Usability Studies with Virtual and Traditional Computer Aided Design Environments

Syed Adeel Ahmed
University of New Orleans

Follow this and additional works at: https://scholarworks.uno.edu/td

Recommended Citation
https://scholarworks.uno.edu/td/477

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UNO. It has been accepted for inclusion in University of New Orleans Theses and Dissertations by an authorized administrator of ScholarWorks@UNO. The author is solely responsible for ensuring compliance with copyright. For more information, please contact scholarworks@uno.edu.
Usability Studies with Virtual and Traditional Computer Aided Design Environments

A Dissertation

Submitted to Graduate faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Engineering and Applied Science

by
Syed Adeel Ahmed
B.S. Osmania University, 1997
M.S. University of New Orleans, 2001

December 2006
Dedication

This dissertation is dedicated to Almighty God, my mother (Shah Zamani Begum), my father (Syed Shakeel Ahmed), my sister (Dr. Asra Tabassum & her kids), my wife (Naveedunnisa Ahmed), my kids (Zoya, Neha, Mukhtadeer/Sohaib Ahmed), my teachers, all my well wishers/friends, and my entire family.

“The best I’ll ever have.”
Acknowledgements

I am most grateful to the Office of Naval Research (ONR) for their support of the research through GCRMTC (Dr. Alley Butler & Dr. Kurt Satter), Dillard University (Dr. Keith Wismar) and to the Naval Research Laboratory (NRL) for the VR equipment provided to the University of New Orleans through Department of Defense (DoD) Property Disposal.

Several University of New Orleans staff and graduate assistants must be credited with much of the CAVE™ systems installation, technical support, and the seemingly endless task of test administration. These exceedingly helpful individuals include: Edward Williams, Steve Attaya, Arthur Butz, Dr. Greg Dobson, and Mathew Hamey.

I must also recognize all my Dillard University students who participated as test subjects in this research. Their assistance was invaluable and sincerely appreciated.

I thank Dr. Kurt Satter, Dr. Alley Butler and their whole families for their precious support. Without their help and encouragement little, if any, progress would have been made.

I also thank the Department Chair, William Lannes, for championing the PhD program and for his support and direction into the administrivia of terminal degree pursuit. Without his guidance and support little progress would have been made.

Last but not least, I acknowledge the expertise and efforts of all my committee members Dr. Donald Barbe, Dr. Jay Hunt, Dr. Zhide Fang, Dr. Charalampidis Dimitrios and Dr. Kurt Satter for their endless encouragement and their mantra of: “You can do it Syed!”
Table of Contents

LIST OF TABLES ........................................................................................................... V

LIST OF FIGURES ........................................................................................................ VIII

ABSTRACT .................................................................................................................... XI

KEYWORDS .................................................................................................................... XI

1. INTRODUCTION ............................................................................................................ 1

   1.1 Background ............................................................................................................... 1
   1.2 VE Interaction: Technology ................................................................................... 6
      1.2.1 Inputs .............................................................................................................. 6
      1.2.2 Output .......................................................................................................... 9
      1.2.3 Immersive Virtual Environments and Their Characteristics ....................... 10
   1.3 Interactions in a Virtual Environment (VE) ........................................................ 11
   1.4 Interaction Devices for Virtual Reality .............................................................. 13
      1.4.1 Head-Mounted Display (HMD) .................................................................... 13
      1.4.2 CAVE™ .................................................................................................... 14
      1.4.3 Data Gloves ............................................................................................... 15
      1.4.4 Pinch Gloves ............................................................................................. 15
      1.4.5 Joystick, Wand, or Flightstick ..................................................................... 16
      1.4.6 Shared Virtual Environments ..................................................................... 17
      1.4.7 ImmersaDesk ............................................................................................ 18
      1.4.8 CrystalEyes ............................................................................................... 20

2. GESTURAL INTERFACE ............................................................................................. 22

   2.1 What is a Gesture .................................................................................................... 22
   2.2 How to use gestures? ........................................................................................... 25
   2.3 Usability and Ease of Use of Gestures ................................................................. 27
   2.4 The Role of Non-Symbolic Gestures .................................................................. 30
   2.5 Task-Based Evaluation of Gesture Interactions ................................................. 31
   2.6 Gesture in Selection ............................................................................................. 32
   2.7 Gesture in Manipulation ...................................................................................... 34
   2.8 Gesture in Travel ................................................................................................. 35
   2.9 Gesture in Traditional User Interfaces .............................................................. 38
   2.10 Guidelines for the Design of Gesture Command Sets ....................................... 40
   2.11 Objective of the Research .................................................................................. 43
3. PROJECT DESCRIPTION .................................................................................. 44

3.1 Background .................................................................................................. 44
3.2 User Centered Development ........................................................................ 49
3.3 Gestural, Voice, and Multimodal Virtual Interfaces ........................................ 52
3.4 General Experimental Procedure .................................................................. 55

4 BENCHMARK 1 (NAVIGATION) ...................................................................... 63

4.1 Description .................................................................................................... 63
4.2 Pass-to-pass Improvements in Elapsed Times ................................................. 63
4.3 B1-Pass-to-Pass Comparison of Elapsed Times Analysis ................................. 64
4.4 Elapsed Times Detailed Statistical Analysis ................................................... 65
4.5 Mann-Whitney Test ....................................................................................... 67
4.6 Pass 3 Statistics ............................................................................................ 67
4.7 User Subjective Overall Environment Ratings ................................................ 68
4.8 B1-Pass to Pass Comparison of Overall Impressions Ratings Analysis .......... 71

5 BENCHMARK 2 (FIND AND REPAIR MANIPULATION) ............................ 72

5.1 Description .................................................................................................... 72
5.2 Benchmark 2 – pass 3 Elapsed timing analysis ............................................. 73
5.3 Pass-to-Pass Comparison of Elapsed Times Analysis ..................................... 74
5.4 Detailed Statistical Analysis .......................................................................... 74
5.5 Pass 3 Statistical Analysis ............................................................................. 74
5.6 User Subjective Overall Environment Ratings .............................................. 76
5.7 Benchmark 2 – pass 3 Overall Impressions ratings analysis ........................... 76
5.8 B2-Pass to Pass Comparison Overall Impressions Ratings Analysis ............... 78

6. BENCHMARK 3 (SPATIAL AWARENESS) ............................................... 79

6.1 Description .................................................................................................... 79
6.2 Benchmark 3, Pass 3, Part 1 & 2 Placement Offsets Analysis ......................... 81
6.3 Detailed Statistical Analysis .......................................................................... 84
6.4 Benchmark 3 Pass 3 Statistics ................................................................. 84
6.5 B3p3-2off –Benchmark 3 Pass 3 Descriptive Statistics .................................. 84
6.6 Benchmark 3 pass 3 Overall Impressions Ratings Analysis .......................... 86
6.7 Detailed Statistical Analysis .......................................................................... 87
6.8 Benchmark 3 Pass 3 Overall Impressions Ratings Statistics .......................... 87
6.9 Benchmark 3 Pass-to-Pass Comparison Analysis ........................................ 89
7. BENCHMARK 4 (FAULT IDENTIFICATION)................................. 91

