simulator platform described in Penhallegon, Perala, Robinson, and Casali (2002), which can be found in Appendix A. The base consisted of a modified Advanced Therapy Products, Inc. (ATP) Interactive Driving Simulator Console with force-feedback steering. To simulate the functionality of a real car, the console was equipped with gas and brake pedals for foot actuated manual control and were capable of simulating the real force and feel of car pedals.

Two different displays were used with the simulator during different trials. The first was an NEC PlasmaSync 50MP1 50 in (diagonal) flat panel monitor. The display can support a multitude of resolution levels on different computer platforms. The second was the IO Systems i-glasses SVGA 3-D Head Mounted Display with the Intersense Intertrax2 headtracking system.

Software. The simulator was controlled by an IBM-compatible personal computer running scene-generating simulation software developed by Dr. Ronald Mourant at Northeastern University in Boston, Massachusetts. The simulated driving environment, or driving course, was a representation of a section of road in Charlottesville, Virginia and began in a virtual Charlottesville DMV building parking lot. Drivers would begin operation of the vehicle in the parking lot, then make a hard left turn to enter the course.

The modeled course consisted of four segments, which included: a 1.2 mile, two-lane urban environment (35 mph speed limit); a 0.6 mile, four-lane divided highway (45 mph speed limit); 1.8 mile, four-lane interstate roadway (65 mph speed limit); and a 2.9 mile, four-lane loop-back road section (65 mph speed limit) that returned the driver to the DMV office at the beginning of the course. Speed limit and other appropriate signs were modeled in the simulation. The software permitted the driver to drive an unlimited number of laps and collected data until the simulation was halted. Data were captured throughout the simulation run, sampled approximately every 125 ms, and saved to a space-delimited text file. Participants averaged approximately 10 minutes to complete one lap around the course with either display type. A more detailed description and technical specifications for the simulator hardware and software, including driving course screenshots, can be found in Penhallegon et al. (2002) (Appendix A).

Pilot Tests

A pilot-test was conducted using a volunteer from the graduate student community at Virginia Tech. The pilot-test involved setting up the hardware and software, running the driving scenario with the various conditions applied, and timing the length of time required for the experimental session. Pretesting ensured that a) data were being gathered and stored properly, and b) simulator-induced sickness symptoms, if elicited, were measurable by the apparatus and were not too severe. As with the experimental participants, the volunteer for the pretest was required to read and sign the Informed Consent Form as shown in Appendix D.

Experimental Procedure

Participants were randomly assigned to begin with one of the two treatment conditions. The general procedure consisted of two sessions, each dedicated to the use of one of the two display types. The two experimental sessions were separated by 11 days, on average.

Participant recruitment and pre-screening. Participants were recruited primarily through e-mail. People were asked to volunteer only if they a) had a valid US driver's license, b) were either between the ages of 18 – 30 or 55 and older, c) had normal vision (which was later verified through a screening test), and d) had never worn a head-mounted display. As volunteers contacted the experimenter and indicated that they met these initial criteria, they were brought into the laboratory for a continuation of the screening process.

Pre-experimental procedures. Before the laboratory screening process began, volunteers were asked to read the informed consent form, which included a general description of the experiment (Appendix D). After the experimenter addressed any questions the participant may have had, the participants were asked to sign the informed consent form. During the informed consent process, participants were told to notify the experimenter if they wished to stop the simulation at any time for any reason.

Participants were first screened by completing the pre-exposure SSQ. Any participant scoring a 7.48 or higher was invited to continue the experiment at a later date. One younger male was rescheduled in this fashion. One older female scored higher than 7.48 but decided to end her participation rather than rescheduling.

Participants were then asked to fill out a participant information form (Appendix B), based on the one used by Yoo (1999). It inquired if, among other things, participants were in a normal state of health, had gotten an adequate amount of sleep the previous night, had not taken any perception-altering drugs in the 24 hours preceding the experiment, and had never been exposed to an HMD-based VE. Only one individual with prior HMD experience was used in the study. This participant had worn an HMD for approximately five minutes at a science museum four years prior, and it was decided that this extremely limited experience was unlikely to affect the results of the experiment.

Participants were then given a near/far visual acuity test and color discrimination test using the Optec 2000 vision tester developed by Stereo Optical Co. Participants were disqualified if they had difficulty with color discrimination and a near/far visual acuity greater than 20/30 (corrected or uncorrected). Difficulty with color discrimination disqualified one older male and one younger male. Limited visual acuity led to the disqualification of two older females. Participants passing the screening process began the Session 1 procedures immediately.

Session 1. The procedures for each session were loosely based on those of Frank (1985). Following the screening process and prior to entering the simulator, participants were administered the TLHTT postural disequilibrium test in accordance with the procedures described earlier. Participants were then assisted into the simulator and adjusted the seat position and backrest tilt until they were comfortable. If the participant was using the Plasma display, the console was moved forward or backward until there was a 7.5 ft distance between the bridge of the participant's nose and the center of the display. This maintained a consistent, 27° horizontal geometric FOV between the direct-view and head-mounted displays. If the participant was using the HMD, it was adjusted on their head until it was tight enough such that the participant did not feel as if it would slip, but not so tight as to be painful.

The headtracker drift problem was also explained to participants using the HMD at this time (Penhallegon et al., 2002). In most cases, participants were responsible for recentering the headtracker and were able to do it without problem. In a few cases, however, the experimenter stood behind the participant and recentered the headtracker as necessary. This was necessary with some of the older participants who had great difficulty doing it on their own due to arthritis, problems with finding the correct button, etc. The number of times the headtracker had to be recentered varied from only once or twice to over a dozen.

The experimenter then described the functions of the simulator controls (steering wheel, turn-signal stalk, and gas and brake pedals) and explained that participants would first be driving one-half of the course loop for simulator familiarization purposes. They were told that data would not be collected at this time. During the familiarization lap, participants were instructed to try to maintain the speed limit as close as possible and to pass other vehicles as necessary. The experimenter stopped the familiarization lap just before participants would have entered the interstate portion of the course.

Participants were then told to relax as the experimenter reinitialized the simulation and reduced the room illumination. Participants were given approximately two minutes to rest as the simulation loaded, and were then notified that the trial was about to start. When the simulator was operational, instructions for the driving task were presented as the participant operated the vehicle. Participants were instructed to maintain a centered, right-lane position, to again observe the posted speed limits as closely as possible, and to pass vehicles as necessary on the highway and loopback sections of the course. Participants were notified when they had to make a turn, and were asked to pass two vehicles on the interstate portion of the course. If the HMD was being used, participants may have been reminded to recenter the headtracker as necessary. Participants drove one lap around the course and completed the simulation by turning back into the DMV parking lot.

Upon completion of the driving task, participants were assisted from the simulator. The TLHTT postural disequilibrium test was then immediately administered and then the participant was asked to fill out the post-exposure SSQ. Participants were then asked to respond in writing to several subjective questions on the Participant Information Form (Appendix B) about their experience. This usually led to unstructured discussions with the experimenter about comfort, ergonomics, and usability of the display system employed for this session as well as potential problems with using a simulator in a DMV environment.

Participants were required to stay in the laboratory for 30 minutes following the simulation exposure. The 30-minute period began at the end of the third post-exposure TLHTT trial. During this time, participants were asked to indicate whether they were feeling any simulator-induced sickness symptoms and if so, were they dissipating. The appointment for a second session was also scheduled at this time. While waiting, participants were offered bottled water and provided with reading material. At the end of the 30 minutes, the TLHTT postural disequilibrium was again administered. Participants were then allowed to leave and instructed to contact the experimenter by telephone or e-mail if they experienced a recurrence of symptoms.

The experimenter contacted participants the next day either by telephone or through e-mail. Participants were asked if any other aftereffects were noted that might be attributed to the experiment, and the duration of these effects. Any additional comments/concerns by the participants were also addressed at this time, and they were again reminded of the date and time for the second session.

Session 2. The procedures for Session 2 were nearly identical to those of the first session, minus the vision testing. A reduced participant information form was provided (Appendix B) and participants were asked to indicate the type and duration of any symptoms that lingered following the first session. Following the post-exposure SSQ, their debriefing included eliciting subjective observations of the two display systems on their participant information form. The remainder of the procedures was the same and participants were then paid for both sessions.

A summary of the procedures of each session in chronological list form is provided in Appendix E.

RESULTS

Data Reduction

As noted earlier, three categories of dependent measures were collected: driving performance, simulator sickness questionnaire scores, and postural stability test scores.

Simulator-induced sickness. As indicated previously, each participant completed a simulator sickness questionnaire (SSQ) following simulator exposure. SSQ scoring, as described in the Experimental Methodology section of this paper, was consistent with the method described by Kennedy, Lane et al. (1993). A Microsoft Excel spreadsheet was created to perform the calculations. A second Excel spreadsheet was created that summarized participants' TSS scores for each session in a table. This table was then imported directly into Statsoft Statistica v6.1 for analysis.

Ataxia. The postural stability test data consisted of the number of seconds that participants were able to hold their position before self-correction. Each test consisted of three trials, which were then averaged to determine participants' final scores for each test. An Excel spreadsheet was developed that calculated the averages for each test and summarized them in a master table. The final data table consisted of a Pre/Post score and a Pre/Final score. The Pre/Post score was the signed difference between the post-exposure average and the pre-exposure average. The Pre/Final score was the signed difference between the final average and the pre-exposure average. A negatively signed difference indicated that the post or final averages were less than the pre-exposure average, which suggested that the participant's ability to balance following the experiment was degraded. A positive difference indicated that they were greater, which suggested that the participant's ability to balance following the experiment was not degraded. The summary table was then imported into Statistica for analysis.

Driving performance measures. Measures of driving performance were described in the dependent variables section and were calculated from data elements sampled by the computer every 125 ms. Recorded elements were written to a space and tab-delimited text file, which was imported into Microsoft Excel for organization, and Statistica for analysis. *Time to complete course* (TCC) was calculated by subtracting the computer system time just before vehicle velocity changed from zero from the computer system time at the "end of the lap" data file flag. This value, in milliseconds, was divided by 1000 to return TCC in seconds.

Lane deviation (in meters) and *steering wheel position* (in radians) were directly collected by the computer. In order to prevent purposeful lane deviations from skewing the averages, lane deviation data were filtered so that data points 0.5 s before and 0.5 s after lane changes were removed. Lane deviation data were converted to millimeters, and each participant's final score consisted of the average lane deviation throughout the entire course (minus lane changes). The final steering angle variance datum for each subject was arrived at in Excel by first converting the angle (in radians) to degrees, then calculating the overall variance of steering wheel position over the entire course. Summary tables of each measure for each subject were created in Excel and imported into Statistica.

The number of collisions with autonomous traffic was originally intended to be used in the analysis as a measure of driving performance. During the experimental trials, it became apparent that collision counts would not be a meaningful measure as the vast majority occurred when the driver slowed down to turn onto either the interstate or into the DMV office and was struck from the rear by an autonomous vehicle. As a lack of rear and side-view mirrors precluded the driver from knowing if an autonomous vehicle was following, collisions occurred randomly and were not the result of poor driving performance. Thus, this measure was not utilized.

Unbalanced data. Two older female participants could not complete the driving course in the first session due to severe nausea experienced during the familiarization lap. The experimenter recorded their answers to the SSQ while they recovered; however, neither felt well enough to perform the post-exposure ataxia test. Additionally, the computer failed to record driving performance data for one younger male. A lack of time precluded finding other participants. As a result, the ataxia and driving performance data sets were unbalanced, and Statistica corrected for this by using the Type III sum of squares for the MANOVA and ANOVA analyses. The selection of the Type III sum of squares is discussed in the next section.

Statistical Analysis

General statistical approach. Analysis of variance (ANOVA) was used to determine the significance of main effects and interactions for the SSQ data. Due to the presence of multiple and interrelated dependent measures, the multivariate analysis of variance (MANOVA) procedure using the Wilks *Lambda* test was performed for the analysis of ataxia and driving performance. This was done because only performing individual ANOVAs in the presence of interrelated dependent measures may result in a Type I error, i.e., the rejection of a true hypothesis. Individual ANOVAs were then performed on each of the dependent measures, and only effects found significant in both the MANOVA and ANOVAs were selected for further analysis.

As the design contained a within-subjects component, three-way ANOVA and MANOVA analyses were first performed with *Session* as an independent, within-subjects variable in order to determine if the second session was affected by practice effects from the first. Table 2 is a sample ANOVA summary table that is based on 24 total participants; $n = 6$.

TABLE 2. Sample ANOVA Summary Table

Source	df	SS	MS	*F*
Between				
Age (A)	1	SS_A	MS_A	$MS_A/MS_{S/AG}$
Gender (G)	1	SS_G	MS_G	$MS_G/MS_{S/AG}$
A × G	1	$SS_{A\times G}$	$MS_{A\times G}$	$MS_{A\times G}/MS_{S/AG}$
S/AG	20	$SS_{S/AG}$	$MS_{S/AG}$	
Within				
Display (D)	1	SS_D	MS_D	$MS_D/MS_{D\times S/AG}$
A × D	1	$SS_{A\times D}$	$MS_{A\times D}$	$MS_{A\times D}/MS_{D\times S/AG}$
G × D	1	$SS_{G\times D}$	$MS_{G\times D}$	$MS_{G\times D}/MS_{D\times S/AG}$
A × G × D	1	$SS_{A\times G\times D}$	$MS_{A\times G\times D}$	$MS_{A\times G\times D}/MS_{D\times S/AG}$
D × S/AG	20	$SS_{D\times S/AG}$	$MS_{D\times S/AG}$	
Total	47	SS_{total}		

The ataxia data set contains two missing cases for the older females that were unable to perform the test. In situations where a dataset is unbalanced but contains no missing cells, Statistica recommends using the Type III sum of squares in hypothesis testing to analyze unbalanced designs with no missing cells. The Statistica Electronic Manual notes:

> *Type I and Type II sums of squares usually are not appropriate for testing hypotheses for factorial ANOVA designs with unequal n's. For ANOVA designs with unequal n's, however, Type III sums of squares test the same hypothesis that would be tested if the cell n's were equal, provided that there is at least one observation in every cell. Specifically, in no-missing-cell designs, Type III sums of squares test hypotheses about differences in subpopulation (or marginal) means. When there are no missing cells in the design, these subpopulation means are least squares means, which are the best linear-unbiased estimates of the marginal means for the design.*
>
> (StatSoft Inc., 2002)

As Type III sums of squares are useful for balanced as well as unbalanced ANOVA designs with no missing cells, they were used exclusively throughout the analysis.

As all independent variables contained only two levels, interactions were analyzed with simple-effects *F*-tests when appropriate. The Scheffé Multiple Comparison Procedure was used for all post-hoc tests. It is generally considered among the more conservative tests, making it appropriate for use in designs with unequal sample sizes (J. Jones, 1998). Due to the unbalanced driving performance data set, the conservative Scheffé post-hoc tests were used to analyze the nature of significant effects and interactions when appropriate.

Simulator-induced sickness results. As data collection progressed, the experimenter noticed that participants seemed to be getting less sick during the second session than the first. Although the experimental design attempted to minimize any potential practice effects from the first session to the second, it was decided to analyze the data using *<u>Session</u>* as a within-subjects variable in order to confirm the presence or absence of any order effects.

The simulator sickness dependent measure was quantified in the form of the SSQ total severity score (TSS). A three-way ANOVA was performed using *<u>Age</u>*, *<u>Gender</u>*, and *<u>Session</u>* as the independent variables. Significant ($p \leq 0.05$) main effects of Age (A) {$F(1,20) = 12.68$, $p = 0.0020$} and Session (S) {$F(1,20) = 10.51$, $p = 0.0041$} were observed. These results indicate that the older participants experienced a greater severity of simulator-induced sickness than the younger persons, and all participants experienced a greater severity of simulator-induced sickness in the first Session as opposed to the second. The complete ANOVA summary table for order effects is provided in Table 3. Means and standard deviations for all possible treatment conditions are shown in Table 4. As the *<u>Age</u>* and *<u>Session</u>* independent variables only contained two levels, and no significant interaction effects were present, further analysis was unnecessary. The patterns of significance are shown in Figures 5 and 6.

TABLE 3. Analysis of Variance Summary Table for Session SSQ TSS Scores

Source	dF	SS	MS	*F*	*p*
Between Subjects					
Age (A)	1	5711.60	5711.60	12.68	**0.0020***
Gender (G)	1	728.52	728.52	1.62	0.2181
A × G	1	564.17	564.17	1.25	0.2764
Error	20	9010.35	450.52		
Within Subjects					
Session (S)	1	2685.62	2685.62	10.51	**0.0041***
S × A	1	420.79	420.79	1.65	0.2142
S × G	1	466.25	466.25	1.82	0.1919
S × A × G	1	29.14	29.14	0.11	0.7392
Error	20	5112.47	255.62		
Total	47				

***Statistically significant effect at $p \leq 0.05$.**

TABLE 4. Descriptive Statistics for Session SSQ TSS Scores*

			Session 1	Session 2
Younger				
	Females	Mean	16.21	2.49
		Standard Deviation	9.94	4.53
		Observations	6	6
	Males	Mean	10.60	6.23
		Standard Deviation	9.87	9.94
		Observations	6	6
Older				
	Females	Mean	52.36	23.69
		Standard Deviation	35.72	17.96
		Observations	6	6
	Males	Mean	29.92	16.83
		Standard Deviation	23.05	19.47
		Observations	6	6

*** Higher total sickness scores indicate greater levels of discomfort**

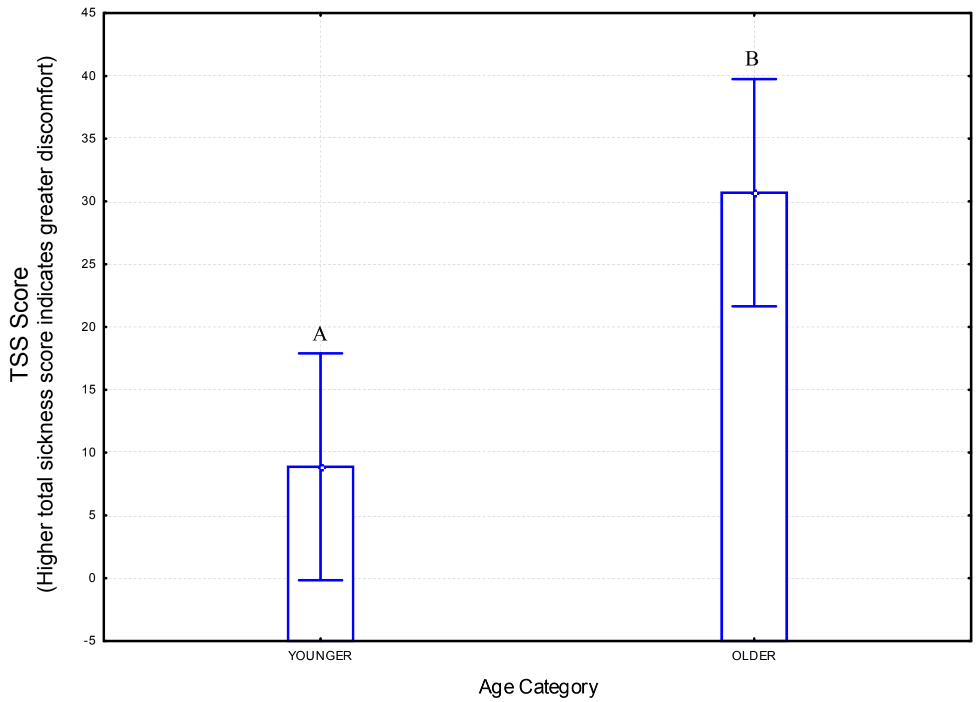

Figure 5. Age Main Effect (p = 0.0020) for SSQ TSS by Session (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