7.1 Description ................................................................................................................... 91
7.2 Benchmark 4, Pass 3, faults count Analysis ................................................................. 92
7.3 B4p3- Benchmark 4 Pass 3 Descriptive Statistics ......................................................... 94
7.4 Benchmark 3 pass 3 Overall Impressions Ratings Analysis ......................................... 94
7.5 Detailed Statistical Analysis ....................................................................................... 95
7.6 Benchmark 4 Pass 3 Overall Impressions Ratings Statistics ....................................... 95

8. CONCLUSIONS AND RECOMMENDATIONS ........................................ 98

8.1 Enhancements for Further Study ................................................................................. 99
8.2 Tracking Map ............................................................................................................. 100
8.3 Future Work ............................................................................................................. 100

References ...................................................................................................................... 103

Appendix A ....................................................................................................................... 116
Appendix B ....................................................................................................................... 141
Appendix C ....................................................................................................................... 166
Appendix D ....................................................................................................................... 195
Appendix 1 – Course Completion Certificates ................................................................. 220
Appendix 2 – Consent Form ............................................................................................ 223
Appendix 3 – Approval for Human Subjects ................................................................. 228
Appendix 4 – Survey used by Satter (2005) .................................................................. 230
Appendix 5 – Medical Condition Questionnaire ............................................................ 232
Appendix 6 – Dillard University Flyer ............................................................................ 234
Vita ................................................................................................................................. 236
List of Tables

Table 4.1: B1-Pass-to-Pass Comparison of Elapsed Times .........................................................65
Table 4.2: B1p3Tim Elapsed Times ..................................................................................................67
Table 4.3: B1p3Ovr Overall Impressions Ratings .........................................................................70
Table 4.4: B1-Pass to Pass Comparison of Overall Impressions Ratings .......................................71
Table 5.1: Pass-to-Pass Comparison of Elapsed Times ..................................................................74
Table 5.2: B2p3Tim Elapsed Times Statistics ...................................................................................75
Table 5.3: B2p3Ovr Overall Impressions Ratings ...........................................................................77
Table 5.4: B2-Pass to Pass Comparison Overall Impressions Ratings ...........................................78
Table 6.1: B3p3-1off Pass 3-Icon 1 Offsets ......................................................................................82
Table 6.2: B3p3-2off Pass 3-Icon 2 Offsets .....................................................................................85
Table 6.3: B3p3Ovr Pass 3 Overall Impressions Ratings Descriptive Statistics ................................88
Table 6.4: B3I1 pass-to-pass Comparison of Offset distances ............................................................89
Table 6.5: B3I2 pass-to-pass Comparison of Offset distances ...........................................................89
Table 6.6: B3 Overall Impressions Ratings pass to pass Comparison .............................................90
Table 7.1: B4p3 Faults Count Statistics ............................................................................................93
Table 7.2: B4p3Ovr Overall Impressions Ratings Statistics .............................................................96
Table 7.3: B4 Pass-to-Pass Comparison of Faults Count .................................................................96
Table 7.4: B4 Overall Impressions Ratings Pass to Pass Comparison .............................................97
Table A-1: B1p1Tim Elapsed Times Statistics ..................................................................................117
Table A-2: B1p1Nav Navigation Ratings Statistics ........................................................................118
Table A-3: B1p1Loc Locating Ratings Statistics .............................................................................119
Table A-4: B1p1Mov Movement Ratings Statistics ........................................................................120
Table A-5: B1p1Gen General Impressions Ratings Statistics .........................................................121
Table A-6: B1p1Ovr Overall Impressions Ratings Statistics ..........................................................122
Table A-7: B1p2Tim Elapsed Times Statistics ..................................................................................123
Table A-8: B1p2Nav Navigation Ratings Statistics ........................................................................124
Table A-9: B1p2Loc Locating Ratings Statistics ..............................................................................125
Table A-10: B1p2Mov Movement Ratings Statistics ......................................................................126
Table A-11: B1p2Gen General Impressions Ratings Statistics .......................................................127
Table A-12: B1p2Ovr Overall Impressions Ratings Statistics .........................................................128
Table A-13: B1p3Tim Elapsed Times Statistics ................................................................................129
Table A-14: B1p3Nav Navigation Ratings Statistics .......................................................................130
Table A-15: B1p3Loc Locating Ratings Statistics ..........................................................................131
Table A-16: B1p3Mov Movement Ratings Statistics ......................................................................132
Table A-17: B1p3Gen General Impressions Ratings Statistics .......................................................133
Table A-18: B1p3Ovr Overall Impressions Ratings Statistics ..........................................................134
Table A-19: B1-3pAvgTim Elapsed Times Statistics ........................................................................135
Table A-20: B1-3pAvgNav Navigation Ratings Statistics .................................................................136
Table A-21: B1-3pAvgLoc Locating Ratings Statistics ...................................................................137
Table A-22: B1-3pAvgMov Movement Ratings Statistics ...............................................................138
Table A-23: B1-3pAvgGen General Impressions Ratings Statistics ...............................................139
Table A-24: B1-3pAvgOvr Overall Impressions Ratings Statistics ..................................................140
Table B-1: B2p1Tim Elapsed Times Statistics ...............................................................................142
Table B-2: B2p1Nav Navigation Ratings Statistics .......................................................................143
List of Figures

Figure 1.1: A Head Mounted Display (HMD).................................................................14
Figure 1.2: Schematic of the CAVE™ System ................................................................14
Figure 1.3: Data Gloves .................................................................................................15
Figure 1.4: Application of Data Gloves ...........................................................................15
Figure 1.5: Pinch Glove System ......................................................................................16
Figure 1.6: Flightstick ...................................................................................................17
Figure 1.7: Networked Virtual Environments .................................................................18
Figure 1.8: The Immersadesk System ............................................................................19
Figure 1.9: Crystal Eye Glasses .......................................................................................21
Figure 3.1: Synthesis of a Design .................................................................................47
Figure 3.2: Completed Design .......................................................................................47
Figure 3.3: Wand Gestural Interface with the Adaptable .................................................48
Figure 3.4: Voice & Glove Gestural Interface with the Adaptable ..................................48
Figure 3.5: Designer Centered Design and Evaluation Process ....................................51
Figure 4.1: B1p3Tim Elapsed Times .............................................................................64
Figure 4.2: B1p3Ovr Overall Impressions Ratings ......................................................69
Figure 5.1: B2p3Tim Elapsed Times .............................................................................73
Figure 5.2: B2p3Ovr Overall Impressions Ratings ......................................................77
Figure 6.0: Icon .............................................................................................................80
Figure 6.1: B3p3-1off Pass 3-Icon 1 Offsets .................................................................82
Figure 6.2: B3p3-2off Pass 3-Icon 2 Offsets .................................................................83
Figure 6.3: B3p3Ovr Pass 3 Overall Impressions Ratings .............................................86
Figure 7.1: B4p3 Faults Count ......................................................................................93
Figure 7.2: B4p3Ovr Overall Impressions Ratings ......................................................96
Figure A-1: B1p1Tim Elapsed Times .............................................................................117
Figure A-2: B1p1Nav Navigation Ratings ....................................................................118
Figure A-3: B1p1Loc Locating Ratings .........................................................................119
Figure A-4: B1p1Mov Movement Ratings ...................................................................120
Figure A-5: B1p1Gen General Impressions Ratings ....................................................121
Figure A-6: B1p1Ovr Overall Impressions Ratings ......................................................122
Figure A-7: B1p2Tim Elapsed Times .............................................................................123
Figure A-8: B1p2Nav Navigation Ratings ....................................................................124
Figure A-9: B1p2Loc Locating Ratings .........................................................................125
Figure A-10: B1p2Mov Movement Ratings ..................................................................126
Figure A-11: B1p2Gen General Impressions Ratings ...................................................127
Figure A-12: B1p2Ovr Overall Impressions Ratings .....................................................128
Figure A-13: B1p3Tim Elapsed Times ...........................................................................129
Figure A-14: B1p3Nav Navigation Ratings ...................................................................130
Figure A-15: B1p3Loc Locating Ratings ......................................................................131
Figure A-16: B1p3Mov Movement Ratings ..................................................................132
Figure A-17: B1p3Gen General Impressions Ratings ...................................................133
Figure A-18: B1p3Ovr Overall Impressions Ratings .....................................................134
Figure A-19: B1-3pAvgTim Elapsed Times ..................................................................135
Figure A-20: B1-3pAvgNav Navigation Ratings ...........................................................136
Figure C-19: B3p3Mov Pass 3 Movement Ratings .........................................................185
Figure C-20: B3p3Gen Pass 3 General Impressions Ratings .......................................186
Figure C-21: B3p3Ovr Pass 3 Overall Impressions Ratings ..........................................187
Figure C-22: B3-3pA-1Off 3 Pass Avg. Part 1 Offsets .................................................188
Figure C-23: B3-3pA-1Off 3 Pass Avg. Part 2 Offsets ...................................................189
Figure C-24: B3-3pAnav 3 Pass Avg. Navigation Ratings ..............................................190
Figure C-25: B3-3pALoc 3 Pass Avg. Locating Ratings ................................................191
Figure C-26: B3-3pAmov 3 Pass Avg. Movement Ratings ............................................192
Figure C-27: B3-3pAgen 3 Pass Avg. General Impressions Ratings ..............................193
Figure C-28: B3-3pAOvr 3 Pass Avg. Overall Impressions Ratings ..............................194
Figure D-1: B4p1 Faults Count ....................................................................................196
Figure D-2: B4p1Nav Navigation Ratings ....................................................................197
Figure D-3: B4p1Loc Locating Ratings .......................................................................198
Figure D-4: B4p1Mov Movement Ratings ...................................................................199
Figure D-5: B4p1Gen General Impressions Ratings ....................................................200
Figure D-6: B4p1Ovr Overall Impressions Ratings .....................................................201
Figure D-7: B4p2 Faults Count ....................................................................................202
Figure D-8: B4p2Nav Navigation Ratings ...................................................................203
Figure D-9: B4p2Loc Locating Ratings .......................................................................204
Figure D-10: B4p2Mov Movement Ratings ................................................................205
Figure D-11: B4p2Gen General Impression Ratings ...................................................206
Figure D-12: B4p2Ovr Overall Impression Ratings .....................................................207
Figure D-13: B4p3 Faults Count ....................................................................................208
Figure D-14: B4p3Nav Navigation Ratings .................................................................209
Figure D-15: B4p3Loc Locating Ratings .....................................................................210
Figure D-16: B4p3Mov Movement Ratings ................................................................211
Figure D-17: B4p3Gen General Impressions Ratings ....................................................212
Figure D-18: B4p3Ovr Overall Impressions Ratings ....................................................213
Figure D-19: B4-3pAvg Faults Count ............................................................................214
Figure D-20: B4-3pAvgNav Navigation Ratings ............................................................215
Figure D-21: B4-3pAvgLoc Locating Ratings ...............................................................216
Figure D-22: B4-3pAvgMov Movement Ratings ..........................................................217
Figure D-23: B4-3pAvgGen General Impressions Ratings ..........................................218
Figure D-24: B4-3pAvgOvr Overall Impressions Ratings .............................................219
Abstract

For both the CAVE™ and the adaptable technology possessed by the University of New Orleans, crystal eye glasses are used to produce a stereoscopic view, and an ascension flock of birds tracking system is employed for tracking of the user’s head position and position of a wand in 3D space.

It is argued that with these immersive technologies along the use of gestures and hand movements should provide a more natural interface with the immersive virtual environment. This allows a more rapid and efficient set of actions to recognize geometry, interaction with a spatial environment, the ability to find errors, or navigate through an environment. The wand interface is used to provide an improved means of interaction. This study quantitatively measures the differences in interaction when compared with traditional human computer interfaces.

This work uses competitive usability in four different Benchmarks: 1) navigation, 2) error detection/correction, 3) spatial awareness, and 4) a “shopping list” of error identifications. This work expands on [Butler & Satter’s, 2005] work by conducting tests in the CAVE™ system, rather than principally employing workbench technology. During testing, the testers are given some time to “play around” with the CAVE™ environment for familiarity before undertaking a specific exercise. The testers are then instructed regarding tasks to be completed, and are asked to work quickly without sacrificing accuracy. The research team timed each task, counted errors, and recorded activity on evaluation sheets for each Benchmark test. At the completion of the testing scenarios involving Benchmarks 1, 2, 3, or 4, the subjects were given a survey document and asked to respond by checking boxes to communicate their subjective opinions.

Keywords

Usability Analysis, Perceptual User Interface, CAVE™ (Cave Automatic Virtual Environments).
1. INTRODUCTION

1.1 Background

The term 'Virtual Reality' (VR) was initially coined by Jaron Lanier, founder of VPL Research. Other related terms include 'Artificial Reality' [Myron Krueger, 1970], 'Cyberspace' [William Gibson, 1984], and, more recently, 'Virtual Worlds' and 'Virtual Environments'. Today, 'Virtual Reality' is used in a variety of ways and often in a confusing and misleading manner. Originally, the term referred to 'Immersive VR.' In immersive VR, the user becomes fully immersed in an artificial, three-dimensional world that is completely generated by a computer. VR is a form of human-computer interface characterized by environmental simulation generated using computer systems. VR requires hardware and software that furnish a sense of (1) immersion, (2) navigation, and (3) manipulation [Helsel, 1992]. VR falls into three major categories: text-based, desktop, and immersive VR. Text-based networked VR involves real-time environments described textually on the Internet where people interact by typing commands and "speak" by typing messages on their computer keyboards. This form of VR has been valuable in distance education [Psotka, 1994]. Desktop VR is an extension of interactive multimedia involving three-dimensional images and adds to the experience of interactive multimedia without being considered immersive. Immersive VR involves a mixture of hardware, software and concepts that allow the user to interact with a three dimensional computer generated "world" [Loeffler & Anderson, 1994].

The specific hardware that currently enables immersive VR includes:

(1) Head Mounted Displays (HMD or 'eyephones') which provide 3D vision of 200 degrees horizontally and 120 degrees vertically [Winn, 1993];
(2) Datagloves which allow the user to interact with the environment by tracking the user’s motion and giving tactile reinforcement to the visual stimuli in the simulated world; and
(3) Wands or other devices which allow the user to manipulate objects in the virtual world.

The major software required for VR includes high resolution image generators which allow real time rendering so the virtual world is updated as the user acts upon it; and software which allows localized stereo sound and in some cases smell and voice recognition [Psotka, 1994].

In addition, Hedberg & Alexander [1994] include sensory and psychological immersion and active learner participation as defining educational factors of VR. Winn [1995] describes the result of VR's mixture of hardware, software and concepts as a phenomenon known as "cognitive presence" involving a "conviction that the virtual world is a valid, though a different form of reality." This phenomenon has been compared to the "suspension of disbelief" humans experience while watching a play or movie, but appears to involve less effort on the part of the audience or user, with far more convincing effects.