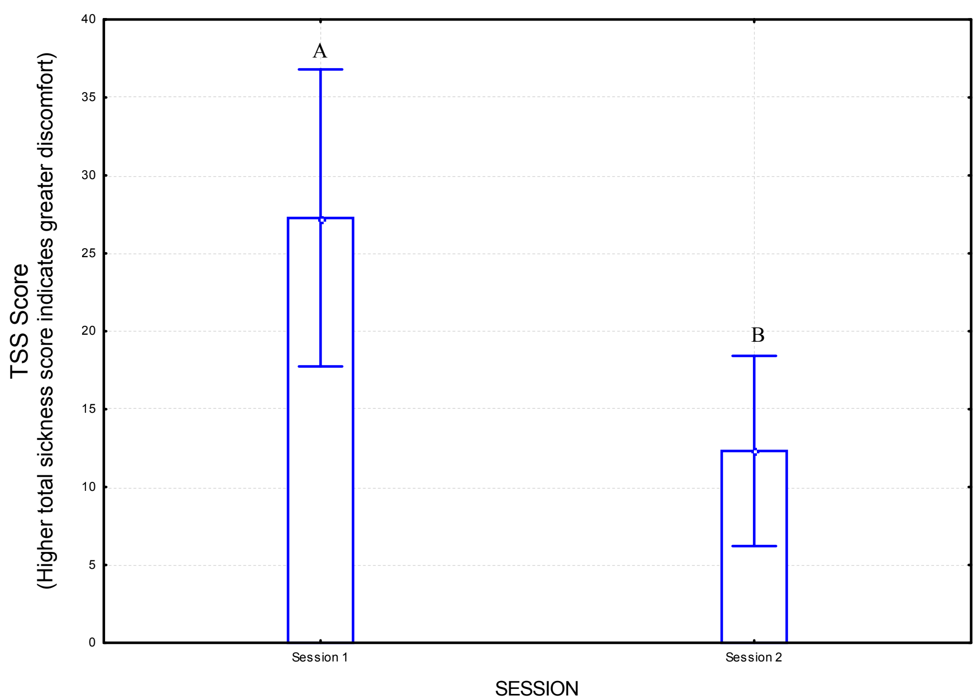

Figure 6. Session Main Effect (p = 0.0041) for SSQ TSS by Session (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

Following the order effects analysis, a three-way ANOVA was performed using Age, Gender, and Display (type) as the independent variables. In this case, the only significant ($p \leq 0.05$) main effect observed was Age (A) {$F(1,20) = 12.68$, $p = 0.0020$} which showed again that the older participants experienced a greater severity of simulator-induced sickness symptoms than the younger participants. The complete ANOVA summary table for order effects is provided in Table 5. A post-hoc analysis was unnecessary as each variable only contained two levels and no significant interaction effects were found.

The within subjects (display type) section of Table 5 shows three F-ratios that are less than one. In general, an F-ratio that is less than one indicates either of two things: a) high error or variability resulting from problems with the measurement tool, or b) there was not enough variability in the results to produce a treatment effect (N. Warner, personal communication, 6/03/03). As the SSQ is a validated measuring tool, it is likely that low F-ratios resulted from the practice effects, which violated the independence assumption of the F-distribution. This caused the minimal variability in the results, which produced the low F-ratios. Means and standard deviations for all possible treatment conditions are shown in Table 6.

TABLE 5. Analysis of Variance Summary Table for Display Type SSQ TSS Scores

Source	dF	SS	MS	*F*	*p*
Between Subjects					
Age (A)	1	5711.60	5711.60	12.68	**0.0020***
Gender (G)	1	728.52	728.52	1.62	0.2181
A × G	1	564.17	564.17	1.25	0.2764
Error	20	9010.35	450.52		
Within Subjects					
Display (D)	1	29.14	29.14	0.08	0.7844
D × A	1	0.00	0.00	0.00	1.0000
D × G	1	728.52	728.52	1.92	0.1808
D × A × G	1	377.67	377.67	0.99	0.3301
Error	20	7578.95	378.95		
Total	47				

***Statistically significant effect at $p \leq 0.05$.**

TABLE 6. Descriptive Statistics for Display Type SSQ TSS Scores*

			Display Type	
Younger			**DV**	**HMD**
	Females	Mean	7.48	11.22
		Standard Deviation	7.10	13.17
		Observations	6	6
	Males	Mean	8.73	8.10
		Standard Deviation	10.75	9.58
		Observations	6	6
Older				
	Females	Mean	30.54	45.50
		Standard Deviation	22.24	38.24
		Observations	6	6
	Males	Mean	29.30	17.45
		Standard Deviation	27.83	12.44
		Observations	6	6

*** Higher total sickness scores indicate greater levels of discomfort**

As shown in Table 3, <u>*Session*</u> was a significant factor, which indicates that practice effects were present in the experiment. Further analysis was performed in order to determine if display type was a significant effect without the contaminating practice effects of Session. For this analysis, all Session 1 data were disregarded, and a completely between-subjects ANOVA was performed using <u>*Age*</u>, <u>*Gender*</u>, and <u>*Display*</u> (type) as independent variables, and Session 2 TSS scores as the dependent variable. Session 1 data were disregarded because participants were still familiarizing themselves with the simulation at that point. When participants returned for Session 2, each had experienced the full course and had the same level of exposure. In this case, again the only significant ($p \leq 0.05$) main effect observed was <u>*Age*</u> (A) $\{F(1,16) = 6.96$, $p = 0.0179\}$; that is, there were no changes in effects from the earlier ANOVA which included data from both sessions. Again, several *F*-ratios were present that were less than one, but was likely a result of an extremely small treatment effect. The ANOVA procedure for display type in Session 2 is provided in table 7.

TABLE 7. Analysis of Variance Summary Table for Session 2 SSQ TSS Scores

Source	dF	SS	MS	F	p
Between Subjects					
Age (A)	1	1515.906	1515.906	6.96	**0.0179***
Gender (G)	1	14.570	14.570	0.07	0.7993
Display (D)	1	168.434	168.434	0.77	0.3924
A × G	1	168.434	168.434	0.77	0.3924
A × D	1	168.434	168.434	0.77	0.3924
G × D	1	210.397	210.397	0.97	0.3405
A × G × D	1	70.521	70.521	0.32	0.5774
Error	16	3487.575	217.973		
Total	23				

***Statistically significant effect at $p \leq 0.05$.**

One additional TSS analysis was performed for the purposes of the ataxia discussion. Session 2 data were analyzed to determine whether older participants still showed significantly higher TSS scores than younger participants when Session 1 data were disregarded. A between-subjects ANOVA analysis was performed using *Age* as the independent variable and Session 2 TSS scores as the dependent measure. In this case, the *Age* (A) main effect was again found to be significant ($p \leq 0.05$) {$F(1,22) = 7.78$, $p = 0.0107$} demonstrating that the older participants experienced significantly greater levels of discomfort than the younger participants. The ANOVA results are shown in Table 8.

TABLE 8. Analysis of Variance Summary Table for Session 2 SSQ TSS Scores by Age

Source	dF	SS	MS	F	p
Between Subjects					
Age (A)	1	1515.906	1515.906	7.78	**0.0107***
Error	22	4288.365	194.926		
Total	23				

***Statistically significant effect at $p \leq 0.05$.**

Ataxia. The two ataxia dependent measures consisted of the signed difference between the average of each participant's trial scores just before and just after operating the simulator (pre-exposure/post-exposure), and just before and 30 minutes after operating the simulator (pre-exposure/final). The participant's postural stability scores were analyzed to determine if display type produced a significant difference between the pre-simulation exposure test and the post-simulation exposure test, and the pre-simulator exposure test and the final test. A negative mean difference indicated that the amount of time a participant could maintain their postural equilibrium was less in the post-exposure or final conditions than in the pre-exposure condition. A positive mean difference indicated that participants were able to maintain their postural equilibrium longer in the post-exposure or final conditions than in the pre-exposure condition. In other words, a positive mean value suggests that the simulation had no negative impact on a participant's ability to maintain postural equilibrium.

Because a significant order effect was discovered during the simulator-induced sickness analysis, a MANOVA analysis of the ataxia data was first performed using *Age* and *Gender* as the between-subjects variables and *Session* as a within-subjects variable. The subsequent MANOVA analysis used *Display* (type) as the within-subjects variable.

The first three-way MANOVA was performed using *Age*, *Gender*, and *Session* as the independent variables. Significant ($p \leq 0.05$) main effects of Age (A) {$F(2,17) = 5.24$, $p = 0.0169$} and Gender (G) {$F(2,17) = 5.55$, $p = 0.0139$} were observed. The complete MANOVA summary table for order effects is provided in Table 9. The results suggest that there were no practice effects present from the first session to the second. Means and standard deviations for all possible treatment conditions are shown in Table 10. The *Age* and *Gender* main effects for each of the dependent measures are the same for both the *Session* and *Display* (type) analyses, and are further examined in the *Display* (type) analysis.

TABLE 9. MANOVA Summary Table for Ataxia by Session

Source	Effect dF	Error dF	Test	Value	*F*	*p*
Between Subjects						
Age (A)	2	17	Wilks Lambda	0.618669	5.24	**0.0169***
Gender (G)	2	17	Wilks Lambda	0.604871	5.55	**0.0139***
A × G	2	17	Wilks Lambda	0.924608	0.69	0.5136
Within Subjects						
Session (S)	2	17	Wilks Lambda	0.814872	1.93	0.1755
S × A	2	17	Wilks Lambda	0.883516	1.12	0.3490
S × G	2	17	Wilks Lambda	0.914056	0.80	0.4659
S × A × G	2	17	Wilks Lambda	0.943649	0.51	0.6108

***Statistically significant effect at $p \leq 0.05$.**

TABLE 10. Descriptive Statistics for Ataxia Scores by Session*

			Session 1		Session 2	
Younger			**Pre/Post**	**Pre/Final**	**Pre/Post**	**Pre/Final**
	Females	Mean	-4.39	4.83	5.22	2.44
		Standard Deviation	9.29	6.50	6.60	4.15
		Observations	6	6	6	6
	Males	Mean	-5.89	2.72	-9.72	-10.94
		Standard Deviation	7.17	9.48	11.99	13.20
		Observations	6	6	6	6
Older						
	Females	Mean	10.92	9.58	6.17	5.08
		Standard Deviation	11.65	9.40	9.12	6.19
		Observations	4	4	4	4
	Males	Mean	2.44	7.56	-4.17	1.72
		Standard Deviation	7.07	13.18	7.35	8.86
		Observations	6	6	6	6

*** Negative mean scores indicate that participants' postural equilibrium was degraded in the post or final tests as opposed to the pretest.**

A second three-way MANOVA was performed using *Age*, *Gender*, and *Display* (type) as the independent variables, with the data collapsed across both sessions. Significant ($p \leq 0.05$) main effects of Age (A) {$F(2,17) = 4.75$, $p = 0.0229$} and Gender (G) {$F(2,17) = 6.42$, $p = 0.0084$} were observed. The complete MANOVA summary table for display type is provided in Table 11. At the 0.05 level of significance, the analysis failed to find a significant difference in the ability of participants to maintain postural equilibrium between display types. Means and standard deviations for all possible treatment conditions are shown in Table 12.

TABLE 11. MANOVA Summary Table for Ataxia by Display Type

Source	Effect dF	Error dF	Test	Value	*F*	*p*
Between Subjects						
Age (A)	2	17	Wilks Lambda	0.641436	4.75	**0.0229***
Gender (G)	2	17	Wilks Lambda	0.569734	6.42	**0.0084***
A × G	2	17	Wilks Lambda	0.975317	0.22	0.8086
Within Subjects						
Display (D)	2	17	Wilks Lambda	0.755327	2.75	0.0921
D × A	2	17	Wilks Lambda	0.909455	0.85	0.4463
D × G	2	17	Wilks Lambda	0.927114	0.67	0.5256
D × A × G	2	17	Wilks Lambda	0.938571	0.56	0.5834

***Statistically significant effect at $p \leq 0.05$.**

TABLE 12. Descriptive Statistics for Ataxia Scores by Display Type*

			DV Display		HMD	
Younger			**Pre/Post**	**Pre/Final**	**Pre/Post**	**Pre/Final**
	Females	Mean	3.11	0.94	-2.28	6.33
		Standard Deviation	7.59	2.17	10.50	6.39
		Observations	6	6	6	6
	Males	Mean	-7.00	-8.06	-8.61	-0.17
		Standard Deviation	9.76	12.52	10.35	13.49
		Observations	6	6	6	6
Older						
	Females	Mean	8.58	7.83	8.50	6.83
		Standard Deviation	13.23	10.38	7.67	5.63
		Observations	4	4	4	4
	Males	Mean	-4.39	-0.17	2.67	9.44
		Standard Deviation	7.17	7.65	7.00	12.60
		Observations	6	6	6	6

*** Negative mean scores indicate that participants' postural equilibrium was degraded in the post or final tests as opposed to the pretest.**

Following the MANOVA, ANOVA analyses were performed on each of the dependent variables. Significant ($p \leq 0.05$) main effects of Age (A) {$F(1,18) = 7.49$, $p = 0.0136$} and Gender (G) {$F(1,18) = 10.25$, $p = 0.0050$} for the PRE/POST condition were observed. The significant ($p \leq 0.05$) main effect of Age (A) $\{F(1,18) = 5.91, p = 0.0258\}$ for the PRE/FINAL condition were observed. The complete ANOVA summary tables for PRE/POST and PRE/FINAL scores are presented in Tables 13 and 14. The significant Age and Gender effects for each dependent variable are portrayed graphically in Figures 7, 8, and 9.

The ataxia *Display* (Type) analysis produced several F-ratios that again, are less than one. Nearly all of these cases, however, occurred for the interactions. The TLHTT test is a validated measuring tool, so it is likely that there was considerable error variability in the within-group results and the experiment was not powerful enough to detect changes in the interaction cases (likely due to small sample size). The only time the F-ratio was less than one for a main effect was the 0.00 for *Display* (type) as shown in Table 13. However, *Display* (type) was greater than one in the pre/final condition (Table 14), which suggests that it had absolutely no effect on participants' ability to maintain balance just after the experiment, but may have had some effect on the participants' rates of recovery. It is possible, however, that a lack of power may again have precluded a significant result.

TABLE 13. ANOVA Summary Table for PRE/POST Ataxia Scores by Display Type

Source	dF	SS	MS	F	p
Between Subjects					
Age (A)	1	605.312	605.312	7.49	**0.0136***
Gender (G)	1	828.493	828.493	10.25	**0.0050***
A × G	1	3.721	3.721	0.05	0.8326
Error	18	1455.501	80.861		
Within Subjects					
Display (D)	1	0.001	0.001	0.00	0.9980
D × A	1	130.185	130.185	1.47	0.2418
D × G	1	79.388	79.388	0.89	0.3571
D × A × G	1	7.538	7.538	0.09	0.7742
Error	18	1599.544	88.864		
Total	43				

***Statistically significant effect at $p \leq 0.05$.**

TABLE 14. ANOVA Summary Table for PRE/FINAL Ataxia Scores by Display Type

Source	dF	SS	MS	*F*	*p*
<u>Between Subjects</u>					
Age (A)	1	412.787	412.787	5.91	**0.0258***
Gender (G)	1	290.928	290.928	4.16	0.0563
A × G	1	68.186	68.186	0.98	0.3364
Error	18	1258.017	69.890		
<u>Within Subjects</u>					
Display (D)	1	319.399	319.399	2.65	0.1208
D × A	1	14.529	14.529	0.12	0.7324
D × G	1	114.552	114.552	0.95	0.3423
D × A × G	1	43.866	43.866	0.36	0.5537
Error	18	2167.679	120.427		
Total	43				

***Statistically significant effect at $p \leq 0.05$.**

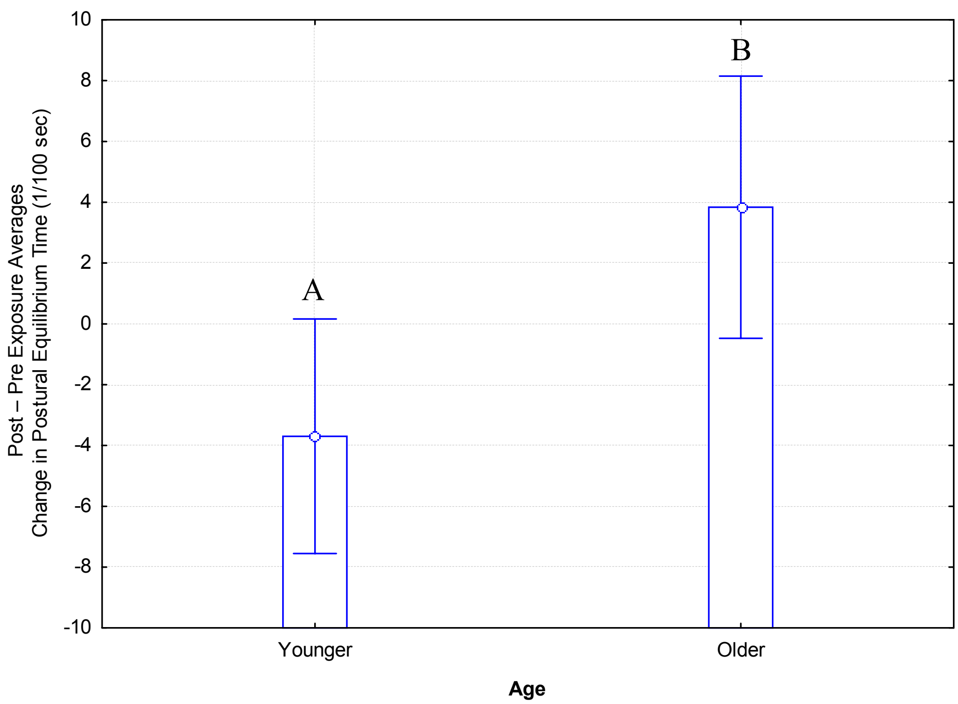

Figure 7. Age Main Effect ($p = 0.0136$) for Pre/Post Ataxia by Display Type (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