The last decade saw the use of computers in almost every field of human activity. One of the main reasons for this was the introduction of human-friendly interfaces that have made computers easy to use and learn. The most successful interface paradigm so far was the Xerox Parc Desktop metaphor. However, the desktop metaphor was best suited to interact with 2D (two dimensional) worlds, but it showed limitations when interacting with 3D (three dimensional) worlds. Recently, researchers in the field of human computer interaction (HCI) focused on this problem and have gradually made possible the development of new input devices and displays for interacting with remote or computer generated worlds.

These new input devices and displays use new paradigms for HCI. Instead of keyboard input, interaction is based on voice, gesture and hand manipulation. Displays are created to closely
match human vision capabilities including development of head mounted display (HMD) concepts. The aim is to simulate operator presence in a computer synthesized world. The aim of virtual reality (VR) systems is to immerse the participant within a computer-generated, virtual environment (VE). Interacting with the VE poses issues unique to VR. With ideal VE system, the participant fully believes he/she is actually performing a task in a “like real” environment. Every component of the task is fully replicated. The VE is visually identical to the real task, but occurs in an artificial created environment. Further, in the ideal virtual environment the participant hears appropriate sounds, smell identical odors, and when they reached out to touch an object, they are able to feel it. For example, in a VR system to examine designs for product assembly, the ideal system would present an experience identical to actually performing the assembly task. Parts and tools would have mass, feel real, and move appropriately with the laws of physics. The participant could interact with every object as if he/she were doing a specific task. The virtual objects would in turn respond to the participant’s action appropriately for the simulated conditions. The result is known as immersion where the user does not notice that the virtual environment is a computer generated simulation. Obviously, current virtual environments are only beginning to approach an ideal immersive system. Participants use specialized equipment, such as tracked displays and gloves, to track movement, interpret actions, and provide input to the VR system. Interactive 3D computer graphics and audio software can generate the appropriate scenes and auditory cues. Finally, the participant receives the VE output (e.g. images, sounds, haptic (tactile) feedback) through visual and audio hardware. Although the term virtual reality is used in many contexts, in the context of this work virtual reality includes only stereoscopic systems with active tracking. Immersive VR is characterized –
though not universally – by participant head tracking (monitoring the participant’s position and orientation) and stereo imagery (providing different views for each eye).

Interestingly, VR human-computer interaction (HCI) issues can be strikingly different than traditional HCI. With the VE, the following important issues influence the HCI approach:

- *The participant views the VE from a first person perspective point of view.*
- *VR interaction strives for a high level of fidelity between the virtual action and the corresponding real action being simulated.* For example, a VR system for training soldiers in close quarters combat must have the participant perform physical actions, and receive visual, audio, and haptic input, similar to the actual ground combat scenario as possible.
- *Typically most – if not all – objects in the VE are virtual.*

Immersive VR systems that satisfy the high fidelity interactions requirements can become an important tool for training, simulation, and education for tasks that are dangerous, expensive, or infeasible to recreate. Examples of a near perfect combination of real and virtual objects are flight simulators. In most state-of-the-art flight simulators, the entire motion platform is real, but a motion platform provides motion sensations, and the visuals of the environment outside the motion platform are virtual. The resulting synergy is so compelling and effective it is almost universally used to train pilots.

Virtual environments have been described as the means to more intuitive and effective HCI for a variety of applications, including the analysis of complex scientific data, medical training, military simulation, phobia therapy, and virtual prototyping [Mine et al, 1997; Stanney et al, 1998]. Among the available interaction paradigms, techniques making use of gesture recognition have frequently been identified as ideal interfaces given their potential to be natural, efficient,
and intuitive means of communication. Since its first application to VR in the 1970’s [Krueger, 1977; Krueger 1983] gesture has been an active area of research. Although the applications of VE technology are numerous, limited research has been conducted that guides the selection and development of successful interaction techniques in virtual environments. Some guidelines for the use and integration of gesture interaction techniques that are motivated by cognitive, perceptual, and human factors research are presented here.

While Turk’s [Turk, 2002] review provides guidance for improving the usability of gestures, his paper does not focus on understanding individual performance requirements and intra-individual limitations of gestures. An examination of the cognitive, perceptual, and human factor motivations for the use of gestures in virtual environments is missing from the literature. Recent reviews have also neglected to examine the role of gestures in specific interaction tasks common to most virtual environments. Instead, reviews have most frequently focused on the variety of hardware solutions available for implementing gesture-based interaction. As a result, the virtual environment community lacks a framework for gesture interaction that could support the identification of appropriate and effective gestures and task-based applications.

Virtual reality (VR) is not a new concept even if the oxymoron "artificial reality" was once introduced by Krueger (1983). Sutherland (1965) introduced the key concepts of immersion in a simulated world, and of complete sensory input and output, which are the basis of current VR research. At MIT, at the beginning of the 1980’s, a limited 3D virtual workspace in which the user interactively manipulates 3D graphical objects spatially corresponding to hand position was developed (Schmandt 1983). In 1984, NASA started the VIVED project (Virtual Visual Environment Display) and later the VIEW project (Virtual Interactive Environment Workstation). As described in Fisher et al. (1986), the aim of the research at NASA is to develop
a multipurpose, multimodal operator interface to facilitate natural interaction with complex operational tasks and to augment operator awareness of large-scale autonomous integrated systems. The application areas of focus are telepresence control, supervision and management of large-scaled information systems, and human factors research. VPL and Autodesk introduced VR to the general public on June 6, 1989 at two trade shows. Both companies presented devices and HMDs for interacting with virtual worlds. Since then, VR has captured the public imagination and much work has been done to explore the possibilities of VR in new areas of application such as medicine, chemistry, and scientific visualization.

VR is more than just interacting with 3D worlds. By offering a realistic simulation to users as an interface metaphor, VR allows operators to perform tasks on remote real worlds, computer generated worlds or any combination of both. The simulated world does not necessarily have to obey natural laws of behavior. Such a statement makes nearly every area of human activity, a candidate for a VR application.

1.2 VE Interaction: Technology

Interaction involves the means for the user to communicate with the virtual environment. This involves providing input to the system and receiving output from the system. This allows the computer system to adjust the virtual environment in real time using a computer to maintain a sense of “immersion” in the virtual world through this input and output interaction. The result is an “immersive environment” with specific characteristics.

1.2.1 Inputs

Tracking and signaling actions are the key means of input into VEs. Tracking is defined as the determination of an object’s position and orientation. Common objects to track include the participant’s head, participant’s limbs, and interaction devices (such as gloves, wands). Most
tracking systems have sensors attached to the objects. Then, other devices track and report the position and orientation of these sensors.

Typically, tracking systems use one or a combination of mechanical, magnetic (Polhemus Fastrak and Ascension Flock of Birds), optical (WorldViz PPT, 3rdTech Hiball), acoustic (Logitech 6D Mouse), inertial (Intersense IS-900), and global position systems (GPS) approaches. Each method has different advantages with respect to cost, speed, accuracy, and size.

Different tasks have varying requirements on the accuracy, speed, and latency of the tracking system’s transmission of position to the computer. VEs that aim for a high level of participant sense of presence have strict head tracking requirements. Researchers estimate that the VR and tracking systems need to accurately determine the participant’s pose and to display the appropriate images in under 90 milliseconds (ms), and preferably under 50 ms. If the lag is too large, the VR system might make the participant disoriented and hamper the quality of interactivity.