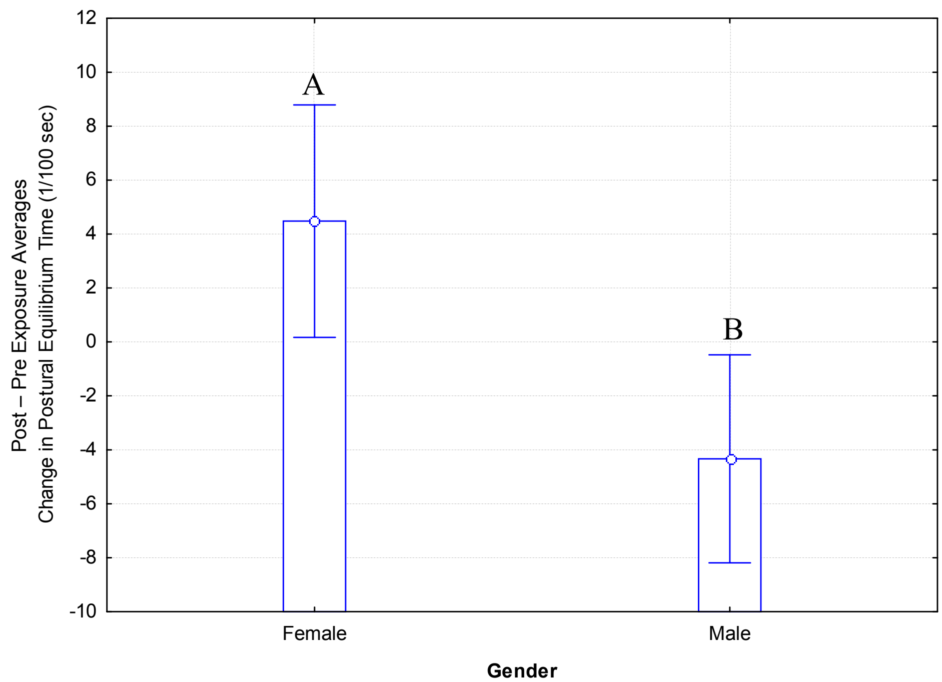

Figure 8. Gender Main Effect ($p = 0.0050$) for Pre/Post Ataxia by Display Type (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

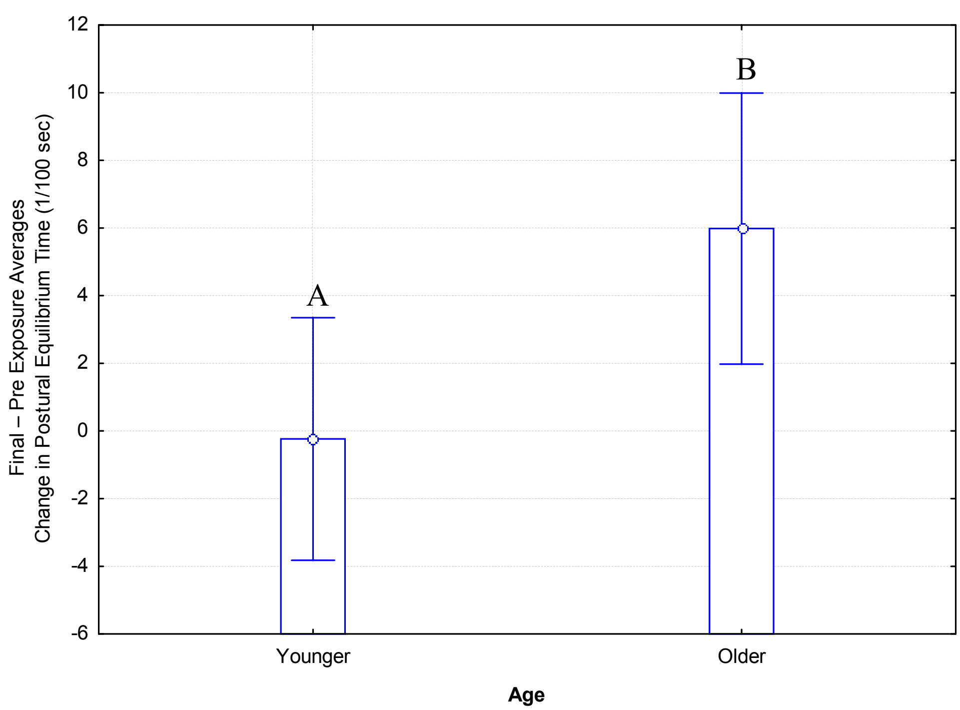

Figure 9. Age Main Effect ($p = 0.0258$) for Pre/Final Ataxia by Display Type (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

Driving performance. Driving performance consisted of three types of dependent measures that were collected by the simulator: time to complete course (TCC), steering wheel angle variance (SAV), and lane deviation (LD). As the three dependent variables were inherently related, a MANOVA analysis was required to investigate the effect of each of the levels of the independent variables on driving performance. Similar to the SSQ and ataxia analyses of the last two sessions, a MANOVA analysis of the driving performance data was first performed using *Session* as a within-subjects variable in order to determine if practice affected the results. The subsequent MANOVA analysis used *Display* (type) as the within-subjects variable.

Before the analysis could begin, the data sets needed to be corrected for an outlier. One participant (an older male) had a steering angle variance data point far greater than all of the other subjects as seen on the steering angle variance plot in Figure 10. This high variance resulted from the city portion of the second Session driving course and this participant's results for the other sections – highway, interstate, and loopback – were consistent with the other participants. The Session 1 results of this participant were also consistent with other participants in the first session. The occurrence of this result in the second Session suggests that this participant had great difficulty *initially* adjusting to the second display type (HMD, in this case) in the highly turn-intensive city portion, and not in the overall driving simulator environment. Although this is a notable observation, this point's inclusion in the data set violates the assumption of data normality.

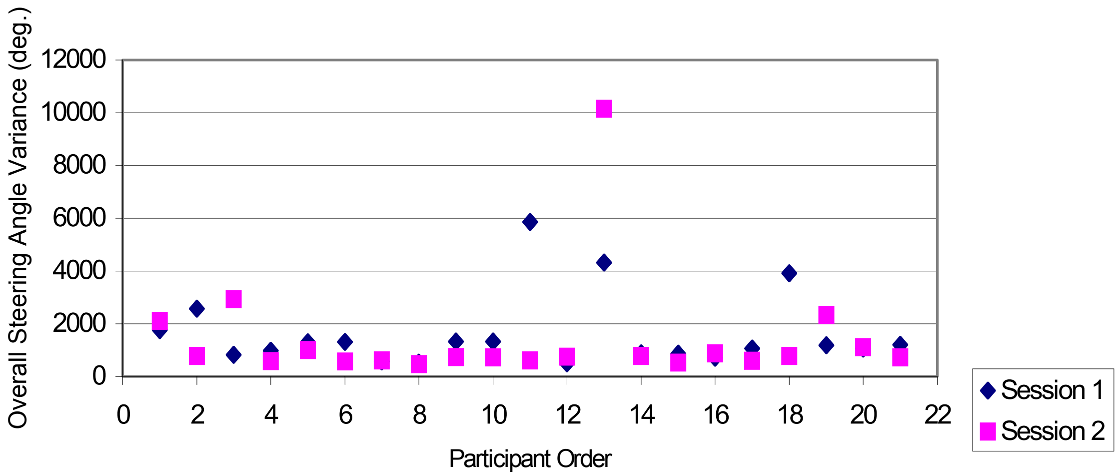

Figure 10. Overall Steering Angle Variance Data by Session

The Dixon Criterion procedure was performed in both the Session and Display cases in order to determine whether this data point was indeed an outlier (Natrella, 1963). The procedure used assumed that a) extreme observations in either direction are considered rejectable and b) the population mean and standard deviation are unknown. For the Session case, n = 21 and r_{22} was calculated to be 0.8156. At a very conservative α value of 0.01, the $r_{1-\alpha/2}$ value is 0.551. Because 0.8156 is greater than 0.551, the Dixon Criteria suggests that the observation can be rejected. A similar result was found in the Display (type) case. This individual participant's results were eliminated in the subsequent MANOVA and ANOVA procedures. This appeared to be the only outlier in the data sets.

A three-way MANOVA was then performed using Age, Gender, and Session as the independent variables. As before, the Type III sum of squares was used to account for incomplete data sets when resolving all main effects and interactions. The Scheffé method of post hoc means comparisons was used to identify statistically significant differences in the means of the dependent measures at each level of the independent variables in cases where the overall model was statistically significant. As noted earlier, the Scheffé test is generally considered very conservative and appropriate for use in designs with unequal sample sizes. A significant ($p \leq 0.05$) main effect of Session (S) {$F(3,14) = 6.22$, $p = 0.0066$} and the interaction Session × Age × Gender (S × A × G) {$F(3,14) = 3.57$, $p = 0.0416$} were observed. The complete MANOVA summary table for order effects is provided in table 15. Means and standard deviations for all possible treatment conditions are shown in table 16.

TABLE 15. MANOVA Summary Table for Driving Performance by Session

Source	Effect dF	Error dF	Test	Value	F	p
Between Subjects						
Age (A)	3	14	Wilks Lambda	0.713480	1.87	0.1804
Gender (G)	3	14	Wilks Lambda	0.926475	0.37	0.7756
A × G	3	14	Wilks Lambda	0.642301	2.60	0.0935
Within Subjects						
Session (S)	3	14	Wilks Lambda	0.428834	6.22	**0.0066***
S × A	3	14	Wilks Lambda	0.980459	0.09	0.9627
S × G	3	14	Wilks Lambda	0.887097	0.59	0.6292
S × A × G	3	14	Wilks Lambda	0.566260	3.57	**0.0416***

***Statistically significant effect at $p \leq 0.05$.**

TABLE 16. Descriptive Statistics for Driving Performance Scores by Session*

			Session 1			**Session 2**	
Younger		**TCC (s)**	**SAV (deg)**	**LD (mm)**	**TCC (s)**	**SAV (deg)**	**LD (mm)**
	Mean	576.9217	1117.3284	574.3907	562.1537	1355.1188	585.3382
Females	Standard Deviation	44.5215	438.0263	140.5012	41.7410	954.3821	136.5380
	Observations	6	6	6	6	6	6
	Mean	579.0220	2120.8320	496.2093	559.3124	620.2046	570.9303
Males	Standard Deviation	50.4853	2248.6578	82.4172	47.3656	108.0643	175.1941
	Observations	5	5	5	5	5	5
Older							
	Mean	716.6133	1902.5131	508.4446	649.8593	800.2221	521.5044
Females	Standard Deviation	138.5123	1345.1198	148.9539	21.6651	233.6422	120.9722
	Observations	4	4	4	4	4	4
	Mean	619.3688	971.9034	669.4639	604.8780	1055.6255	611.5496
Males	Standard Deviation	51.4664	216.2673	114.3485	16.1007	726.6572	88.4527
	Observations	5	5	5	5	5	5

*** Key to scores is provided in the text.**

Following the MANOVA, an ANOVA analysis was performed on each of the dependent variables. The results are presented in tables 17, 18, and 19. Again, there are several instances in which the F-ratios are less than one. The small F-ratios may have resulted from large variances within the groups and a lack of power on the part of the experiment to resolve them. As they were present primarily in the SAV and LD measures, there was likely a large degree of variability within the groups' abilities to control the simulated vehicle. In other words, some older females (for example) controlled the vehicle very well, some not so well, and the experiment lacked the power to afford a generalization about older females' driving performance in the simulator.

TABLE 17. ANOVA Summary Table for TCC Scores by Session

Source	dF	SS	MS	*F*	*p*
Between Subjects					
Age (A)	1	60100	60100	12.25	**0.0030***
Gender (G)	1	12514	12514	2.55	0.1297
A × G	1	12256	12256	2.50	0.1335
Error	16	78470	4904		
Within Subjects					
Session (S)	1	8199	8199	4.39	0.0525
S × A	1	1339	1339	0.72	0.4099
S × G	1	1371	1371	0.73	0.4045
S × A × G	1	2004	2004	1.07	0.3160
Error	16	29918	1870		
Total	39				

***Statistically significant effect at $p \leq 0.05$.**

TABLE 18. ANOVA Summary Table for SAV Scores by Session

Source	dF	SS	MS	F	p
Between Subjects					
Age (A)	1	142960	142960	0.13	0.7237
Gender (G)	1	101227	101227	0.09	0.7660
A × G	1	545357	545357	0.49	0.4923
Error	16	17669489	1104343		
Within Subjects					
Session (S)	1	3186620	3186620	3.19	0.0933
S × A	1	36531	36531	0.04	0.8509
S × G	1	186827	186827	0.19	0.6714
S × A × G	1	5236099	5236099	5.23	**0.0361***
Error	16	16008223	1000514		
Total	39				

***Statistically significant effect at $p \leq 0.05$.**

TABLE 19. ANOVA Summary Table for LD Scores by Session

Source	dF	SS	MS	F	p
Between Subjects					
Age (A)	1	4330	4330	0.19	0.6657
Gender (G)	1	15376	15376	0.69	0.4191
A × G	1	72305	72305	3.24	0.0910
Error	16	357608	22350		
Within Subjects					
Session (S)	1	1020	1020	0.09	0.7662
S × A	1	10430	10430	0.94	0.3478
S × G	1	32	32	0.00	0.9581
S × A × G	1	11116	11116	1.00	0.3328
Error	16	178313	11145		
Total	39				

***Statistically significant effect at $p \leq 0.05$.**

The *Session* variable was not found to be significant in any of the ANOVAs. As it was found to be significant in the MANOVA, however, a Scheffé post-hoc means comparison on *Session* in all three of the dependent variables was performed using the within-subject error term. As shown in Table 20, the average time to complete course (TCC) was significantly greater in Session 2 than Session 1, and lane deviation (LD) was significantly greater in Session 1 than

Session 2. There was no significant difference between the steering angle variance (SAV) means. The *Session* main effect for each dependent variable is portrayed graphically in Figures 11, 12, and 13.

TABLE 20. Scheffé Post Hoc Analysis of Session Main Effect in Driving Performance

	TCC		SAV		LD	
Session	S1 Mean 616.00	S2 Mean 985.54	S1 Mean 589.67	S2 Mean 565.42	S1 Mean 1488.90	S2 Mean 575.52
1		**0.0234***		0.4554		**0.0060***
2	**0.0234***		0.4554		**0.0060***	
df	16		16		16	
Error Within MS	217,500		10039		833,100	

***Statistically significant effect at p ≤ 0.05.**

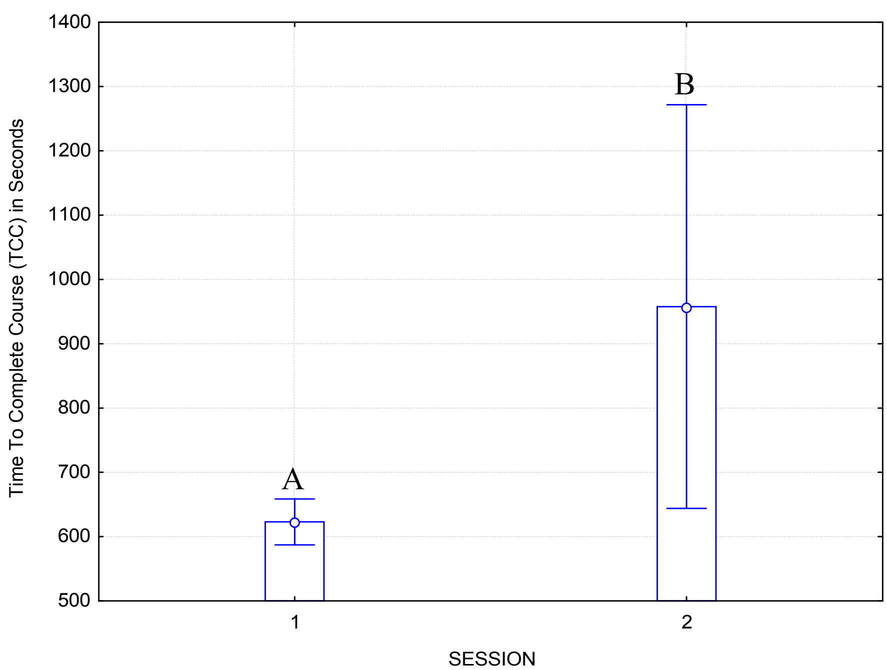

Figure 11. Session Main Effect (p = 0.0066) Least Squares Means for TCC by Session (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

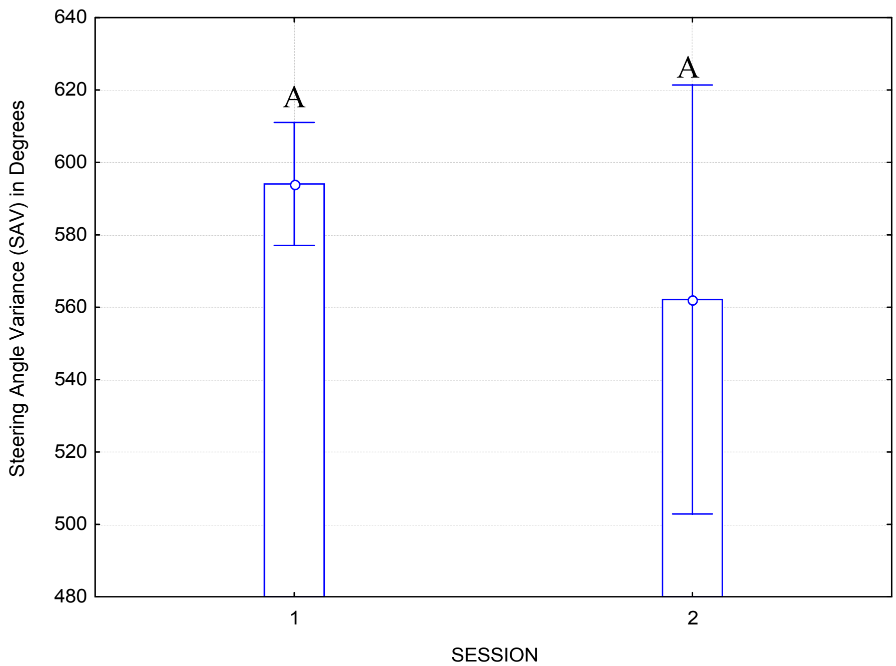

Figure 12. Session Main Effect (p = 0.0066) Least Squares Means for SAV by Session (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

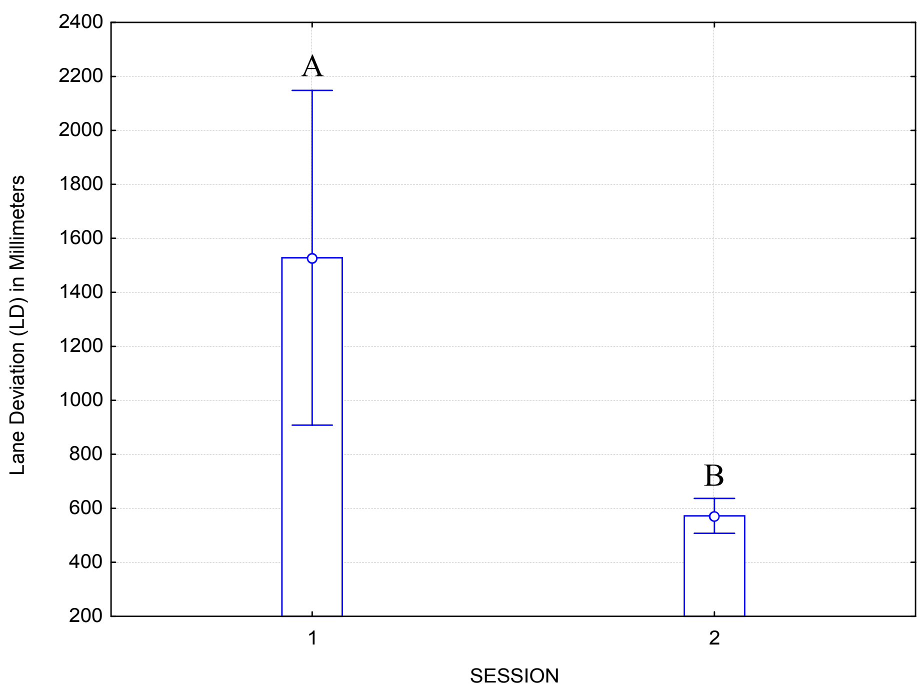

Figure 13. Session Main Effect (p = 0.0066) Least Squares Means for LD by Session (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

The MANOVA also found a significant 3-way interaction among Session × Age × Gender (S × A × G) {$F(3,14) = 3.57$, $p = 0.0416$}. It is a weak effect, but still significant at $p \leq 0.05$. Session × Age × Gender was also found significant in the SAV ANOVA analyses {$F(1,16) = 5.23$, $p = 0.0361$}. The 3-way interaction was decomposed using simple-effects *F*-tests in order to partition the variance on one dimension and save power (J. Casali, personal communication, 1/27/2003).

First, the effect was decomposed into all possible two-way interactions at each level of each variable. The degrees of freedom, sum of squares, and mean squares were calculated using the Type III sum of squares procedure in Statistica. The results were then imported into Microsoft Excel where the *F*-ratios and *p*-values were manually calculated using the error-term from the original 3-way interaction in the SAV ANOVA. The results are shown in Table 21.

TABLE 21. Two-Way Simple Effects *F*-Tests for the S × A × G Interaction

Source	dF	SS	MS	*F*	*p*
Session 1					
A × G	1	4580564	4580564	4.58	**0.0481***
Error	16	26800310	1000514		
Session 2					
A × G	1	1200892	1200892	1.208	0.2895
Error	16	6877402	1000514		
Age Level 1 (Younger)					
G × S	1	4121040	4121040	4.12	0.0594
Error	16	16008223	1000514		
Age Level 2 (Older)					
G × S	1	1562919	1562919	1.56	0.2293
Error	16	16008223	1000514		
Gender Level 1 (Female)					
A × S	1	2154982	2154982	2.15	0.1616
Error	16	16008223	1000514		
Gender Level 2 (Male)					
A × S	1	3137704	3137704	3.14	0.0956
Error	16	16008223	1000514		

***Statistically significant effect at $p \leq 0.05$.**

At this point, however, Statistica returned a “Matrix Inversion” error when performing the A × G calculation. After some investigation, it was found that Age × Gender gives a matrix with coefficients that contain multicolliniarity; i.e., some of the factors that go into the matrix are not independent. With the help of Statistica technical support, a simple correlation analysis was performed (Table 22). A 0.95 correlation was found between Older Males (SAV Session 2) and Younger Males (SAV Session 1), and a perfect (-1) negative correlation between Young Females (SAV Session 1) and Young Males (SAV Session 1).

TABLE 22. SAV Correlation Matrix
(*N*=4; Casewise deletion of missing data)

	YM1	YM2	OF1	OF2	YF1	YF2	OM1	OM2
YM1	1.00	0.33	0.07	-0.33	**-1.00***	-0.28	0.32	**0.95***
YM2	0.33	1.00	0.91	-0.50	-0.37	-0.45	-0.38	0.04
OF1	0.07	0.91	1.00	-0.12	-0.14	-0.67	-0.73	-0.21
OF2	-0.33	-0.50	-0.12	1.00	0.29	-0.50	-0.56	-0.20
YF1	**-1.00***	-0.37	-0.14	0.29	1.00	0.36	-0.24	-0.94
YF2	-0.28	-0.45	-0.67	-0.50	0.36	1.00	0.81	-0.15
OM1	0.32	-0.38	-0.73	-0.56	-0.24	0.81	1.00	0.47
OM2	**0.95***	0.04	-0.21	-0.20	-0.94	-0.15	0.47	1.00

***Statistically significant effect at $p \leq 0.05$.**

As the simple effects on the three-way interaction show the effect restricted to Session 1, the paired differences (*t*-tests) were calculated among the four Age and Gender means on Session 1 to resolve the 3-way interaction (R. Williges, personal communication, 2/21/2003). Statistica was used to perform the calculation (*t*-test for independent groups). The results are shown in Table 23. Younger females are denoted as YF, younger males as YM, older females as OF, and older males as OM. None of the combinations proved to be significant which is likely a result of the correlation shown above and the small sample sizes.

TABLE 23. *t*-Tests for the Session 1 A × G Interaction

Group 1	Group 2	Group 1 Mean	Group 2 Mean	*t*-value	df	*p*	Group 1 Valid *n*	Group 2 Valid *n*	Group 1 Standard Deviation	Group 2 Standard Deviation
YF	YM	1117.33	2120.83	-1.08	9	0.3082	6	5	438.0263	2248.658
OF	OM	1902.51	971.90	1.55	7	0.1653	4	5	1345.120	216.2673
YF	OM	1117.33	971.90	0.67	9	0.5179	6	5	438.0263	216.2673
OF	YM	1902.51	2120.83	-0.17	7	0.8698	4	5	1345.120	2248.658
YF	OF	1117.33	1902.51	-1.36	8	0.2105	6	4	438.0263	1345.120
OM	YM	971.90	2120.83	-1.14	8	0.2883	5	5	216.2673	2248.658

***Statistically significant effect at p ≤ 0.05.**

A three-way MANOVA was then performed using <u>*Age*</u>, <u>*Gender*</u>, and <u>*Display*</u> as the independent variables. As before, the Type III sum of squares procedure was used to account for incomplete data sets when resolving all main effects and interactions. A significant ($p \leq 0.05$) main effect of <u>*Display*</u> (D) {$F(3,14) = 6.27$, $p = 0.0064$} was observed. The complete MANOVA summary table for the Display Type analysis is provided in Table 24. Means and standard deviations for all possible treatment conditions are shown in Table 25.

TABLE 24. MANOVA Summary Table for Driving Performance by Display Type

Source	Effect dF	Error dF	Test	Value	*F*	*p*
<u>Between Subjects</u>						
Age (A)	3	14	Wilks Lambda	0.695185	2.05	0.1536
Gender (G)	3	14	Wilks Lambda	0.834723	0.92	0.4548
A × G	3	14	Wilks Lambda	0.910407	0.46	0.7151
<u>Within Subjects</u>						
Display (D)	3	14	Wilks Lambda	0.426617	6.27	**0.0064***
D × A	3	14	Wilks Lambda	0.923471	0.39	0.7643
D × G	3	14	Wilks Lambda	0.853365	0.80	0.5133
D × A × G	3	14	Wilks Lambda	0.703153	1.97	0.1649

***Statistically significant effect at *p* ≤ 0.05.**

TABLE 25. Descriptive Statistics for Driving Performance Scores by Display Type*

		DV Display			HMD		
Younger		**TCC (s)**	**SAV (deg)**	**LD (mm)**	**TCC (s)**	**SAV (deg)**	**LD (mm)**
	Mean	565.2915	1303.1256	594.0278	573.7838	1169.3215	565.7010
Females	Standard Deviation	48.0714	932.7794	145.0452	38.7454	505.7858	130.1366
	Observations	6	6	6	6	6	6
	Mean	570.2498	2018.4015	516.7743	568.0846	1080.9119	496.9381
Males	Standard Deviation	59.5114	2319.5434	134.5499	38.6073	862.1334	59.9511
	Observations	5	5	5	5	5	5
Older							
	Mean	673.5353	970.6442	543.0083	692.9373	1732.0909	486.9408
Females	Standard Deviation	67.1849	270.1543	185.2688	133.6451	1488.1024	22.6942
	Observations	4	4	4	4	4	4
	Mean	619.8594	1114.3283	667.1047	604.3874	913.2006	613.9088
Males	Standard Deviation	20.0711	725.2366	91.2917	49.8679	167.5596	113.5470
	Observations	5	5	5	5	5	5

*** Key to scores is provided in the text.**

Following the MANOVA, an ANOVA analysis was performed on each of the dependent variables. Again, there are several instances in which the *F*-ratios are less than one. As with the <u>Session</u> analysis, it is possible that the small *F*-ratios may have resulted from large variances within the groups and a lack of power on the part of the experiment to resolve them. The results are presented in tables 26, 27, and 28.

TABLE 26. ANOVA Summary Table for TCC Scores by Display

Source	dF	SS	MS	*F*	*p*
<u>Between Subjects</u>					
Age (A)	1	60100	60100	12.25	**0.0030***
Gender (G)	1	12514	12514	2.55	0.1297
A × G	1	12256	12256	2.50	0.1335
Error	16	78470	4904		
<u>Within Subjects</u>					
Display (D)	1	64	64	0.03	0.8735
D × A	1	4	4	0.00	0.9703
D × G	1	1269	1269	0.52	0.4832
D × A × G	1	359	359	0.15	0.7076
Error	16	39402	2463		
Total	39				

***Statistically significant effect at $p \leq 0.05$.**

TABLE 27. ANOVA Summary Table for SAV Scores by Display

Source	dF	SS	MS	*F*	*p*
<u>Between Subjects</u>					
Age (A)	1	433541	433541	0.35	0.5647
Gender (G)	1	1431	1431	0.00	0.9735
A × G	1	1037996	1037996	0.83	0.3764
Error	16	20063193	1253950		
<u>Within Subjects</u>					
Display (D)	1	159854	159854	0.13	0.7195
D × A	1	1629894	1629894	1.36	0.2602
D × G	1	1910005	1910005	1.60	0.2245
D × A × G	1	15457	15457	0.01	0.9109
Error	16	19138987	1196187		
Total	39				

***Statistically significant effect at $p \leq 0.05$.**

TABLE 28. ANOVA Summary Table for LD Scores by Display

Source	dF	SS	MS	*F*	*p*
<u>Between Subjects</u>					
Age (A)	1	11579	11579	0.58	0.4558
Gender (G)	1	6756	6756	0.34	0.5675
A × G	1	96535	96535	4.87	**0.0423***
Error	16	317163	19823		
<u>Within Subjects</u>					
Display (D)	1	15173	15173	1.63	0.2199
D × A	1	2286	2286	0.25	0.6270
D × G	1	79	79	0.01	0.9277
D × A × G	1	19	19	0.00	0.9642
Error	16	148924	9308		
Total	39				

***Statistically significant effect at $p \leq 0.05$.**

The only variable found significant in the MANOVA was <u>*Display*</u>, although it was not found to be significant in any of the ANOVAs. As such, it was analyzed in the context of all three dependent measures. A Scheffé post hoc means comparison on <u>*Display*</u> was performed in Statistica using the within-subject error term. As shown in Table 29, the average time to complete course (TCC) was significantly greater with the HMD than with the DV, but lane deviation (LD) was significantly greater with the DV than with the HMD. There was no significant difference between the steering angle variance (SAV) means. The ***Display*** main effect is portrayed graphically in Figures 14, 15, and 16.

TABLE 29. Scheffé Post Hoc Analysis of Display Type Main Effect in Driving Performance

	TCC		SAV		LD	
Display	DV Mean 601.82	HMD Mean 1195.7	DV Mean 603.84	HMD Mean 582.78	DV Mean 1368.2	HMD Mean 544.81
DV		**0.0065***		0.5527		**0.0117***
HMD	**0.0065***		0.5527		**0.0117***	
df	16		16		16	
Error Within MS	360,800		12,056		837,500	

***Statistically significant effect at *p* ≤ 0.05.**

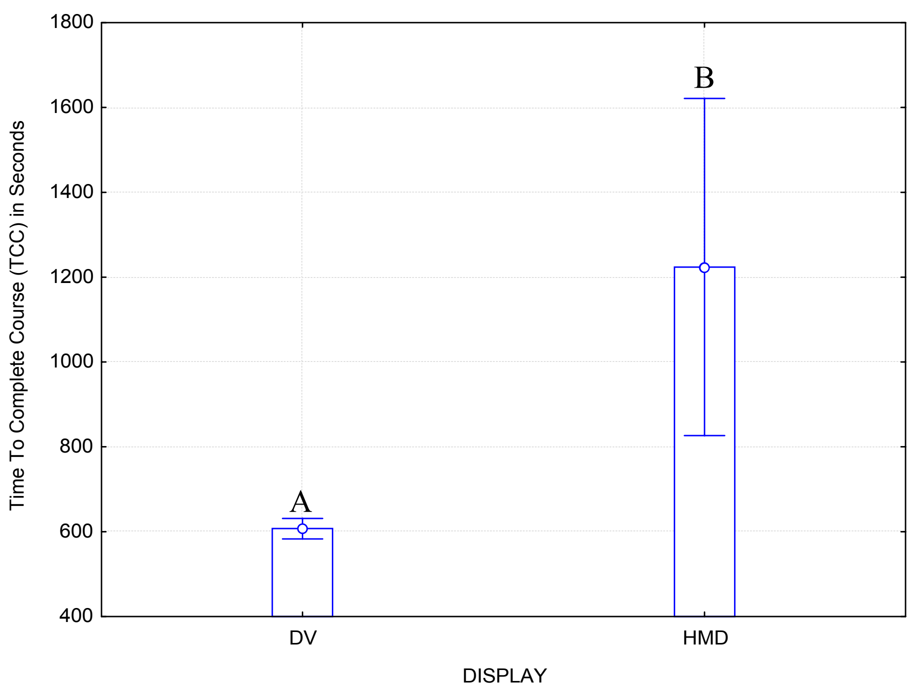

Figure 14. Display Type Main Effect ($p = 0.0064$) Least Squares Means for TCC by Display Type (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

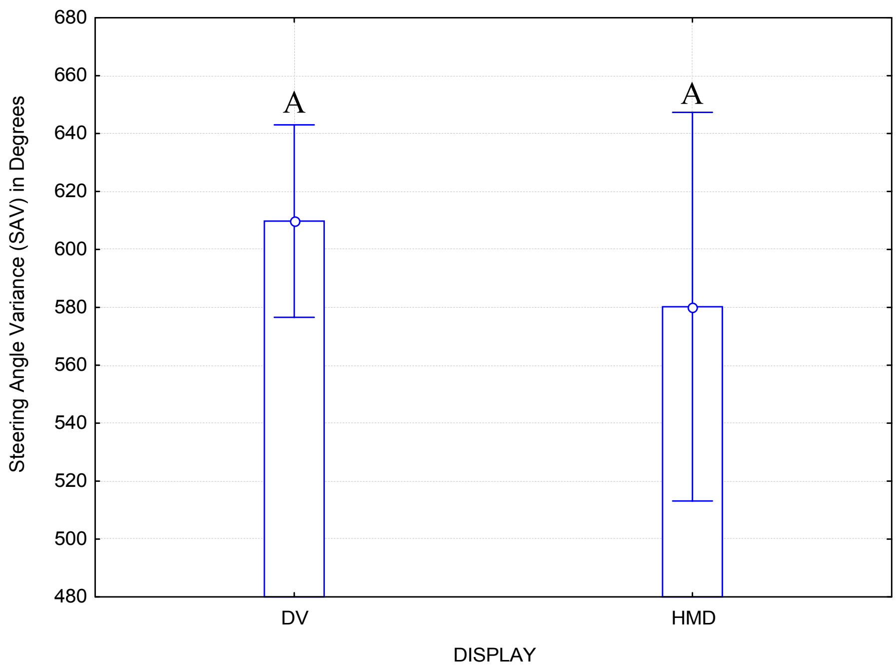

Figure 15. Display Type Main Effect (p = 0.0064) Least Squares Means for SAV by Display Type (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

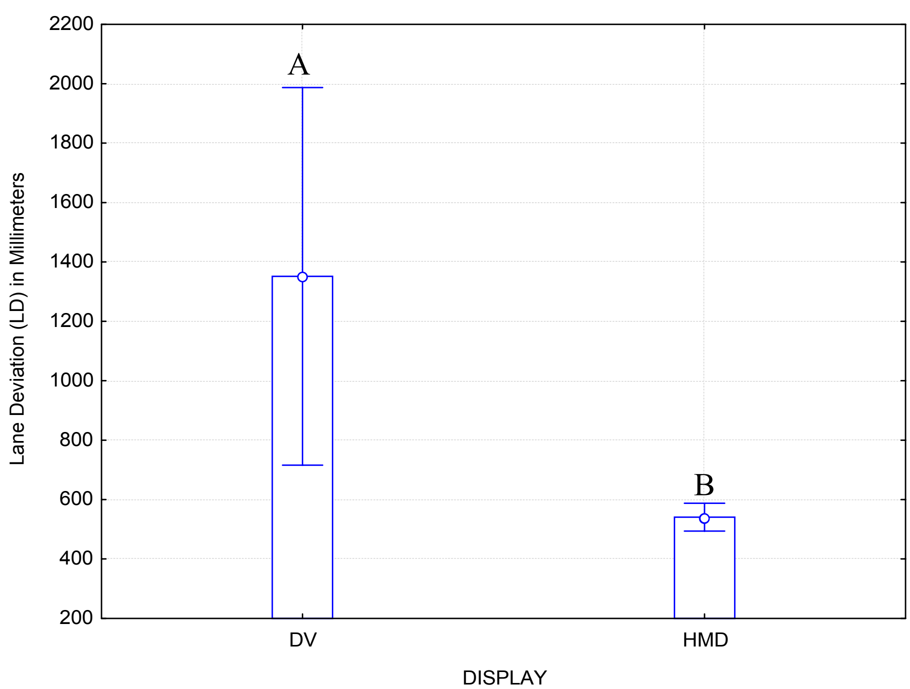

Figure 16. Display Type Main Effect (p = 0.0064) Least Squares Means for LD by Display Type (Vertical bars denote 0.95% confidence intervals; means with different letters are significantly different at $p \leq 0.05$)

Subjective preference. In addition to the numeric data, this study also collected subjective observations with regard to the two display types. Participants were asked which of the two display types they preferred and why. One question on the participant information form asked whether participants found the HMD physically comfortable and another asked which of the two displays they would prefer to use in a simulator-based, DMV driving test and why. Tables 30 and 31 summarize participants' answers.