Tracking the participant’s limbs allows the VR system to (1) present an avatar, a virtual representation of the user within the VE, and (2) rough shape information of the participant’s body pose. The presence of an avatar increases a participant’s sense of presence. The accuracy and speed requirements for limb tracking are typically lower than that of head tracking.

The popular conception of an avatar comes from science fiction novels about adventures in virtual worlds. One of the most popular of these was Neal Stephenson’s *Snow Crash*, published in 1992. The concept became a reality when the World Wide Web enabled fast transmission of graphical content over the Internet.
Finally object tracking, usually accomplished by attaching a sensor, allows a virtual model of an object to be registered with a physical real object. For example, attaching a tracker to a fork allows an associated virtual fork to be naturally manipulated. Since each sensor reports the pose information of a single point, most systems use one sensor per object and assume the real object is rigid in shape and appearance.

Since humans use hands for many interaction tasks, tracking and obtaining inputs from the hand is a natural evolution for a data glove. A tracked glove reports position and pose information of the participant’s hand to the VR system. Tracked gloves can also report pinching gestures (Fakespace Pinchglove), button presses (buttons built into the glove) and finger bends (Immersion CyberTouch). These glove actions are associated with virtual actions such as grasping, selecting, translation, and rotation. Tracked gloves provide different kinds of inputs and most importantly, are very natural to use. Glove disadvantages include sizing problems (most are a one size fits all), limited feedback (issues with haptic feedback and detecting gestures), and hygiene complications with multiple users.

The most common interaction devices are tracked mice (sometimes called bats) and joysticks. They are similar to a regular mouse and joystick, but with an integrated 3 or 6 degrees-of-freedom (DOF) tracking sensor that reports the device’s position and possibly orientation. Tracked mice and joysticks have many buttons for the participant to provide input, and they are cheap, and easily adaptable for different tasks. However, they might not provide the required feel for a given task or seem awkward and unnatural. So tracked mice and joysticks must be designed properly and used in appropriate applications.
1.2.2 Output

Given the system inputs, the resulting VE (visuals, audio, tactile information) is provided to the participant. For example, as the participant changes their head position and orientation, the tracking system passes that information to the VR system’s rendering engine. 3D stereoscopic rendering views for the VE are generated from the updated pose information.

The visual output is typically presented either in a head-mounted display (HMD) or a stereoscopic projected environment using shutter glasses. HMDs are head-worn helmets with integrated display devices. The helmet has two screens located a short distance from the participant’s eyes. HMDs can be thought of as the participant “carrying” around the display. The shutter glasses by comparison are lightweight, but require an infrared emitter to create the stereoscopic display.

The workbench environment provides one image for the right and left eye. The images are back projected onto a screen to avoid shadows. In contrast, CAVE™ environments have multiple back projected display walls and possibly direct projection onto the floor. With the CAVE™, the VE rendering is based on the viewer’s location, but projected onto each respective screen (such as forward, right, left, down).

VR systems can use either stereo headphones or multiple speakers to output audio. Given the participant’s position, sounds sources, and VE geometry, stereo audio is presented to the user.

VR haptic (tactile) information is presented to the participant through active feedback devices. Examples of force feedback devices include a vibrating joystick (e.g. vibrating when the participant collides with a virtual object) and the Sensible Phantom, which resembles a 6 DOF pen. Active feedback devices can provide a high level of HCI fidelity. Typically, VE participants prefer a traditional interface (Graphical User Interface or GUI) for interaction tasks.
and a VR interface (Perceptual User Interface or PUI) for visual tasks. By adding haptic feedback to VR design systems participants can make faster decisions [Vance 2001]. A collaborative use of virtual 3D display systems along with stereoscopic systems helps to detect design errors faster when using VE interfaces [Satter 2005].

1.2.3 Immersive Virtual Environments and Their Characteristics

With input and output, along with computation power to develop and present visual and other sensory information in real time, the user can experience “immersion” in the virtual environment. The unique characteristics of immersive virtual reality can be summarized as follows:

- Head-referenced viewing provides a natural interface for the navigation in three-dimensional space and allows for look-around, walk-around, and fly-through capabilities in virtual environments.
- Stereoscopic viewing enhances the perception of depth and the sense of space.
- The virtual world is presented in full scale and relates properly to the human size.
- Realistic interactions with virtual objects via data glove and similar devices allow for manipulation, operation, and control of virtual worlds.
- The convincing illusion of being fully immersed in an artificial world can be enhanced by auditory, haptic, and other non-visual technologies.
- Networked applications allow for shared virtual environments.
1.3 Interactions in a Virtual Environment (VE)

Though recent improvements in the fidelity of displays and the precision of tracking equipment have brought the users closer to simulating real-world environments, interaction with VEs still differs dramatically from real-world interaction. VEs tend to be impoverished versions of the physical world, providing incomplete sensory cues and inconsistent world models. Often, displays provide no feedback to secondary sensory modalities such as haptic channels. Even when a sensory channel is simulated, a full set of sensory cues may not be present, for example, dynamic shadows may be missing from visual displays and echoes may be missing from auditory displays. Virtual events may also be inconsistent with the world model, and lack reference from naturally occurring constraints [Mine et al, 1997]. For example, a table moved in the virtual world may not be constrained to slide across the floor, as it would be in the real world.

For a virtual simulation to succeed, the participant must construct a mental model of the world and its characteristics from the available cues of the visual, aural, and tactile displays. Once constructed, this mental model must be reinforced through subsequent interactions, requiring consistency from the VE implementation. There exists no unified framework for VE interaction, no desktop-style metaphor familiar to the majority of participants, and no optimal interaction technique for all possible task and input devices in VEs [Poupyrev et al., 1997]. Most frequently, the use of VE requires training. Interaction techniques in VEs may be unfamiliar to novice users. Interaction devices may not be self-revealing, providing limited clues to their use [Norman, 2002]. Finally, cognitive, perceptual, and motor differences between participants may impact the effectiveness of a VE and its perception by the participant [Poupyrev et al, 1997].

Several methodologies exist to help in the design, evaluation, and application of interaction techniques for VEs. The two most common of these include a sequential approach to usability
evaluation and the testbed evaluation method [Bowman, et al, 2001; Bowman, et al, 1999]. Sequential evaluation is a usability engineering approach based on enhanced versions of several existing 2D and GUI usability evaluation methods. In order to evaluate a VEs user interface, the developer performs user task analysis, heuristic evaluation, formative evaluation, and summative evaluation either serially or iteratively. In contrast, testbed evaluation describes a process of empirically assessing interaction techniques for VEs in a generic context through the description of taxonomy of interactions for the tasks being evaluated. The testbed evaluation method is widely used in the VR community as a means to identify low-level interaction tasks and their optimal implementation. Both methodologies note that the choice of an interaction paradigm is dependant upon the conditions of the task, the given display device, the chosen input device, and the interaction method.