TABLE 30. Participant Subjective Display Type Preferences

Age	Gender	Session 1 Display	HMD Comfortable?	Preferred Display
Younger	F	DV	No	DV
Younger	F	HMD	Yes	DV
Younger	F	HMD	Yes	DV
Younger	F	DV	No	DV
Younger	F	DV	No	HMD
Younger	F	HMD	Yes	DV
Younger	M	HMD	Yes	HMD
Younger	M	HMD	Yes	DV
Younger	M	HMD	Yes	HMD
Younger	M	DV	No	HMD
Younger	M	DV	Yes	HMD
Younger	M	DV	Yes	HMD
Older	F	DV	Yes	HMD
Older	F	HMD	No	DV
Older	F	HMD	Yes	Neither
Older	F	DV	Yes	HMD
Older	F	DV	Yes	DV
Older	F	HMD	No	DV
Older	M	DV	No	DV
Older	M	DV	Yes	HMD
Older	M	HMD	Yes	DV
Older	M	HMD	Yes	DV
Older	M	DV	Yes	HMD
Older	M	HMD	No	HMD

TABLE 31. Summary of Participants' Subjective Preferences

	Young Females		Younger Males		Older Females		Older Males		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%
DV Display										
Preferred Display	5	83%	1	17%	3	50%	3	50%	12	50%
HMD										
Uncomfortable	3	50%	1	17%	2	33%	2	33%	8	33%
Preferred Display	1	17%	5	83%	2	33%	3	50%	11	46%

A Chi-Squared Goodness-of-Fit test was employed to determine the significance of participant display type preference. As shown in Table 31, 12 participants preferred the direct-view display and 11 preferred the HMD. The participant that did not express a preference for either display was not used in the calculation, which resulted in an expected value of 11.5 for the sample. The results of this test $\{\chi^2\ (1) = 0.02,\ p = 0.8875\}$ indicate that there was no significant overall preference for display type.

DISCUSSION

Simulator-induced sickness

There is very little literature available that has directly investigated age and gender susceptibility to simulator-induced sickness in conjunction with display type. One of the goals of this study was to determine which display type – direct-view (DV) or head-mounted – induced greater feelings of discomfort in various population types when performing a DMV-type driving task. Because the literature is sparse in this area, it is difficult to discuss the results of this portion of the study in context with previous research.

Session main effect. Because the experiment was designed with display type as a within-subjects variable, it was important to determine if practice effects carried over from the first session to the second. An analysis was first performed that collapsed across *Display* and used *Session* as an independent, within-subjects variable. Unfortunately, the analysis did find a significant session main effect, which indicates that the results from the second session were contaminated by those from the first. Participants experienced significantly greater discomfort in the first session than they did in the second – regardless of display type.

Kennedy and Fowlkes (1992) suggest that repeated measures designs are not appropriate for the analysis of simulator-induced sickness due to the robustness of the adaptation effect. They support this conclusion by citing studies that examined simulator-induced sickness with multiple, long-duration exposures in a short time frame. It was felt that this study could largely avoid these adaptation effects because: 1) simulator exposure in each session was limited to 15 minutes, 2) a week elapsed between trials, and 3) two radically different display types were employed that effectively created new and different environments.

The significance of *Session* as a main effect supports the findings in the literature of the robustness of the adaptation effect. It would seem that adaptation is an even more robust effect than even Kennedy and Fowlkes (1992) have suggested, as it can occur even with short exposures separated by long gaps in time. Subsequent analysis examined the significance of *Display* (type) as a within-subjects variable. Table 5 shows that it was not found to be significant, but even if it had been, its significance would have been suspect due to the observed practice effects. *As such, this study cannot conclude that one display type is more prone to producing simulator-induced sickness effects than the other.*

Age main effect. Despite the significance of the *Session* main effect, the significant *Age* main effect is still an important finding. In this study, the older participants experienced significantly more simulator-induced sickness discomfort than the younger participants. The physiological effects of aging on the visual system have been well established. Due to reduced visual acuity contrast sensitivity and reduced visual accommodation resulting from age, it was hypothesized that the older participants would experience significantly greater discomfort with the HMD than the DV. This would have been demonstrated by a significant Display × Age interaction effect, which did not occur. *Older participants experienced greater discomfort than younger persons – regardless of display type.*

This finding supports the work of Liu et al. (1999), who reported that simulator-induced sickness actually increased with age in their study involving an HMD-based, medium-fidelity driving simulator. Kolasinski (1995) reiterates that conflicts are thought to occur between an actual pattern of stimuli and an expected pattern of stimuli stemming from repeated, real-world experiences. In the case of flight simulators, Kennedy et al. (1990) suggested that the pilot's experience with the sensory aspects of actual flight might lead to a greater sensitivity in a simulator experience versus what he or she would expect from the real world task. It is not unreasonable to suggest that users of a driving simulator might react the same way; that is, older persons with greater driving experience may be more sensitive to sensory conflicts and may thus show a greater incidence of sickness than those with less.

However, there is likely a point of diminishing returns with age and experience. Ignoring any other external factors, a person who, for example, is 25 years old and has been driving for nine years is unlikely to have a significantly different expected pattern of stimuli than someone who is 40 years old and has been driving for 24 years. Although expectations of the driving task may be influenced by numerous other factors such as type of driving (city vs. highway), style of driving (defensive vs. aggressive), etc., for an overall population, it is unlikely that the occurrence of simulator-induced sickness in a driving simulator in all age ranges results only from real-world experience. It is more likely that the simulator-induced sickness produced in older persons would probably be a result of the stress of being in an unfamiliar situation rather than expectations of real-world conditions.

It would not be prudent, however, to completely discount the effect of experience. Although participants were screened in order to minimize experience with VR/HMD type display systems in simulators, it would have been very difficult to find (younger) participants with no video game experience. Four of the 12 older participants noted at least one experience with a driving-type video or arcade game in their lifetimes (with one unknown). Ten of the 12 younger participants also noted that they had at least one experience in their lifetimes with these types of games (also with one unknown). Seven younger participants and two older participants noted that they had had a driving video game experience in the last 10 years. These younger participants are also likely to have experienced other types of games with fast-moving visual scenes, which may have increased their overall adaptation to the vection effect. This difference lends some support to the theory that one's susceptibility to simulator-induced sickness is more of a function experience with simulator devices (Kennedy et al., 1990) rather than physiological changes in the visual system due to aging.

The effects of experience, however, may be dampened by age-related decreases in the sensitivity of the vestibular system. Shupert (1992) notes that the vestibular system has been shown to degrade in persons over 55, and is most pronounced in persons over 70 years of age. An older person who has been driving for 10 years or so with a severely degraded vestibular system may "forget" what the driving gravito-inertial forces feel like, resulting in visual/vestibular conflicts that may not be as pronounced as those of younger persons. This would lead one to expect that simulator-induced sickness would occur less with older people; however, this study found that it occurred more. This is likely due to a relatively healthy older participant pool. An experiment that uses older persons showing a loss in vestibular function may experience different results.

One possible use of the driving simulator is for driver license testing, as discussed in Peck and Wachtel (1993). The users of a DMV driving simulator will likely be individuals retesting for a license due to any factor that may call into question their ability to safely operate an automobile. As users will likely have a range of real-world experience with the driving task, and older persons demonstrate a greater propensity to experience simulator-induced sickness, it is important that a DMV considering the use of driving simulator devices should also allow for alternatives when testing an older population.

Gender main effect. This study did not find a significant difference in the degree of simulator-induced sickness experience between males and females. This supports the findings of Kolasinski (1996) and S. A. Jones (1998) who also found no significant difference in susceptibility to simulator-induced sickness based on gender. As discussed earlier, this is a topic in the literature with contradictory findings, but this study was not able to reject the hypothesis that there is no difference between the genders.

Ataxia

This study sought to determine if display type affected the rates of recovery for ataxic effects. Because the <u>*Session*</u> variable was not found to be a significant effect for ataxia, it is unlikely that a participant's participation in the first session affected his or her results for the second. As such, the following discussion is based on the <u>*Display*</u> analysis.

In the previous section, an ANOVA analysis on <u>*Age*</u>, <u>*Gender*</u>, and <u>*Display Type*</u>, was performed on each of the dependent variables following the MANOVA. For the purposes of this discussion, it is useful to reintroduce the results of the ataxia statistical analysis. As noted earlier, a negative mean difference indicates that the amount of time a participant could maintain their postural equilibrium was less in the post-exposure or final conditions than in the pre-exposure condition. A positive mean difference indicates that participants were able to maintain their postural equilibrium longer in the post-exposure or final conditions than in the pre-exposure condition. In other words, a positive mean value would indicate that the simulation had no negative impact on a participant's ability to maintain postural equilibrium. Table 32 provides a summary of the results.

TABLE 32. Summary of ANOVA Age and Gender Main Effects Results in Ataxia

	PRE/POST		PRE/FINAL	
Age	Younger Mean -3.69	Older Mean 2.90	Younger Mean -0.24	Older Mean 5.72
p-value	**0.0136***		**0.0258***	
Gender	Female Mean 3.67	Male Mean -4.33	Female Mean 5.12	Male Mean 0.26
p-value	**0.0050***		0.0563	

***Statistically significant effect at $p \leq 0.05$.**

From the signed average of the summary data, the results show that younger participants could not maintain their postural stability as long in the post and final conditions as they could in the pre-exposure condition. Also, there was no apparent degradation in the ability of older participants to maintain their postural stability in the post and final conditions than they could in the pre-exposure condition. This strongly suggests that the simulation did not affect older participants' postural stability and the time increases likely resulted from Session-to-Session practice of the TLHTT test. More significantly, Table 32 shows that the younger participants had a significant reduction in postural equilibrium just after, and 30 minutes following simulator exposure as compared to the older participants.

The conclusion is that older people are less likely to have their immediate and long-term postural equilibrium affected by a simulation environment (as measured by ataxia) than younger people. To put it another way, older persons are less likely to be affected by a simulation environment just after exposure than younger persons, and will tend to readjust more quickly to the real world environment following the simulation.

Of particular interest is the finding that although the older participants showed significantly higher rates of simulator-induced sickness discomfort than the younger participants, their postural equilibrium was less likely to be affected by the simulation. This supports the results of Gower et al. (1988), Kennedy et al. (1995), and Kolasinski and Gilson (1999) who found that participants who have adapted to simulator environments have reported a decrease in simulator-induced sickness symptoms and a concurrent worsening of post-exposure ataxia. As the significance of the *Session* variable in the simulator-induced sickness analysis showed, participants did adapt to the simulation. However, younger people demonstrated significantly lower TSS scores in Session 2 than older participants (Table 8), which suggests that they showed a greater degree of adaptation to the simulation. Thus it follows that because older participants had lower levels of adaptation to the simulation environment (as evidenced by greater discomfort levels), they would demonstrate lower post-exposure ataxia than younger people. This is consistent with the findings of this study.

Although it is likely that this effect would disappear in time due to adaptation over repeated exposures, it has implications for DMVs considering use of a driving simulator device before an on-road skills test (Peck & Wachtel, 1993). A younger person performing well in a simulator is more likely to be disoriented after the simulation than an older person, and may perform worse in the road test as a result. Of course, an older person may also perform poorly on a road test due to a simulator-induced headache and feelings of nausea. Organizations such as DMVs or driver education schools must carefully weigh the benefits that simulators will add to drivers license testing (or training) against the potential for degraded performance that may be elicited during the on-road portion.

Male participants had significantly more reduction in postural equilibrium just after simulator exposure than did female participants (table 31). In other words, male participants could not hold their balance as long in the post-exposure condition as they could in the pre-exposure condition. Female participants were actually able to hold their balance longer in the post-exposure condition than they could in the pre-exposure condition. There was no significant difference between female and male participants between the final and pre-exposure conditions. The overall conclusion is that the males' postural equilibrium is significantly more affected than the females' just after simulation exposure; however, both groups were, on average, able to fully recover 30 minutes following the exposure.

Driving Performance

This study used time to complete course (TCC), steering angle variance (SAV), and lane deviation (LD) as indices of overall driving performance. As discussed earlier, all participants were instructed to maintain the posted speed limits as closely as possible. Participants who took longer to complete the course were not as comfortable driving at the posted speed limits, and thus may have had greater difficulty operating the simulator. A higher SAV implied that the steering wheel was manipulated more often than a lower total SAV, suggesting that operators with higher variances made more steering corrections and did not steer the vehicle as "smoothly" as an operator with a lower variance. Lane deviation measured the average distance a participant's vehicle deviated from the lane centerline. A higher average LD suggests that the participant had greater difficulty in smoothly controlling the vehicle.

Session effect. In order to examine the possibility of practice effects contaminating the driving performance results, a three-way MANOVA was performed using *<u>Session</u>* as the within-subjects, independent variable. As shown in Table 20, the average TCC and LD variables are significantly different from Session 1 to Session 2, but steering angle variance was not. From this, one may conclude that participants required more time to complete the course in Session 2, regardless of age, gender, or display type; however, their lane deviation scores significantly decreased from Session 1 to Session 2. Although one may expect that TCC would correlate positively to LD, the results suggest that participants took greater care in their driving during the second Session (the lower LD) than the first, which required them to take more time to complete the course.

The experimenter did sense that some participants were unhappy with their performance in the first Session and wished to "do better" during the second. It is likely that participants judged it more important to maintain steady control of the vehicle rather than maintain the posted speed limits so they drove slower, but had less lane deviation. Despite the care taken to provide a practice run for participants in order to even out practice effects, it still seems that Session proved to be a contaminating effect.

Session × Age × Gender (S × A × G) was a significant effect in the SAV ANOVA. After a simple effects interaction decomposition and a series of *t*-tests, none of the effects proved to be significant. This is likely due to the small sample sizes and the correlation found in the Younger Male/Younger Female and Younger Male/Older Male cases.

Display main effect. Following the Session analysis, a three-way MANOVA was performed using *Display* (Type) as the within-subjects, independent variable. As shown in Table 29, the average time to complete course and lane deviation variables are significantly different between display types. It is apparent that participants required more time to complete the course using the HMD regardless of age or gender; however, their lane deviation scores significantly decreased when they drove the course with the HMD. Steering angle variance did not change significantly.

It may be tempting to attribute the improved performance to an increase in presence afforded by the HMD (Hendrix & Barfield, 1996); however, this conclusion is questionable for at least two reasons. First, the 26˚ field of view of the i-glasses HMD does not approach the recommended 100˚ FOV required for true immersion (Levine & Mourant, 1995; Stanney et al., 1998). Second, one must also consider the corresponding increase in the time to complete the course. It is more likely that the participants took more care in driving the course with the HMD because it was thought to be more difficult to control the vehicle with this display and headtracking system.

Despite the "drift" of the headtracker unit (Penhallegon et al., 2002) and the varying needs to recenter it, it is unlikely that the headtracker negatively affected driving performance in a significant way. Although one might expect that participants with higher rates of recentering may have found it more difficult to stay centered in the travel lane, this study found that lane deviation was significantly less with the HMD than the direct-view display. Although these results are also affected by the care in which the driver operated the vehicle, the fact that drivers did seem to control the vehicle better with the HMD indicates that headtracker recentering did not significantly affect driving performance. Any firm conclusions regarding the effect of *Display* on driving performance are suspect, however, due to the significance of the *Session* variable as described above.

Subjective Preference

As shown in Tables 30 and 31, the subjective preference results show an apparent preference for the direct-view display on the part of the younger female participants (83%) and an apparent preference for the HMD on the part of the younger males (83%). 33% of participants reported that the HMD was "uncomfortable," the majority being female participants. The older male participants were evenly divided as to whether they would prefer to operate a simulator with either display type and females were divided three-to-two in favor of the direct-view, with one participant noting that neither display was suitable, in its present form, for this type of application. Despite the within-group differences, however, the results of the chi-squared test suggest that there was no apparent overall preference for one display type or the other.

In addition, more than half of the participants (16 out of 23, or 70%) preferred the display type that they used in the second Session. As participants were significantly less sick in the second Session as opposed to the first, it is likely that the degree of sickness they felt influenced their choice. Some of the subjective comments indicated as much, with more than one participant explaining that a display was preferred because it did not make them feel as “dizzy” or “nauseous.” This does not necessarily hold, however, for the younger participants who were largely in favor of a particular display type – regardless of which they used first. Based on the subjective comments, participants who preferred the HMD did so because it gave them greater peripheral and side-to-side vision, and they generally felt more immersed in a driving environment. They felt that the HMD/headtracking system enhanced their situational awareness, which made them more aware of hazards and obstacles in the driving environment. This is consistent with the findings of Hendrix and Barfield (1996) who suggested that HMDs provide a greater sense of presence than direct-view displays.

Participants who preferred the direct-view display generally felt that they were able to better control the vehicle with this type. While wearing the HMD, some participants, particularly those in the older group, had difficulty separating the steering of the vehicle from the turning of the head. They reported feeling that the vehicle would turn in the direction in which they were looking; not where they were actually steering. Although it was explained to them that as long as the steering wheel was held steady and in a neutral position, the vehicle would travel in a straight line irrespective of where they were looking, this instinct proved difficult for some participants to overcome. Consequently, they felt more in control of the vehicle with the fixed direct-view display. This is more of an issue with the HMD/headtracker system, however, than for the HMD alone.

The majority of participants found the HMD to be comfortable enough in terms of its physical ergonomics (67%). The younger females were evenly split, but four out of six older males and older females found the HMD to be comfortable. Only one younger male found the HMD uncomfortable. The most common complaints included weight of the device and a tendency for it to feel like it was not securely affixed to the head – even when correctly adjusted. Several participants suggested a third strap that would go over the top of the head in addition to those around the sides. Two older participants also noted that the device sat very close to their eyeglasses, which made them feel uncomfortable.

Implications for Display Type Selection and Use

Although this study was unable to draw clear conclusions with respect to display type selection, several observations will be of interested to persons or institutions looking to employ simulation devices for commercial, public, or research purposes. The selection of display type is perhaps the single most important decision a designer can make with regards to maximizing simulator utility and minimizing the potential for simulator-induced sickness. Each class, HMD and direct-view, affects viewers in different ways, and either may be the more appropriate choice for differing applications.

As discussed in the literature review section of this study, HMDs place unique stressors on the visual system. They sit quite close to the eyes, placing a greater demand on the accommodation/convergence system; they typically have high contrast ratios with the blacked-out housing, lower resolutions, and limited fields of view; and may have transport delays associated with headtracking. Widescreen, direct-view displays, such as the one used in this study, may create a greater sensation of vection in the peripheral vision, leading to greater severity of simulator-induced sickness.

Despite the reduced ocular stress afforded by a direct-view display, the designer must still ensure careful integration into a simulator. The simulation software for this study was optimized to run at either a 1024 x 768 or 800 x 600 resolution. The NEC widescreen plasma display, however, had a native resolution of 1365 x 768. In order to fill the screen with the image, the display "stretched" the video signal, which resulted in various aliasing artifacts. These included the disappearance and reappearance of lane markers on the course roads and occasional flickering of the simulated foliage. Although not studied directly, these could have acted as contributing factors to the simulator-induced sickness discomfort felt by participants. As a result, improper integration of a direct-view display may result in severe discomfort by its viewers.

The selection of display type must ultimately be based on the intended *uses* and *users* of the simulator. A simulator intended to be used often by the general public, such as in a DMV-type application, may wish to preclude the use of an HMD due to practical issues such as hygiene, ruggedness of the device, the additional cognitive demands of headtracking, and the resulting high learning curve for controlling the vehicle. A single direct-view display may be easier to employ than an HMD even if it results in a greater weight and footprint; however, designers should ensure that the resolution of the video signal matches that of the display.

The appropriate display type selection for research applications should also take the research goals into account during the selection. Driving simulator research that is intended to study issues such as target detection in the forward visual field may not need the additional presence afforded by an HMD. The greater degree of practicality inherent in a direct-view display may be the most important issue. However, research that requires a strong sense of “being there” in the virtual environment, and cannot afford multiple direct-view displays, may find an HMD to be the more appropriate choice.

Experimental Limitations

The fundamental limitation of this study was that it used a repeated measures design to investigate simulator-induced sickness. Even with short exposures and no less than a week between sessions, it was found that participants adapted to the simulation environment. Although this is a notable finding in and of itself, it presents a difficulty when trying to attribute differences in simulator-induced sickness to either of the two display types.

This study was also limited by the simulator device itself. The VT-UVA-Carilion Driving Simulator was constructed as a “proof-of-concept” type device with “medium” fidelity in the spectrum of driving simulators. Several participants noted that the sensation of operating the simulator was far removed from that of driving a car, so it is difficult to extrapolate driving performance results to the real-world environment. In particular, the simulator lacked rear-view capability, a seatbelt, and a working gearshift.

Additionally, the “drift” of the headtracker unit (Penhallegon et al., 2002) and the varying needs to recenter it may have contributed to anxiety and workload of the participants. Occasionally during a simulation run, participants were required to reach up and press the recentering button on the headtracker. For some of the older participants, the experimenter had to perform this operation as they found it difficult to control the vehicle with one hand. The number of times the headtracker needed to be recentered varied by participant – sometimes it only needed to be recentered once or twice, other times it was close to a dozen. The need to recenter appeared to occur randomly.

As noted earlier, this is unlikely to have contaminated the driving performance data. However, the sudden scene changes resulting from a recentering operation could have produced feelings of discomfort or post-exposure postural disequilibrium. The simulator-induced sickness results showed no clear difference between display types; however, this is likely due to the practice effects resulting from the experimental design. A determination of the exact effect of headtracker recentering would need to be performed in a controlled environment.

The chosen HMD may also limit the generalizability of the results. The i-glasses are a relatively inexpensive, consumer HMD with limited resolution and FOV. Although this is the type of HMD likely to be chosen for a cost-conscious DMV, it may not be appropriate to generalize the results of this study when more sophisticated devices are employed.

Finally, the relatively limited subject pool could be a potential limiting factor. An effort was made to recruit participants randomly; however, most were from the Blacksburg area and the younger persons were all college students. The backgrounds of the older participants were more diversified, but most were college-educated and had a basic comfort level with technology. It is possible that the subject pool was skewed and that individuals that were willing to participate in this experiment were not representative of the overall population.

CONCLUSIONS

A review of the literature suggests that HMDs create a greater sense of presence in virtual environments, which should improve driving performance. However, greater presence can also lead to greater sensory conflicts resulting in increased incidence of simulator-induced sickness, VRISE, and postural instability. This study hoped to be able to recommend a particular display type that would be appropriate for use with particular age/gender groups in a DMV or training-application driving simulator. Unfortunately, because the experiment used a repeated measures design, simulator-induced sickness and driving performance results for display type are suspect due to the presence of practice effects.

Despite this, the study showed that older participants experienced significantly increased simulator-induced sickness discomfort than the younger participants – regardless of display type. This suggests that DMVs or driver training schools considering the use of medium fidelity simulator devices to augment or replace on-road driver license testing must be aware of potential difficulties when used by older persons. There was no significant difference found between genders, however.

Perhaps the most important finding of this study was that although the older participants showed significantly higher rates of simulator-induced sickness discomfort than the younger participants, they recovered their postural equilibrium significantly faster than the younger people, at least as measured by the TLHTT test. This supports the results of those who suggest that participants who have adapted to simulator environments have reported a decrease in simulator-induced sickness symptoms and concurrent worsening of post-exposure ataxia. Although it is likely that this effect would disappear over time, it has implications for those who are considering the use of a driving simulator device before an on-road skills test. Display type was not found to affect the degree of ataxia experienced by participants.

A seemingly contradictory finding was shown in driving performance as participants required more time to complete the course using the HMD, but showed a significant decrease in the degree of lane deviation. Although this may be due in part to an increase in presence afforded by the HMD, a more likely explanation is that participants took more care in driving the course with the HMD because it was more difficult to control the vehicle with this display and headtracking system.

At this time, it is unlikely that a low-cost driving simulator such as the one developed by Virginia Tech, the University of Virginia, and Northeastern University would deliver great utility to the Department of Motor Vehicles. Due to the inherent limited-fidelity of the device driven by cost constraints, operators did not really feel as though they were controlling an actual automobile. As a result, the proof-of-concept simulator is too crude to replace on-road driver license testing. Validation for less demanding applications, such as screening for who should and should not take an on-road skills test, is also not likely to happen in the short term. Still, a basic simulator may be useful to a DMV for scenario or rule-based testing such as on-road navigation, hazard identification and avoidance, determining when other vehicles can legally be passed, and street sign and traffic signal compliance.

SUGGESTIONS FOR FUTURE RESEARCH

Unfortunately the main thrust of this study was somewhat marred by the use of the repeated subjects experimental design. As participants did appear to benefit from practice as they operated the simulator in the second session, it is not possible to assign differences in the main effects to the type of display used. More specifically, this research was not able to show a significant difference between the direct-view display and the HMD in terms of simulator-induced sickness, ataxia, or driving performance. A future study seeking to quantify the effect of display type should use a completely between-subjects design.

One positive conclusion of this outcome is that it shows that adaptation is an even more robust effect than Kennedy and Fowlkes (1992) have suggested, as it can occur even with short exposures separated by long gaps in time. A possibility for future research may be a study of how long these gaps need to be before practice effects disappear – if at all. Quantifying rates of adaptation for a particular display type could be extremely useful for institutions seeking to use them as a tool in the public domain.

Additionally, although not directly measured, the experimenter noted through personal observation that participants appeared to be much more at ease in the second session than the first. Although the experimenter tried to make participants feel comfortable during both sessions, participants knew what to expect from the simulation and from the overall experimental protocol in the second session. As such, they did not feel that the experiment might surprise them at times and thus appeared to be more relaxed and confident. This may have been a reason why simulator-induced sickness was not as prevalent in the second session. A future study might look at the relationship between anxiety level and simulator-induced sickness – particularly with older populations.

This HMD system used in this study included a headtracker device. As some of simulator-induced sickness symptoms may be attributable to the presence of the headtracker, it would be difficult to discuss simulator-induced sickness findings with regard to the HMD only. Future research might look at whether the addition of headtracking increases or decreases one's chances of feeling ill, and compare pure display types for the effect on simulator-induced sickness, i.e. a direct-view display to an HMD without headtracking. By eliminating the headtracker, a future study might be able to isolate the effects of the display system itself.

Summary of Recommendations:

- Avoid the use of repeated measures designs when studying the effect of equipment characteristics on simulator-induced sickness.
- Minimize headtracker drift when using an HMD as the primary display system.
- Provide peripheral vision, rear and side-view mirrors, and a seatbelt to improve the fidelity of the simulation.
- Structure research as to minimize the amount of time older people are subjected to the virtual environment.
- Maintain display resolution consistency between the output video signal of the computer and the primary display.
- Use a direct-view display when practicality is more important than presence.
- Use driving environment software with realistic effects such as non-segmented curves, discrete trees, three-dimensional structures, and strong capabilities for customizing the driving environment.

REFERENCES

Anderson, G. J., & Braunstein, M. L. (1985). Induced self-motion in central vision. *Journal of Experimental Psychology: Human Perception and Performance, 11*(2), 122-132.

Arthur, E. J., Hancock, P. A., & Chrysler, S. T. (1993). *Spatial orientation in real and virtual worlds.* Paper presented at the Human Factors and Ergonomics Society 37th Annual Meeting, Seattle.

Baltzley, D. R., Kennedy, R. S., Berbaum, K. S., Lilienthal, M. G., & Gower, D. W. (1989). The time course of postflight simulator sickness symptoms. *Aviation, Space, and Environmental Medicine, 60*(11), 1043-1048.

Barfield, W., & Hendrix, C. (1995). The effect of update rate on the sense of presence within virtual environments. *Virtual Reality, 1*(1), 3-16.

Barfield, W., Hendrix, C., Bjorneseth, O., Kaczmarek, K. A., & Lotens, W. (1995). Comparison of human sensory capabilities with technical specifications of virtual environment equipment. *Presence: Teleoperators and Virtual Environments, 4*(4), 329-356.

Barfield, W., & Kim, Y. (1991). Computer graphics programming principles as factors in the design of perspective displays. In H. J. Bullinger (Ed.), *Human aspects in computing: Design and use of interactive systems and work with terminals* (pp. 93-97). Amsterdam: Elsevier Science Publishers B. V.

Barfield, W., & Weghorst, S. (1993, August). *The sense of presence within virtual environments: A conceptual framework.* Paper presented at the 5th International Conference on HCI, Orlando.

Barfield, W., Zeltzer, D., Sheridan, T. B., & Slater, M. (1995). Presence and performance within virtual environments. In W. Barfield & T. A. Furness (Eds.), *Virtual environments and advanced interface design* (pp. 473-513). New York: Oxford University Press.

Benson, A. J. (1987, May). *Aetiological factors in simulator sickness.* Paper presented at the Motion Cues in Flight Simulation and Simulator Induced Sickness: AGARD Conference Proceedings No. 433, Neuilly Sur Seine, France.

Biocca, F. (1992). Will simulation sickness slow down the diffusion of virtual environment technology? *Presence: Teleoperators and Virtual Environments, 1*(3), 334-343.

Brinkley, J. (1998, September). Flatter is better? *Stereophile Guide to Home Theater, 4,* 80-91.

Brinkley, J. (2000, July/August). Fujitsu Plasmavision Slimscreen PDS 4221. *Stereophile Guide to Home Theater, 6,* 72-75.

Bruno, P. (1993, August). *Eye discomfort and "VDT" work.* Paper presented at the 5th International Conference on HCI, Orlando.

Casali, J. G. (1980). *Evaluation of various design alternatives influencing discomfort in a driving simulator.* Paper presented at the Manned Systems Design: Methods, Equipment, and Applications, Freiburg, West Germany.

Casali, J. G. (1986). *Vehicular simulation-induced sickness, Volume I: An overview* (Technical Report NTSC-TR-86-012). Orlando, FL: Naval Training Systems Center.

Casali, J. G., & Frank, L. H. (1986). Perceptual distortion and its consequences in vehicular simulation: Basic theory and incidence of simulator sickness. *Transportation Research Record*(1059), 57-65.

Casali, J. G., & Wierwille, W. W. (1980). The effects of various design alternatives on moving-base driving simulator discomfort. *Human Factors, 22*(6), 741-756.

Casali, J. G., & Wierwille, W. W. (1986). Potential design etiological factors of simulator sickness and a research simulator specification. *Transportation Research Record*(1059), 66-74.

Cobb, S. V. G., Nichols, S., Ramsey, A., & Wilson, J. R. (1999). Virtual reality-induced symptoms and effects (VRISE). *Presence: Teleoperators and Virtual Environments, 8*(2), 169-186.

Crosby, T. N., & Kennedy, R. S. (1982). *Postural disequilibrium and simulator sickness following flights in a P-3C operational flight trainer.* Paper presented at the 53rd Annual Scientific Meeting of the Aerospace Medical Association, Bal Harbor.

Czaja, S. J. (1996). Aging and the acquisition of computer skills. In W. A. Rogers & A. D. Fisk & N. Walker (Eds.), *Aging and Skilled Performance* (pp. 201-220). Mahwah, NJ: Lawrence Erlbaum Associates.

Dichgans, J., & Brandt, T. (1978). Visual vestibular interactions: Effects on self-motion perception and postural control. In R. M. Held & H. W. Leibowitz & H. L. Teuber (Eds.), *Handbook of Sensory Physiology* (Vol. 8, pp. 755-804). Berlin: Springer-Verlag.

Diewald, W. (1995, July-August). New developments in advanced driving simulators. *TR News,* 23-26.

DiZio, P., & Lackner, J. R. (1992). Spatial orientation, adaptation, and motion sickness in real and virtual environments. *Presence: Teleoperators and Virtual Environments, 1*(3), 319 - 328.

Ebenholtz, S. M. (1992). Motion sickness and oculomotor systems in virtual environments. *Presence: Teleoperators and Virtual Environments, 1*(3), 302-304.

Ehrlich, J. A. (2000). *Effect of viewing conditions on sickness and distance estimation in a virtual environment* (1112). Alexandria: U.S. Army Research Institute for the Behavioral and Social Sciences.

Ellis, S. R., & Bucher, U. J. (1994). *Distance perception of stereoscopically presented virtual objects optically superimposed on physical objects by a head-mounted see-through display.* Paper presented at the Human Factors and Ergonomics Society 38th Annual Meeting, Nashville.

Ellis, S. R., & Menges, B. M. (1995). *Judged distance to virtual objects in the near visual field.* Paper presented at the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego.

Fowlkes, J. E., Kennedy, R. S., & Lilienthal, M. G. (1987, October 19-23). *Postural disequilibrium following training flights.* Paper presented at the Human Factors and Ergonomics Society 31st Annual Meeting, New York.

Frank, L. H. (1985). *Simulator sickness: An upcoming problem.* Unpublished dissertation proposal, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Frank, L. H., & Casali, J. G. (1986). *Simulator sickness: A review of its costs, countermeasures, and prediction.* (SAE No. 861782): Society of Automotive Engineers Technical Paper Series.

Frank, L. H., Casali, J. G., & 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), 201-217.

Frank, L. H., Kennedy, R. S., McCauley, M. E., & Kellogg, R. S. (1983). *Simulator Sickness: Reaction to a transformed perceptual world. I. Scope of the problem* (NAVTRAEQUIPCEN TN-65). Orlando, FL: Naval Training Equipment Center.

Fregly, A. R. (1974). Vestibular ataxia and its measurement in man. In H. H. Kornhuber (Ed.), *Handbook of Sensory Physiology, Part 2, Psychophysics and General Interpretation* (pp. 321-360). New York: Springer-Verlag.

Gigante, M. A. (1994). Virtual reality: enabling technologies. In R. A. Earnshaw & M. A. Gigante & H. Jones (Eds.), *Virtual reality systems* (pp. 15-25). London: Academic Press.

Gower, D. W., Lilienthal, M. G., Kennedy, R. S., & Fowlkes, J. E. (1988). *Simulator sickness in US Army and Navy fixed and rotary-wing flight simulators.* Paper presented at the AGARD Medical Symposium on Motion Cues in Flight Simulation and Simulator Induced Sickness, Neuilly-sur-Seine, France: NATO Advisory Group for Aerospace Research and Development.

Guedry, F. E. (1970, August 19-21). *Conflicting sensory orientation cues as a factor in motion sickness.* Paper presented at the Fifth Symposium on the Role of the Vestibular Organs in Space Exploration, Pensacola.

Guerrier, J. H., Manivannan, P., Pacheco, A., & Wilkie, F. L. (1995). *The relationship of age and cognitive characteristics of drivers to performance of driving tasks on an interactive driving simulator.* Paper presented at the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego.

Held, R. M., & Durlach, N. I. (1987). *Telepresence, time delay, and adaptation*. NASA Conference Publication 10032.

Held, R. M., & Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators and Virtual Environments, 1*(1), 109-112.

Hendrix, C., & Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments, 5*(3), 274-289.

Hettinger, L. J., Berbaum, K. S., Kennedy, R. S., Dunlap, W. P., & Nolan, M. D. (1990). Vection and simulator sickness. *Military Psychology, 2*(3), 171-181.

Hettinger, L. J., & Riccio, G. E. (1992). Visually induced motion sickness in virtual environments. *Presence: Teleoperators and Virtual Environments, 1*(3), 306-310.

Howarth, P. A., & Costello, P. J. (1997). The occurrence of virtual simulation sickness symptoms when an HMD was used as a personal viewing system. *Displays, 18*, 107-116.

Jones, J. (1998). *Math 170 - Introduction to Statistics Lecture Notes*. Richland Community College. Retrieved, from the World Wide Web: http://www.richland.cc.il.us/james/lecture/m170/ch13-dif.html

Jones, S. A. (1998). *Effects of restraint on vection and simulator sickness.* Unpublished dissertation, University of Central Florida, Orlando, FL.

Kalawsky, R. S. (1994). Critical aspects of visually coupled systems. In R. A. Earnshaw & M. A. Gigante & H. Jones (Eds.), *Virtual reality systems* (pp. 203-212). London: Academic Press.

Kawara, T., Ohmi, M., & Yoshizawa, T. (1996). Effects on visual functions during tasks of object handling in virtual environment with a head mounted display. *Ergonomics, 39*(11), 1370-1380.

Kellogg, R. S., & Gillingham, K. K. (1986). *United States Air Force experience with simulator sickness, research and training.* Paper presented at the Human Factors and Ergonomics Society 30th Annual Meeting, Dayton.

Kennedy, R. S., Berbaum, K. S., & Smith, M. G. (1993, October 11-15). *Methods for correlating visual scene elements with simulator sickness incidence.* Paper presented at the Human Factors and Ergonomics Society 37th Annual Meeting, Seattle.

Kennedy, R. S., & Fowlkes, J. E. (1992). Simulator sickness is polygenic and polysymptomatic: Implications for research. *The International Journal of Aviation Psychology, 2*(1), 23-38.

Kennedy, R. S., & Frank, L. H. (1985). *A review of motion sickness with special reference to simulator sickness* (Technical Report NAVTRAEQUIPCEN 81-C-0105-16). Orlando, FL: Naval Training Equipment Center, Human Factors Division.

Kennedy, R. S., Hettinger, L. J., Harm, D. L., Ordy, J. M., & Dunlap, W. P. (1996). Psychophysical scaling of circular vection (CV) produced by optokinetic (OKN) motion: Individual differences and effects of practice. *Journal of Vestibular Research, 6*(5), 331-341.

Kennedy, R. S., Hettinger, L. J., & Lilienthal, M. G. (1990). Simulator sickness. In G. H. Crampton (Ed.), *Motion and Space Sickness* (pp. 317-341). Boca Raton, FL: CRC Press.

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

Kennedy, R. S., Lane, N. E., Lilienthal, M. G., Berbaum, K. S., & Hettinger, L. J. (1992). Profile analysis of simulator sickness symptoms: Application to virtual environment systems. *Presence: Teleoperators and Virtual Environments, 1*(3), 295-301.

Kennedy, R. S., Lanham, D. S., Drexler, J. M., & Lilienthal, M. G. (1995). A method for certification that aftereffects of virtual reality exposures have dissipated: Preliminary findings. In A. C. Bittner & P. C. Champney (Eds.), *Advances in industrial ergonomics and safety* (Vol. VII, pp. 231-237). London: Taylor & Francis.

Kennedy, R. S., Lanham, D. S., Massey, C. J., Drexler, J. M., & Lilienthal, M. G. (1994). Gender differences in simulator sickness incidence: Implications for military virtual reality systems. *SAFE Journal, 25*(1), 69-76.