An interaction technique’s success is most frequently measured by participant performance. This metric may include both task performance as a measure of the quality of task completion (e.g. speed, accuracy) and technique performance as a measure of the experience of the participant (e.g. usability, ease of use, learn ability, fun, and user comfort) [Bowman, 2002]. The former measure is more quantitative, while the latter is typically measured more qualitatively. Use of these two metrics implies the potential for trade-off between task and technique performance. Correspondingly, computer interaction modalities are frequently characterized by ease versus expressiveness tradeoffs. Thus, the efficiency with which the user can remember commands is often limited when the number of commands (i.e., the expressiveness) increases [Martin, 1989].
1.4 Interaction Devices for Virtual Reality

Presently, a set of devices, hand measurement hardware, HMDs, 3D audio systems, shutter glasses, and speech recognition systems are available. Also, many research labs are working on developing new devices such as force-feedback devices, tactile gloves, eye-tracking devices, or on further improving existing devices such as HMDs and tracking systems.

1.4.1 Head-Mounted Display (HMD)

The head-mounted display (HMD) was the first device providing its wearer with an immersive experience. Evans and Sutherland demonstrated a head-mounted stereo display in 1965. It took more than 20 years before VPL Research introduced a commercially available HMD, the famous "EyePhone" system (1989). A head-mounted display (HMD) is shown in Figure 1.1.

A typical HMD houses two miniature display screens and an optical system that channels the images from the screens to the eyes, thereby, presenting a stereo view of a virtual world. A motion tracker continuously measures the position and orientation of the user's head and allows the image-generating computer to adjust the scene representation to the current view. As a result, the viewer can look around and walk through the surrounding virtual environment. To overcome the often-uncomfortable intrusiveness of a head-mounted display, alternative concepts (e.g., shutter glasses) for immersive viewing of virtual environments are available.
1.4.2 CAVE™

The CAVE™ (Cave Automatic Virtual Environment) was developed at the University of Illinois at Chicago and provides the illusion of immersion by projecting stereo images on the walls and floor of a room-sized cube. Several persons wearing lightweight stereo glasses can enter and walk freely inside the CAVE™. A head tracking system continuously adjusts the stereo projection to the current position of the leading viewer. A CAVE™ system schematic is shown in Figure 1.2.
1.4.3 Data Gloves

Data gloves provide feedback to the computer regarding the motion of the digits on a left or right hand. The data gloves can be tracked, and this provides the location of the hand in addition to the position of all digits. A data glove, shown in Figure 1.3, allows for interactions with the virtual world as illustrated in Figure 1.4.

![Figure 1.3: Data Gloves](image1)

![Figure 1.4: Application of Data Gloves](image2)

1.4.4 Pinch Gloves

Pinch gloves are VR peripheral devices used for interacting with a simulation. Resembling winter gloves these devices (two gloves) fit on the left and/or right hand. Each pinch glove is lined with wires extending to all five fingertips of each hand. When the thumb of either hand touches one of the remaining four fingers the pinky, ring, middle, or index, a circuit is completed and a signal is sent to the computer, and hence to the application, which in turns sends an event triggering a user defined logical action. The pinch glove system provides a method of recognizing natural gestures. Recognizable gestures have natural meaning to the user. A pinching gesture can be used to grab a virtual object, and a finger snap between the middle finger and thumb can be used to initiate an action. It is a hand-gesture interface system that allows
developers and users of immersive applications to use hand interaction to work within the virtual environment.

Such actions may be selecting objects, opening and closing doors, moving objects, etc., much as hands are used in real world everyday activities. [Fakespace, 1997]. Figure 1.5 provides a photograph of the pinch glove system

![Pinch Glove System](image)

**Figure 1.5: Pinch Glove System**

1.4.5 Joystick, Wand, or Flightstick

A joystick, also called a wand or flight stick is simply a device that is formed to fit comfortably in the hand. It typically contains a tracking device to determine the location of the joystick (or wand or flightstick), and there are usually switches and buttons on the device that can be pressed to signal specific operations to the computer system. Figure 1.6 provides a photograph of a joystick system developed by cannibalizing a commercial toy PC joystick.
1.4.6 Shared Virtual Environments

In the example illustrated below in Figure 1-7, three networked users at different locations (anywhere in the world) meet in the same virtual world by using a BOOM device (Binocular Omni Orientation Monitor is similar to a head-mount except that there's no fussing with a helmet as shown in the figure 1.7 to the top left.), a CAVE™ system, and a Head-Mounted Display, respectively. All users see the same virtual environment from their respective points of view. Each user is presented as a virtual human (avatar) to the other participants. The users can see each other, communicate with each other, and interact with the virtual world as a team [Beier, 2004].

Figure 1.6: Flightstick
Today, the term 'Virtual Reality' is also used for applications that are not fully immersive. The boundaries are becoming blurred, but all variations of VR will be important in the future. This includes mouse-controlled navigation through a three-dimensional environment on a graphics monitor, stereo viewing from the monitor via stereo glasses, stereo projection systems, and others. Apple's QuickTime VR, for example, uses photographs for the modeling of 3D worlds and provides pseudo look-around and walk-through capabilities on a graphics monitor. [Beier, 2004]

1.4.7 ImmersaDesk

The ImmersaDesk was developed in 1994 at Electronic Visualization Laboratory (EVL) of the University of Illinois at Chicago. It is a drafting table format VR display. It features a 67x50-inch rear-projected screen at a 45-degree angle. Up to 5 users wear shutter glasses to view high resolution, stereoscopic, head tracked images. The ImmersaDesk screen mostly fills a user's field of view, and at the same time enables the user to look forward and down. One user's head is tracked, allowing an accurate perspective to be generated based on that user’s position. A tracked wand is also used, so that the user can interact with the environment. The system is equipped
with stereo sound. The ImmersaDesk cabinet is on wheels and folds up and fits through doors [Czernuszenko et al. 1997]. Figure 1.8 shows the Immersadesk system.

Figure 1.8: The Immersadesk System

The projector is located in the lower section and a pop-up mirror folds the optics. The screen can rotate to the transportation position, where it is enclosed in the ImmersaDesk body, allowing the system to have a footprint of 34 inch depth by 73 inch width.

Only one graphics pipe is needed for the ImmersaDesk, which allows the use of a less expensive mid-range workstation. The decision to tilt the screen 45 degrees came from experience with the CAVE™. In the CAVE™, users usually look at images that are displayed on the walls. This might suggest that the CAVE™ floor is not important. However, if the floor is not used, a significant part of the VR experience is lost. This idea resulted in the design of the ImmersaDesk to support looking down, as well as forward [Czernuszenko et al. 1997].

The ImmersaDesk user stands close to the screen, creating a 110 degree horizontal field of view. In that case, parts of the screen are viewed at large angles. In order for the entire image to appear...
uniformly bright, one has to choose a low gain screen. This screen is made of a clear plexiglass with sprayed on backcoating [Czernuszenko et al. 1997].

All parts of the ImmersaDesk cabinet are made of wood or stainless steel, so as to minimize interference with the electromagnetic tracking system. The ImmersaDesk uses the same CAVE™ library software as is used in the CAVE, to generate accurate perspective projection, and to read tracker and input devices. Therefore, applications developed for the CAVE can be run on the ImmersaDesk, and vice versa, without any code changes [Czernuszenko et al. 1997].