Kennedy, R. S., Lilienthal, M. G., & Fowlkes, J. E. (1991, August 12-14). *What needs doing about simulator sickness?* Paper presented at the AIAA Flight Simulation Technologies Conference, New Orleans.

Kennedy, R. S., & Stanney, K. M. (1996). Postural instability induced by virtual reality exposure: Development of a certification protocol. *International Journal of Human-Computer Interaction, 8*(1), 25-47.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (Third ed.). Upper Saddle River, NJ: Prentice Hall.

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

Kolasinski, E. M. (1996). *Prediction of simulator sickness in a virtual environment.* Unpublished Dissertation, University of Central Florida, Orlando, FL.

Kolasinski, E. M., & Gilson, R. D. (1999). Ataxia following exposure to a virtual environment. *Aviation, Space, and Environmental Medicine, 70*(3), 264-269.

Koretz, J. F., Handelman, G. H., & Brown, N. P. (1984). Analysis of human crystalline lens curvature as a function of accommodative state and age. *Vision Research, 24*(10), 1141-1151.

Leibowitz, H. W., & Post, R. B. (1982). The two modes of processing concept and some implications. In J. Beck (Ed.), *Organization and representation in perception* (pp. 343-363). Hillsdale, NJ: Lawrence Erlbaum Associates.

Levine, O. H., & Mourant, R. R. (1995, January). *A driving simulator based on virtual environments technology.* Paper presented at the Transportation Research Board 74th Annual Meeting, Washington DC.

Liu, L., Watson, B., & Miyazaki, M. (1999). VR for the Elderly: Quantitative and qualitative differences in performance with a driving simulator. *Cyberpsychology & Behavior, 2*(6), 567-576.

Ma, J., Hollerbach, J. M., & Hunter, I. W. (1993). Optical design for a head-mounted display. *Presence: Teleoperators and Virtual Environments, 2*(3), 185-202.

McCauley, M. E., & Sharkey, T. J. (1992). Cybersickness: Perception of self-motion in virtual environments. *Presence: Teleoperators and Virtual Environments, 1*(3), 311-318.

McDowall, I. (1994, May). 3D stereoscopic data for immersive displays. *AI Expert, 9,* 18-21.

McGreevy, M. W., & Ellis, S. R. (1986). The effect of perspective geometry on judged direction of spatial information instruments. *Human Factors, 28*(4), 439-456.

McLane, R. C., & Wierwille, W. W. (1975). The influence of motion and audio cues on driver performance in an automobile simulator. *Human Factors, 17*(5), 488-501.

McLean, J. R., & Hoffman, E. R. (1975). Steering reversals as a measure of driver performance and steering task difficulty. *Human Factors, 17*(3), 248-256.

Mead, S. E., Spaulding, V., Sit, R. A., Meyer, B., & Walker, N. (1997). *Effects of age and training on world wide web navigation strategies.* Paper presented at the Human Factors and Ergonomics Society 41st Annual Meeting, Albuquerque.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems, E77-D*(12), 1321-1329.

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

Morgan, M. W. (1986). Changes in visual function in the aging eye. In A. A. Rosenbloom & M. W. Morgan (Eds.), *Vision and Aging: General and Clinical Perspectives* (pp. 121-134). New York: Professional Press Books/Fairchild Publications.

Natrella, M. G. (1963). *Experimental Statistics*. National Bureau of Standards Handbook 91. Washington, D.C.: U.S. Department of Commerce.

Neveu, C., Blackmon, T., & Stark, L. (1998). Evaluation of the effects of a head-mounted display on ocular accommodation. *Presence: Teleoperators and Virtual Environments, 7*(3), 278-289.

Nichols, S. (1999). Physical ergonomics of virtual environment use. *Applied Ergonomics, 30*, 79-90.

Okhravi, N. (1997). *Manual of primary eye care*. Oxford: Butterworth-Heinemann.

Olsen, E. C. B. (1995). *Can a simulator serve the purpose of an on-the-road examination? A study of the promises and pitfalls of replacing the on-the-road evaluation.* Paper presented at the Silicon Valley Ergonomic Institute's 1st Annual Conference (ErgoCon '95), San Jose.

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

Pausch, R., Crea, T., & Conway, M. (1992). A literature survey for virtual environments: Military flight simulator visual systems and simulator sickness. *Presence: Teleoperators and Virtual Environments, 1*(3), 344-362.

Pausch, R., Shackelford, M. A., & Proffitt, D. (1993). *A user study comparing head-mounted and stationary displays.* Paper presented at the IEEE Symposium on Research Frontiers in Virtual Reality, San Jose.

Peck, R. C., & Wachtel, J. (1993). *New technology for an old problem: A report on a prospective study to evaluate a simulator-based approach to driver licensing* (Committee Report). Washington DC: Transportation Research Board.

Peli, E. (1998). The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display. *Vision Research, 38*, 2053-2066.

Penhallegon, W. J., Perala, C. H., Robinson, G. S., & Casali, J. G. (2002). *Development of the VT-UVA-Carilion Driving Simulator* (ISE Department Technical Report 200201, Audio Lab Number 3/18/02-1-HP). Blacksburg, VA: Virginia Tech, Grado Department of Industrial and Systems Engineering, Auditory Systems Laboratory.

Pierce, J. S., Pausch, R., Sturgill, C. B., & Christiansen, K. D. (1999). Designing a successful HMD-based experience. *Presence: Teleoperators and Virtual Environments, 8*(4), 469-473.

Pimentel, K., & Teixeira, K. (1995). *Virtual reality: Through the new looking glass* (Second ed.). New York: McGraw-Hill.

Reason, J. T., & Brand, J. J. (1975). *Motion sickness*. London: Academic Press.

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

Rinalducci, E. J. (1996). Characteristics of visual fidelity in the virtual environment. *Presence: Teleoperators and Virtual Environments, 5*(3), 330-345.

Robinett, W. (1992). Synthetic experience: A proposed taxonomy. *Presence: Teleoperators and Virtual Environments, 1*(2), 229-247.

Roehl, B. (1996, March/April). A comparison of low-cost HMDs. *Virtual Reality Special Report,* 19-22.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1999). Navigating large scale virtual environments: What differences occur between helmet-mounted and desk-top displays? *Presence: Teleoperators and Virtual Environments, 8*(2), 157-168.

Rushton, S., Mon-Williams, M., & Wann, J. P. (1994). Binocular vision in a bi-ocular world: New generation head-mounted displays avoid causing visual deficit. *Displays, 15*(4), 255-260.

Rushton, S., & Riddell, P. M. (1999). Developing visual systems and exposure to virtual reality and stereo displays: Some concerns and speculations about the demands on accommodation and vergence. *Applied Ergonomics, 30*, 69-78.

Russell, G., & Miles, R. (1993). Volumetric visualization of 3D data. In S. R. Ellis (Ed.), *Pictorial Communications in Virtual and Real Environments (2nd ed.)* (pp. 131-137). London: Taylor and Francis.

Sato, K., Kimura, M., & Abe, A. (1992). Intelligent manipulator system with nonsymmetric and redundant master-slave. *Journal of Robotic Systems, 9*(2), 281-290.

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

Shupert, C. L. (1992). *The aging vestibular system*. Vestibular Disorders Association. Retrieved, from the World Wide Web: http://www.vestibular.org/aging.html

Slater, M., & Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments, 2*(3), 221-233.

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

Stanney, K. M., Kennedy, R. S., & Drexler, J. M. (1997). *Cybersickness is not simulator sickness.* Paper presented at the Human Factors and Ergonomics Society 41st Annual Meeting, Albuquerque.

Stanney, K. M., Kennedy, R. S., Drexler, J. M., & Harm, D. L. (1999). Motion sickness and proprioceptive aftereffects following virtual environment exposure. *Applied Ergonomics, 30*, 27-38.

Stanney, K. M., Mourant, R. R., & Kennedy, R. S. (1998). Human factors issues in virtual environments: A review of the literature. *Presence: Teleoperators and Virtual Environments, 7*(4), 327-351.

StatSoft Inc. (2002). STATISTICA (data analysis software system), Version 6. www.statsoft.com

Stoffregen, T. A., & Riccio, G. E. (1991). An ecological critique of the sensory conflict theory of motion sickness. *Ecological Psychology, 3*(3), 159-194.

Stuart, R. (1996). *The design of virtual environments*. New York: McGraw-Hill.

Thomley, K. E., Kennedy, R. S., & Bittner, A. C. (1986). Development of postural equilibrium tests for examining environmental effects. *Perceptual and Motor Skills, 63*, 555-564.

Uliano, K. C., Kennedy, R. S., & Lambert, E. Y. (1986). *Asynchronous visual delays and the development of simulator sickness.* Paper presented at the Human Factors and Ergonomics Society 30th Annual Meeting, Dayton.

Weir, D. H., & Clark, A. J. (1995, February 27 - March 2). *A survey of mid-level driving simulators.* Paper presented at the SAE International Congress and Exposition, Detroit.

Wilson, J. R. (1996). Effects of participating in virtual environments: a review of current knowledge. *Safety Science, 23*(1), 39-51.

Wilson, J. R. (1997). Virtual environments and ergonomics: needs and opportunities. *Ergonomics, 40*(10), 1057-1077.

Witmer, B. G., Bailey, J. H., & Knerr, B. W. (1996). Virtual spaces and real world places: Transfer of route knowledge. *International Journal of Human-Computer Studies, 45*, 413-428.

Witmer, B. G., & Singer, M. J. (1994). *Measuring presence in virtual environments* (Technical Report 1014). Alexandria, VA: US Army Research Institute.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225-240.

Worden, A., Walker, N., Bharat, K., & Hudson, S. (1997). *Making computers easier for older adults to use: area cursors and sticky icons.* Paper presented at the Human Factors in Computing Systems (CHI 97).

Yeow, P. T., & Taylor, S. P. (1989). Effects of short-term VDT usage on visual functions. *Optometry and Vision Science, 66*(7), 459-466.

Yoo, Y. H. (1999). *Prediction and quantification of individual differences in susceptibility to simulator sickness In fixed-base simulators.* Unpublished Dissertation, University of Central Florida, Orlando, FL.

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

Appendix A

Development of the VT-UVA-Carilion Driving Simulator

Report withheld.
Please contact author for a copy.

Appendix B

Participant Information Forms

Session 1 Participant Information Form

Participant #______

Instructions: Please fill in the blank or circle the appropriate response.

1. What is your age? ________ years

2. What is your gender? Female Male

3. Are you currently in your usual state of good fitness? Yes No

4. How many hours did you sleep last night? __________ hours

5. Was it sufficient? Yes No

6. Please indicate all medications/substances you have used in the past 24 hours. Circle all that apply:

    a. None
    b. Sedatives or tranquilizers
    c. Aspirin, Tylenol, other analgesics
    d. Anti-histamines
    e. Decongestants
    f. Other ______________________________________

7. Do you have a history of seizures? Yes No

8. Do you have a history of motion sickness? If yes, please explain.

9. How many hours per week do you use a computer? ________ hours

10. My level of confidence in using computers is: 1 2 3 4 5
    low average high

11. I enjoy video games (home or arcade) 1 2 3 4 5

disagree unsure agree

12. I am _____ at playing video games. 1 2 3 4 5

bad average good

13. How many hours per week do you play video games? _______ hours per week

14. Have you ever experienced an HMD-based virtual reality game or entertainment?

Yes No

15. Visual acuity ___________

16. Color-blind: Yes No

To be filled out following the experiment

17. You used the (HMD / large-screen display) this session.

18. If you used the HMD, what did you think of it?

19. Was it comfortable? Yes No

If not, why not?

20. If you used the large-screen display, what did you think of it?

Session 2 Participant Information Form

Participant #_____

Instructions: Please fill in the blank or circle the appropriate response.

1. How many hours did you sleep last night? _________ hours

2. Was it sufficient? Yes No

3. Please indicate all medications/substances you have used in the past 24 hours. Circle all that apply:

   a. None
   b. Sedatives or tranquilizers
   c. Aspirin, Tylenol, other analgesics
   d. Anti-histamines
   e. Decongestants
   f. Other ________________________________

4. How many times in the last month have you experienced an HMD-based virtual reality game or entertainment? _______ times

5. Did you experience any lingering discomfort following the previous session? If so, please list your specific symptoms below and note how long they lasted.

Symptom	Duration
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________

To be filled out following the experiment

6. You used the (HMD / large-screen display) this session.

7. If you used the HMD, what did you think of it?

8. Was it comfortable? Yes No

   If not, why not?

9. If you used the large-screen display, what did you think of it?

10. If you had to use a simulator such as this for a DMV-based driving test, which display type would you prefer to use and why?

Appendix C

Pre- and Post-Exposure Simulator Sickness Questionnaires (SSQ)

from

(Kennedy, Lane et al., 1993)

and

(Yoo, 1999)

Participant #_____

Pre-Exposure Simulator Sickness Questionnaire

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

		(0)	(1)	(2)	(3)
1.	General Discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Headache	None	Slight	Moderate	Severe
4.	Eye Strain	None	Slight	Moderate	Severe
5.	Difficulty Focusing	None	Slight	Moderate	Severe
6.	Increased Salivation	None	Slight	Moderate	Severe
7.	Sweating	None	Slight	Moderate	Severe
8.	Nausea	None	Slight	Moderate	Severe
9.	Difficulty Concentrating	None	Slight	Moderate	Severe
10.	Fullness of Head	None	Slight	Moderate	Severe
11.	Blurred Vision	None	Slight	Moderate	Severe
12.	Dizzy (Eyes Open)	None	Slight	Moderate	Severe
13.	Dizzy (Eyes Closed)	None	Slight	Moderate	Severe
14.	Vertigo*	None	Slight	Moderate	Severe

* Vertigo is a disordered state in which the person or his/her surroundings seem to whirl dizzily; giddiness

15.	Stomach awareness**	None	Slight	Moderate	Severe

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

16.	Burping	None	Slight	Moderate	Severe

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

Participant #_____

Post-Exposure Simulator Sickness Questionnaire

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

		(0)	(1)	(2)	(3)
1.	General Discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Headache	None	Slight	Moderate	Severe
4.	Eye Strain	None	Slight	Moderate	Severe
5.	Difficulty Focusing	None	Slight	Moderate	Severe
6.	Increased Salivation	None	Slight	Moderate	Severe
7.	Sweating	None	Slight	Moderate	Severe
8.	Nausea	None	Slight	Moderate	Severe
9.	Difficulty Concentrating	None	Slight	Moderate	Severe
10.	Fullness of Head	None	Slight	Moderate	Severe
11.	Blurred Vision	None	Slight	Moderate	Severe
12.	Dizzy (Eyes Open)	None	Slight	Moderate	Severe
13.	Dizzy (Eyes Closed)	None	Slight	Moderate	Severe
14.	Vertigo*	None	Slight	Moderate	Severe

* Vertigo is a disordered state in which the person or his/her surroundings seem to whirl dizzily; giddiness

15.	Stomach awareness**	None	Slight	Moderate	Severe

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

16.	Burping	None	Slight	Moderate	Severe

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

Appendix D

Request for Approval of Research Proposal

Participant Informed Consent

Request for Approval of Research Proposal
Grado Department of Industrial and Systems Engineering
Virginia Polytechnic Institute and State University

Title: Effects of display type, age, and gender on driving performance and the susceptibility to simulator-induced sickness in a medium-fidelity, fixed-base driving simulator

Principle Investigators: Dr. John G. Casali, John Grado Professor, ISE Department Head
Dr. Gary S. Robinson, Senior Research Associate, ISE
Mr. William J. Penhallegon, Graduate Candidate, ISE

I. Project Justification

Simulators have been developed and employed over the centuries for training purposes and to help us understand human actions and decisions in real world situations. They provide many advantages over field-testing in terms of cost savings, reduction of hazards, and reduction of time to complete tasks under observation. One of the most important goals in simulation is a strong sense of *presence*, or a convincing level of similarity to the real world environment being simulated. Creating this level of realism is one of the major challenges facing any type of virtual environment simulation. Head-mounted displays (HMDs) are becoming increasing attractive to simulator designers as they have been shown to increase the sense of presence and improve operator task performance over a traditional, desktop display. However, HMDs are also associated with greater stress on the visual system, which can induce motion sickness-like symptoms in some individuals. This phenomenon is often referred to as simulator-induced sickness and symptoms range from fatigue and drowsiness, to, more rarely, postural disequilibrium, nausea, and vomiting.

Although direct-view displays have been employed in simulators for years, HMDs have only recently become sufficiently advanced and cost-effective enough to be considered as a simulator's primary display device. However, as HMDs are still an emerging technology, there is little research that directly compares the two devices – particularly in terms of simulator-induced sickness and use with elderly populations. This study will compare the use of a head-mounted display with that of a direct-view display in a medium-fidelity driving simulator to test for differences in driving performance and simulator-induced sickness based on the age and gender of a population. Additionally, this study will examine ataxic effects (post-simulation postural disequilibrium) of HMD use in a driving simulator to determine if they are any more pronounced than those experienced with the use of a direct-view display.

As the Virginia DMV is interested in placing driving simulators in regional offices for license testing, and as driving simulator designers are becoming increasingly interested in the use of HMDs, it is important that potential side effects such as simulator-induced sickness in the entire range of potential users is well understood. This research will be a first step in investigating any increased risk of HMD use in devices intended for use by the general public. The potential liability inherent in sickness and ataxic effects following exposure makes understanding and prediction of individual susceptibility very important. As HMD technology improves, virtual reality (VR) use in simulators as well as in a variety of other professional and consumer applications requires that operators understand any risks beyond those with traditional direct-view systems. The results of this research may help engineers and researchers build safer, more effective (realistic), and lower-cost driving simulators for use in studying the effects of "humans in the loop" in virtual environments.

Problem Statement

It has been shown that head mounted displays can create a greater sense of presence in a virtual environment, however, HMDs can also place greater stress on the visual system and may put people more at risk to simulator-induced sickness. Simulator-induced sickness is the tendency of many vehicular simulators, including both driving and flight devices, to induce acute, residual, and sometimes aftereffect motion sickness-like symptoms of discomfort in operators and passengers. Symptoms include, but are not limited to: general discomfort, apathy, drowsiness, headache, disorientation, fatigue, pallor, sweating, salivation, stomach awareness, nausea, and, in rare cases, vomiting. This research will attempt to determine if display type (HMD, Direct-view) has an effect on simulator-induced sickness and driving performance in a medium-fidelity, PC-based driving simulator.

The primary goal of this research is not about predicting the onset of simulator-induced sickness, but determining how one display type affects the population expected to use the simulator as compared to the other display type. Each type, HMD or direct-view, may be usable for simulation applications in and of itself, but one may be more appropriate for certain population types than the other. As a DMV driving simulator would be intended for use by the entire population of potential drivers license applicants, the goal of this research will be to provide a general recommendation for the use of a particular display type, with certain parameters, for use with specific populations.

Age and gender are factors in motion sickness that show inconsistent results in the literature. Additionally, there is very little research overall on the effects of gender and age to susceptibility of *simulator-induced* sickness in a traditional simulator or virtual reality environment. As senior citizens would be a portion of potential users of the DMV driving simulator, it is important to understand not only how their driving performance compares against a younger population, but also if they are more susceptible to simulator-induced sickness. Several studies have found that females are more susceptible to motion sickness while others have called this conclusion into question. However, none have directly investigated age and gender susceptibility to simulator-induced sickness in conjunction with display type in a driving simulator. Few studies have had the benefit of a driving simulator that allows the interchangeable use of both a direct-view display and an HMD. The DMV simulator, co-developed by Virginia Tech, Northeastern University, and the University of Virginia, will allow participants to perform the same driving tasks with both direct-view and head-mounted displays. The direct comparison of the two display technologies in a driving simulator task makes this research unique.

II. Procedures

Participants

Approximately 24 participants will be recruited for this study: 6 elderly males, 6 elderly females, 6 non-elderly males, and 6 non-elderly females. Participants are expected to encompass a wide range of driving abilities, habits, and preferences. All participants will be over 18 years of age and will be required to hold a valid driver's license. Elderly participants will range from 55 to 80 years in age. All participants will be compensated for their time and effort and will be recruited by flyers posted around the Virginia Tech Campus, by word of mouth, and (possibly) by postings in the Virginia Tech / Blacksburg USENET newsgroups. Participants will be compensated for their time at $8 per hour.

Prior to actual selection, prospective participants must undergo and successfully pass a screening process. Participants will be screened to ensure that they are in a normal state of health, have no history of epilepsy or seizures, have normal color vision, have not been exposed to a virtual environment in the last 48 hours, and have not taken any drugs (decided on case-by-case basis) in the last 24 hours. Screening shall be accomplished through the use of a participant information form, and standard static visual acuity and colorblindness tests. Participants will be required to have a minimum visual acuity of 20/30 (corrected or uncorrected). Once the prospective participant passes the screening process, he or she will be considered an active participant in the experiment. Participants that do not pass the screening process will be compensated for their time. No other exclusions will be used in selecting participants. Before the screening begins, volunteers will be asked to read the informed consent form, which includes a general description of the experiment. After the experimenter addresses any questions the participant may have, he or she will be asked to sign the informed consent form and the screening tests will begin.

Apparatus

The driving simulator that will be used in this study will be located on the fifth floor of Whittemore Hall on the Blacksburg Virginia Tech campus. It consists of a modified, fixed-base Advanced Therapy Products, Inc. (ATP) Interactive Driving Simulator Console with force-feedback steering and an acoustic seat for the simulation of vehicle frequency and bounce dynamics. To simulate the realism and functionality of a real car, the console is equipped with gas and brake pedals for lower extremity activity and simulates the real force and feel of car pedals. Two different displays will be used with the simulator during different trials. The first is a commercial NEC PlasmaSync 50MP1 50” (diagonal) flat panel monitor. The display can support a multitude of resolution levels on different computer platforms. The second is a commercially-available Head Mounted Display; the IO Systems i-glasses SVGA 3-D.

The simulator is controlled by an IBM-compatible personal computer running simulation software developed at Northeastern University in Boston, Massachusetts. The simulated driving environment, or driving course, is a direct representation of the course used by the DMV office in Charlottesville, Virginia. As the simulator software and hardware are still under development, details are subject to change.

Procedure

Pilot-testing. A series of pilot tests will be conducted using volunteers from the graduate student community at Virginia Tech. The pilot tests will involve setting up the hardware and software, running the driving scenarios with the various conditions applied, and timing the trials and setup between each trial. This will allow the experimenter to get a better idea of what will be involved in running actual participants, ensure data is being gathered and stored properly, and will allow for problems to present themselves so they may be addressed before the actual experiments begin. Pretesting will also ensure that: A) simulator-induced sickness symptoms can be elicited from the apparatus, and B) the driving scenario can be comfortably performed with both the direct-view display and the HMD. Volunteers for the pilot tests will also be asked to read and sign the informed consent form.

Experimental procedure. Participants will be randomly assigned to begin with one of the two treatment conditions. The general procedure will consist of two sessions, each dedicated to the use of one of the two display types. The two experimental sessions will be separated by at least one week.

Session 1. Prior to entering the simulator, participants will be asked to complete the paper-based Simulator Sickness Questionnaire (SSQ) and then be administered a test of postural disequilibrium (PD). The PD tests will have the participant stand with arms crossed and eyes closed and one leg behind the other – heel-to-toe. PD will be assessed by the time it takes a participant to move a foot. The Investigator or an assistant will stand close to the participant in the event the participant loses his or her balance.

Following the SSQ, the participant will be assisted into the simulator and the experimenter will explain the functions of the controls and how the experiment will proceed. In accordance with other driving simulator research performed at Virginia Tech, the participant will operate the simulator for three minutes in the experimenter's presence to "get a feel" for vehicle handling and simulator operation. Participants will then be told to relax as the experimenter initializes the recording equipment and room illumination will be reduced. Participants will be given a short time to rest then will be notified that the trial is about to start. During the trial, the Investigator will orally present instructions for the driving task as the participant operates the vehicle. If, for any reason the subject wishes to stop or an emergency develops, the simulation will be halted and the appropriate action will be taken. If the participant shows signs of discomfort, ataxia, etc., he or she will be monitored until it seems reasonable to let the participant leave, or he or she will be offered medical assistance if the experimenter and/or participant believes it to be necessary.

Upon completion of the driving task, the participant will be assisted from the simulator. The postural disequilibrium test will be immediately administered and then the participant will be asked to complete a second SSQ. Following that, if requested by the participant, he or she will be debriefed on the experiment and asked not to discuss or disclose any aspect of the experiment with anyone else for a period of 6 months.

Participants will be required to stay in the laboratory for 30 minutes following the simulation exposure. During this time, they will be asked to indicate whether they are feeling any simulator-induced sickness symptoms and if those symptoms are dissipating. At the end of the 30 minutes, the postural disequilibrium test will again be administered. Upon determination of the participant's well being, they will be paid, allowed to leave, and told to contact the Investigator if they experience a recurrence of symptoms. If a participant continues to exhibit symptomology, he or she will be monitored until symptoms abate, or will be offered medical assistance if the experimenter and/or participant believes it to be necessary. The experimenter will contact the participant the following day and ask if any other aftereffects were noted that may be attributed to the experiment. Any additional comments/concerns by the participants will also be addressed at that time.

Session 2. The procedures for session 2 will be nearly identical to those of the first session, minus the screening. A slightly different participant information questionnaire will be provided and participants will be asked to provide any indications of symptoms that they may still be experiencing from the first session. The rest of the procedures remain the same. The data from each participant's driving scenario will be collected by the simulation software and written to a local file on the simulator computer for later analysis. An analysis of variance (ANOVA) will be used to determine statistical significance between the factors being studied.

III. Risks and Benefits

Potential Risks

The plasma display and HMD used by the simulator are commercially available. Participants in this experiment are at no greater risk than if they play an immersive video game either at an arcade or in the home. The likelihood and seriousness of potential risks associated with this experiment are based on previous research in virtual environments and simulators. Visual simulator side-effects such as headaches, eyestrain, nausea, sweating, drowsiness, and dizziness may occur in some participants during the experiment or could develop within 24 hours after the experiment. Previous vehicular simulator studies typically find symptoms occurring in roughly half of the participants. Symptoms are almost always minor and short-lived, and usually persist no longer than 24 hours. However, in very rare cases, intense nausea or disorientation can be experienced during or shortly after the experiment, again usually lasting no longer than 24 hours. If the effects of simulator-induced sickness become too overwhelming for a participant to continue, the session will be halted and the participant will be allowed to rest and continue later, or quit the experiment altogether.

Minimizing Risks

The literature shows little evidence to support a clear method of preventing or adequately predicting simulator-induced sickness, however, the severity of simulator-induced sickness effects and post-simulation ataxia is often a function of the duration of the simulation. To minimize the possibility of severe instances of simulator-induced sickness and ataxia, the driving task will be kept short – no longer than 15 to 30 minutes. Participants will also be required to remain at the lab for 30 minutes following the simulation exposure so that symptoms can be monitored. If symptoms persist, participants can wait at the lab until they dissipate, receive a ride home from the experimenter, or, if necessary, the experimenter will arrange for medical assistance. Participants exhibiting symptoms will not be allowed to drive themselves home. There will be a one week rest period between sessions to allow for any latent effects to dissipate before continuing.

Benefits

The proposed research may help engineers and researchers build more effective (realistic) low-cost driving simulators for use in studying the effects of "humans in the loop" in virtual environments. This research should provide a basis of comparison between HMD and direct-view displays and suggest appropriate situations and users of each. In the immediate, this research will benefit the Virginia Department of Motor Vehicles by suggesting who may or may not be eligible for skills testing in a driving simulator. The benefits to the participants include a better understanding of driving simulators, experience with VR systems and the essential components necessary to create a realistic virtual driving environment, and insight into how experimental research is designed and conducted.

IV. Informed Consent

See attached consent form.

V. Confidentiality/Anonymity

Individual responses from this study will be kept strictly confidential. Personal information will be collected only for screening purposes and will not be disclosed without subject approval. Each participant will be assigned a number that will identify that participant through to the conclusion of the research. The name/number key will be kept locked in either the laboratory or the graduate student office and will be destroyed following the experiment. Participant data will be combined with information from other people taking part in the study and individual responses will not be revealed to anyone except members of the research team without the participant's written consent.

VIRGINIA TECH
GRADO DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING

Informed Consent for Participants of Investigative Projects

Research Title: Effects of display type, age, and gender on driving performance and simulator-induced sickness in a medium-fidelity driving simulator.

Principle Investigators: Dr. John G. Casali, John Grado Professor, ISE Department Head
Dr. Gary S. Robinson, Senior Research Associate, ISE
Mr. William J. Penhallegon, Graduate Candidate, ISE

WHY WE'RE DOING THIS RESEARCH

We are doing this research to help engineers and researchers build more realistic, low-cost, virtual-reality based driving simulators that can be used by the majority of the general population. In the immediate, this research will benefit the Virginia Department of Motor Vehicles by suggesting who may or may not be eligible for driver license skills testing in a driving simulator.

HOW WE'RE DOING THIS RESEARCH

You will be operating a fixed-base driving simulator two different times: once while viewing a direct-view, wide-screen monitor and again while wearing a head-mounted display (HMD). We will first check your visual acuity and color vision. You will then fill out a simple questionnaire that asks how you are currently feeling. We will then test your postural stability by asking you to stand with one leg behind the other, heel-to-toe with your eyes closed and arms crossed. You will then operate the driving simulator for approximately 20 minutes, which will be similar to a driving-based arcade game. When you are finished, we will again test your postural stability and ask you to complete the questionnaire again. Sometimes people experience disorientation following simulator exposure – we will ask that you remain in the lab for 30 minutes so we can make sure that you are ok. We will then give you the postural stability test one more time and if you are feeling fine, you can leave.

You will be asked to come back a week later to do the experiment again. Procedures will be identical; however, you will be operating the simulator with the other display type. You will again be asked to remain on site for 30 minutes after you operate the simulator so that we can make sure that you are ok. We will call you the days following each session to follow-up and make sure that you are not experiencing any problems. Your participation in this study ends when we call you after the second session.

WHERE WE'RE DOING THE RESEARCH AND HOW LONG IT WILL TAKE

The experiment will be conducted at Virginia Tech's Environmental and Safety Engineering Laboratory located in room 539, Whittemore Hall. The experiment will consist of two sessions: 1) pre-experiment screening and first simulator operation, and 2) second simulator operation a week later. The first session will take approximately 2 hours to complete and the second will take approximately 1 hour and 30 minutes. We will work with you to choose the days that fit most conveniently within your schedule.

RISKS TO YOU

As mentioned before, there is a potential risk to you in terms of possible disorientation or nausea that can occur during the experiment or could develop within 24 hours after the experiment; however, these risks are no greater than those inherent in operating an immersive or virtual reality-based video game. Symptoms are similar to motion sickness and range from mild to moderate headache, dizziness, disorientation, equilibrium disruption, and slight nausea. Previous studies typically find that symptoms occur in roughly half of the participants. Symptoms are almost always minor, if they occur at all, and usually persist for no longer than 24 hours. As a person who has been in a boat for most of a day might feel as if the room is momentarily swaying, you may feel some residual effect of having been in the simulator. However, in very rare cases, intense nausea or disorientation can be experienced during or shortly after the experiment, again usually lasting no longer than 24 hours.

We are making every effort to minimize all possible risks but if the experiment becomes too overwhelming for you to continue, the session will be stopped and you will be allowed to rest and continue later, or quit the experiment altogether. If symptoms persist, you can wait at the lab until they dissipate, receive a ride home from the experimenter, or, if necessary, the experimenter will arrange for medical assistance. You will not be allowed to drive yourself home if you are showing signs of disorientation. If symptoms reoccur after you have left the session, you should lie down and rest until they subside. Also, please contact one of the experimenters as we are concerned with your well-being and are interested in knowing if you experience delayed reactions.

MAINTAINING YOUR CONFIDENTIALITY

We will make every effort to prevent people who are not on the research team from seeing your data. Any personal information collected for screening purposes will not be disclosed without your approval. We will not put your name on the questionnaires, but instead will assign your responses a code number as identification for any written reports. The name/number key will be kept locked in either the laboratory or the graduate student office and will be destroyed following the experiment. Your data will be combined with information from other people taking part in the study and individual responses will not be revealed to anyone without your written consent except to members of the research team.

PAYMENTS AND REWARDS

Although the experimenters make no promise of benefits beyond monetary compensation, following the experiment you may have a greater understanding of driving simulators, virtual environments, and how experimental research is designed and conducted. You will be paid $8.00 an hour for your time during both the screening process and the experiment. Any fraction of time less than an hour will be pro-rated on an $8.00 per hour basis.

FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time for any reason without penalty. Circumstances may arise that the experimenter will determine that you should not continue as a participant in the study (an illness, for example). There will be no negative consequences to you if you do not complete the study and you will still be paid for your time.

ASKING QUESTIONS AND RECEIVING MORE DETAILED INFORMATION

Please feel free to ask any questions that might come to your mind right now. Later, if you have any questions about this study you can e-mail William Penhallegon at wpenhall@vt.edu or call 231-9086 (office) or 552-3871 (home). You can also contact the faculty advisors: Dr. John Casali at 231-9081 and Dr. Gary Robinson at 231-2680. If you have any questions about your rights as a research volunteer or whom to contact in the event of a research-related injury, you may contact the Chair of the Institutional Review Board, Dr. David Moore, at 231-4991. This research has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects at Virginia Polytechnic Institute and State University and by the Grado Department of Industrial and Systems Engineering. We will give you a copy of this consent form to take with you.

YOUR PERMISSION

By signing this form, I am freely volunteering to participate in this research project. I have read the informed consent and fully understand the procedures and conditions of the experiment. I have had all my questions answered, and I hereby give my voluntary consent to be a participant in this research study. I agree to abide by the rules of the experiment and my responsibilities are: remain (illegal) drug and alcohol free for 24 hours before the experiment, operate a driving simulator, and notify the experimenters if I begin to feel ill in any way. I understand that I may withdraw from the study at any time, without consequence.

_______________________________________ ________________

Signature of person agreeing to take part in the study Date

____

Printed name of person taking part in the study

_______________________________________ ________________

Signature of person obtaining informed consent Date

Appendix E

Session Chronological Procedures

Session 1 Procedure

1. Review Informed Consent and have participant sign

2. Distribute participant information questionnaire. If HMD is used, take this time to calibrate.

3. Administer pre-exposure SSQ.

4. Vision tests:
   - Visual acuity (far, near)
   - Color Discrimination

5. Administer pre-exposure TLHTT postural disequilibrium test

6. Seat participants in simulator. Explain controls, driving scenario, procedures.

7. If Plasma display is used, adjust display distance.

8. Participant drives one half of the loop to get “feel” of simulator.

9. Participant gets two minutes to relax as simulation loads; then is informed that trial will begin.

10. Trial begins. Participant operates vehicle for one lap and is instructed to maintain a centered right-lane position and to obey posted speed limits. Investigator verbally provides specific instruction at decision points (turns, passing). Participants are instructed to pass two vehicles on the interstate scenario.

11. Trial ends. Participant is assisted from simulator.

12. Post-exposure TLHTT postural disequilibrium test is again administered. Start 30-minute timer.

13. Post-exposure SSQ administered

14. Participant is kept for 30 minutes following evaluation.
    - Solicit subjective observations of the display system.
    - Schedule second session.
    - Final TLHTT postural stability test is administered at the end of the 30-minute period.
    - Participant is allowed to leave.

15. Next day follow-up call.
    - Ask if any aftereffects were noted that may be attributed to the experiment.
    - Solicit any additional comments/concerns by the participants.
    - Confirm second session.

Session 2 Procedure

1. Distribute participant information questionnaire. If HMD is used, take this time to calibrate.

2. Administer pre-exposure SSQ.

3. Administer pre-exposure TLHTT postural disequilibrium test.

4. Seat participants in simulator. Explain controls, driving scenario, procedures.

5. If Plasma display is used, adjust display distance.

6. Participant drives one half loop to get “feel” of simulator.

7. Participant gets two minutes to relax as simulation loads; then informed that trial will begin.

8. Trial begins. Participant operates vehicle for one lap and is instructed to maintain a centered right-lane position and to obey posted speed limits. Investigator verbally provides specific instruction at decision points (turns, passing). Participants are instructed to pass two vehicles on the interstate scenario.

9. Trial ends. Participant is assisted from simulator.

10. Post-exposure TLHTT postural disequilibrium test is again administered. Start 30-minute timer.

11. Post-exposure SSQ administered.

12. Participant is kept for 30 minutes following evaluation.
    - Solicit subjective observations of the display system.
    - Schedule second session.
    - Final TLHTT postural stability test will be administered at the end of the 30-minute period.
    - Participant is paid and allowed to leave.

13. Next day follow-up call.
    - Ask if any aftereffects were noted that may be attributed to the experiment.
    - Solicit any additional comments/concerns by the participants.

VITA

William J. Penhallegon was born in 1974 in Towson, Maryland. He received his Bachelor of Science degree in Industrial and Systems Engineering from Virginia Polytechnic Institute and State University in December 1996. Upon completion of the program, he began working for the Naval Air Systems Command (NAVAIR) at Patuxent River, Maryland in the Human Systems Research and Engineering Division. His primary responsibilities include human factors consulting during weapons system design phases and human factors test and evaluation support. He has a selected passenger flight qualification and occasionally serves in a flight test engineering capacity. He is also experienced in developing cost analysis models for avionics programs.

In the Summer of 2003, Mr. Penhallegon completed a Master of Science Degree program in Industrial and Systems Engineering (Human Factors Engineering Option) at Virginia Tech. As a graduate student he was a recipient of the UPS Fellowship. His interests include simulation, situational awareness, aviation and display technologies, and multimedia development. He is also a member of the Human Factors and Ergonomics Society.

Mr. Penhallegon returned to his position at the Patuxent River Naval Air Test Center in the Fall of 2002.