1.4.8 CrystalEyes

CrystalEyes glasses allow the user to have a left and right eye view of the scene. The CrystalEye glasses use LCD shutters to occlude either the left or right eye at a rate of up to 120 Hz, which is so fast that it can not be perceived by the user. The CrystalEye glasses remain in synchronization with the computer image using an infrared emitter, with the infrared receiver found above the nose on the glasses. Using real time tracking as the user’s head moves from side to side, closer to or furthers away from the monitor, the image on the display changes its perspective, giving the convincing illusion that the image is a real object. The user's hands are left free to manipulate the data. CrystalEyes systems presents the viewer with a high-resolution, full-color, stereoscopic virtual world. With six degrees of freedom, rapid response and flicker-free viewing, images come alive; and because of headtracking, the image remains in concert with the user’s head movement naturally. Additional users can simultaneously view the virtual world. CrystalEyes systems support multiple viewers. CrystalEyes stereoscopic eyewear is as comfortable and lightweight as a pair of eyeglasses. [VR Depot, 2005]. Figure 1.8 illustrates a user wearing CrystalEye glasses.
With these different pieces of hardware and with the various configurations that can be developed, virtual environments based on real time, tracked stereoscopic viewing can maintain a sense of immersion in the virtual environment that is akin to the suspension of disbelief that occurs when an individual becomes lost in a play or movie. Essentially, the individual believes in the existence of the world created, and the experience is similar to a real sense of immersion in the artificial world.
2. GESTURAL INTERFACE

2.1 What is a Gesture?

Perceptual, cognitive, and usability evaluations have motivated the use of gestures as an interaction technique in virtual environments for their potential to increase speed and accuracy in task performance as well as improve usability. A gesture may be defined as a physical movement of the hands, arms, face, and body with the intent to convey information. Gesture recognition, then, consists not only of the tracking of human movement, but also the interpretation of that movement as semantically meaningful commands. Interpretation can vary in resolution and encompass a range of large and small scale motions, including tracking in which the subject is viewed as a single object, tracking of the subject as an articulated kinematics structure, and tracking of small-scale movements such as facial expression and hand gestures.

Gestures are identified by their function, their linguisticity, and their role in communication. Gestures are grouped according to function as semiotic, ergotic and epistemic [Cadoz, 1994].

Semiotic gestures convey meaningful information by facilitating communication. Semiotic gestures are frequently derived from shared cultural experience. Waving good-bye or giving someone the “finger” are examples of semiotic gestures in the USA. Ergotic gestures include manipulations of the physical environment, and are frequently related to the notion of work. The act of putting an object onto a shelf is ergotic gesture. Epistemic gestures involve the process of discovering the environment through tactile experiences. Judging the weight of an object by holding it in one hand or is an example of epistemic gesture.

Gestures are also classified according to their linguisticity. This classification forms a continuum ranging from gesticulation as the most multisemiotic to sign language as a well-defined linguistic system and an autonomous semiotic gesture set [Kendon, 1988]. Proceeding down the scale from
sign language towards decreasing autonomy are emblematic, pantomime, and language-like gestures. Emblematic gestures are symbolic gestures which, as noted above, are frequently culturally-specific representations. Pantomime gestures accompany speech by depicting objects or actions. For instance, a pantomime occurs when a person describes the size of a fish. Language-like gestures are those gestures which are fully integrated into speech, often replacing a particular word or phrase. These frequently may be full-body gestures including posture and facial expressions. Gesticulations, the least autonomous of the semiotic gestures, include spontaneous and idiosyncratic movements of the body that accompany speech. Gesticulations are seldom culturally defined and almost never occur in the absence of speech.

Semiotic gestures are further classified by their role in communication as iconic, metaphoric, deictic, and beat-like [McNeill, 1992]. Iconic gestures are representative of an action, object or event, while metaphoric gestures depict a common metaphor rather than depicting the event or object directly. Pointing gestures which are used to indicate people, objects, or directions are deictic gestures. Beat-like gestures are small, emphatic gestures generally performed with the hand or the head.

Another method of classifying gesture makes use of four dichotomies: act-symbol, opacity-transparency, autonomous semiotic-multisemiotic, and centrifugal-centripetal [Nespoulous, 1986]. Examination of these dichotomies provides insight into the range of communication and command capabilities of gesture, as well as some of the issues related to the use of gesture-based interaction techniques.

The first of these, the act-symbol dichotomy, describes how gestures may be defined by action or may be symbolic in nature. Action (ergotic) gestures occur when a person performs a task, for example, playing the piano. Symbolic or semiotic gestures are representative of a concept or
emotion, such as giving someone “the finger” or a “thumbs-up” signal. Action gestures are particularly applicable to use in gesture-based interaction techniques for commands which parallel a real world task. Symbolic gesture sets may also be effectively implemented when their recognition rate is high across the user population.

The opacity-transparency dichotomy expresses that gestures may not be easily accessible to all individuals, and that many gestures lack universality. A higher rate of recognition across individuals and cultures is associated with transparency. While the notion of cross-cultural gestures seems plausible, few if any known body motions or gestures have been identified to have the same meaning in all societies [Birdwhistell, 1970]. American Sign Language (ASL) is an example of an opaque gesture set. Though many signs are reminiscent of the words or concepts they represent, an observer unfamiliar with the language would be unable to interpret a signed communication. As a result, the use of sign language gestures in an interaction technique would require advanced training and make it difficult for novice users to interact with the application.

The next dichotomy in gesture classification is autonomous semiotic versus multisemiotic. This dichotomy refers to the function of a gesture as either a member of a language or as accompaniment to language. ASL is an example of an autonomous, semiotic system because it is as a complete gestural language unto itself. Gestures which accompany or enhance speech are multisemiotic in nature. Gestures which indicate relative size or location during conversation or add emphasis to a key point would be classified as multisemiotic. Gesture recognition methods typically rely on the creation of a semiotic gestural command set which is specific to the application and tasks. This command set may supplement other interaction techniques such as voice recognition or may exist as the key form of communication between the user and the VE.
Finally, the centrifugal-centripetal dichotomy describes a gesture’s direction of intent. Centrifugal gestures are directed towards specific objects or people, whereas centripetal gestures are not. A pointing gesture to indicate which object is being discussed would be centrifugal in nature, whereas a gesture using the hands to indicate relative size or location would be centripetal.

2.2 How to use gestures?

Incorporating hand gestures into an existing GUI permits exploration of how it meshes with current interface technology and how it compares to current interface devices. Using this knowledge one can determine if future systems can benefit from using gesture, and if so, what changes must be made to accommodate gesture.

While the interface envisioned here looks in many ways similar to standard GUIs of today, this thesis does not attempt to argue that simply using gesture as a direct mouse replacement in a current GUI gives any real advantages. Indeed, many aspects of current interfaces are tuned to complement specifically the capabilities of the mouse and keyboard, and as such are not well suited for gesture. It is argued, however, that in the context of an appropriately designed interface, gesture can offer real advantages as an interface modality.

While gesticulation and other multisemiotic categories of gesture play a key role in interpersonal communication, gesture-driven interfaces most frequently make use of emblematic and pantomime gestures. Emblematic and pantomime gestures are less spontaneous than gesticulation, and also tend to be culturally specific, and therefore learned rather than innate. Along with ergotic gestures and deictic gestures, these types of semiotic gestures are clearer in their intended meaning, making them more appropriate for expressions of intent in virtual
environments. Within an individual cultural group, these gestures may also be more consistent in their performance, thus improving tracking and recognition.

Applications that used ergotic (action-representative) gestures have included general object manipulation, manipulation of molecules, interaction with scientific visualization data, manipulation of financial and n-dimensional data sets [Feiner et al., 1990], specification of 3D mouse interactions [Weimer et al, 1992], and robotic control [Sturman et al, 1994]. VEs making use of semiotic gestures have included deictic-style navigation [Bolt, 1980;], natural navigation [Krueger, 1993], sign language interpretation [McGuire et al., 2004; Fels et al, 1995], robotic control [Waldherr et al., 2000], and control of multimedia presentations [Baudel et al., 1993]. A final category of gesture-based devices relies on semiotic gesticulations for interpretive purposes in controlling aural or visual interactive performance spaces [e.g. Paradiso et al., 1997].
2.3 Usability and Ease of Use of Gestures

Gesture-based interaction techniques have found both support and criticism in the area of usability. Gestures have been offered as a means of “natural interaction,” for their properties of direct interaction, their flexibility, and their reality of experience. However, gestures have also been described as imprecise, non-ergonomic, and not self-revealing.

For new VR users, knowledge of how to manipulate objects or travel through an immersive environment, or even the knowledge that such tasks might be available, is often not self-revealing. VR applications are primarily designed with an expert user in mind, and new applications tend to require training on the interaction techniques even for those familiar with the technology. Information about how to interact with the VE is seldom “stored in the world” [Norman, 2002], with the physical devices giving clues to their use. Gesture recognition overcomes this problem by making use of emblematic, pantomime and natural ergotic gestures to represent common actions or semantic expressions [Kendon, 1972]. Examples of these might include holding the hand up palm out for “stop” or both hands out in front with fists clenched to mimic the act of riding a bicycle. These are symbolic gestures which carry clear and familiar, although, sometimes culturally specific meaning, and can be associated with VE tasks to provide both new and experienced users with a set of intuitive commands.

While gesture commands are frequently selected for their ability to parallel a real-world task, the ability of the VE to perform this task is not always immediately clear to a new user. Gestures themselves are not self-revealing, lacking the discoverability afforded by menu and button-based paradigms [Baudel et al., 1993]. For this reason, their use in VEs must be prefaced with an explanation of the interaction technique’s capabilities. Visual reminders must be available to provide guidance to the user and enable learning of more complex or less intuitive techniques.
As gesture recognition techniques improve, gesture has the ability to be a highly flexible means of interaction. Instead of confining the user to an arbitrary, rigidly defined gesture set, gesture interaction techniques can adapt to changing user characteristics, including changes due to fatigue, level of experience or disability [Williams et al, 1990]. Allowing for flexibility in input styles would ensure a broader user base and could help reduce repetitive strain and overexertion. Improvements in recognition could also help combat complaints regarding the lack of precision in gesture-based interactions. While some variability in gesture systems may be due to limited precision of the tracking equipment itself [Mine et al, 1997], blame may also lie with the developer. Gestures may be tracked at the level of the body, the hand and arm, the hand and fingers, or the head and face [Turk, 2002]. The choice of resolution imposes constraints on the precision and repeatability of the gesture commands. For example, the average user on a full-body scale cannot replicate the fine motor control and accurate positioning that is possible with the fingers. In his development of a hand-based gesture system, Wexelblat [1995] quantified movements in more “expressive” joints of the hand (e.g. the index finger) as more significant to determinations of gesture than the movements of less expressive joints (e.g. the pinkie). Thus, movements in the pinky must be larger to be recognized by the system than what would be required of the higher-precision index finger. At the implementation level, Mulder [1996] argued that lower precision and computational power was required for semiotic gestures than for ergotic motions. The latter requires the accurate detection of more complex, continuously changing motions through dynamic gesture recognition, while the former could be tracked by recognizing a discrete number of postures or positions through static gesture recognition.

As the user is constantly being tracked in a gesture-based system, the potential for “immersion syndrome” is high [Baudel et al., 1993]. Gesture systems are passive in that they are always on
and monitoring the user’s motions. Thus, gestures can be evaluated by the system whether or not they were intended for interpretation. This constant monitoring may limit the user’s ability to communicate via gesture to other devices or with other people in the environment, if gestures are not well selected and defined.

Body-centered and arm and hand-based gestures have the potential to suffer from additional precision degradation due to ergonomic factors. Care must be taken to design gestures that do not require awkward posturing, repetitive motion, and excessive repetition. As gestures do not introduce external forces on the body, they could be described as ergonomically superior interaction methods; however, the lack of registration with a physical surface provides no external frame of reference, and no means to steady the body. Under such conditions, gesture has the potential to cause tension and fatigue. Increased arm fatigue due to tension or posturing has been found to degrade performance [Baudel et al., 1993]. To reduce fatigue, gesture commands must minimize effort and be quick and easy to execute. User discomfort may also occur in cases in which users must wear a glove or other tracking devices on the body and be linked to the system via wires. Improvements in computer vision techniques continue to alleviate these problems by providing effective, wireless solutions.

In addition to impacting a virtual environment’s usability, ease of use, and physical comfort, gesture based interaction techniques has the potential to impart desirability, reality of experience, and a sense of fun. Bowman [2002] described the success of a VE as capable of being measured in task performance and technique performance. While the former metric deals with quality of task completion and accuracy, the latter describes the qualitative experience of the user. In this category, gestures have the potential to excel beyond wand and button-based controllers by providing involvement, attention, interest, and realism. High reports of presence have been
correlated with reality of experience, including successfully supported action in the virtual environment [Slater, 2004; Zahoric et al., 1998].

2.4 The Role of Non-Symbolic Gestures

While the selection of intuitive gestures favors learning and ease of use, it reinforces the trade-off between ease and expressiveness [Martin, 1989]. Complex and non-symbolic gestures may be difficult to learn and retain, but have the potential to provide a greater range of expression and control than do emblematic or pantomime gestures [Baudel et al., 1993]. Developing gestures for command and control tasks such as save, load, and change color commands may be particularly challenging, as these have no consistent physical manifestation. Wexelblat [1995] noted that poorly chosen mapping of gesture input to commands might provide minimal functional gain over a button or key-based system. Instead of enabling ease of use, such techniques may induce cognitive load due to requirements for the user to memorize the gesture set. In the desktop domain, evidence points toward a gain in efficiency and control once gestures are learned. Through the use of shortcut key commands, for example, an expert user can easily outperform a less experienced typist who must move back and forth from mouse to keyboard input. Though initially unintuitive, these command sets provide a physical mnemonic and reduced motion set, which improves accuracy and performance.

Great success has been shown in the application of gestures to desktop interfaces, with the most notable example being the TouchStream keyboard and iGesture pad by FingerWorks [FingerWorks, 2002]. These systems use a technology called MultiTouch to sense and interpret the motion of multiple fingers on touch imaging surfaces. The TouchStream keyboard allows users to combine touch-typing, pointing operations, and intuitive hand gestures on a single, zero-force device. The result is that gesture shortcuts can be performed anywhere on the keyboard’s
surface, eliminating hand movements to the mouse or the use of hotkey sequences. Finger combinations and a direction of motion define gestures. For example, the user can open a new file by touching the keyboard with the thumb and first three fingers and rotating on the pad in a counterclockwise direction in a manner similar to opening a jar. While some gestures are relatively symbolic, others are more arbitrary. To copy a selected item, for example, the user touches the keyboard with the thumb and middle finger. To paste copied text, the user touches the keyboard with the same finger combination and then spreads the fingers on the pad.

TouchStream’s gesture set had been designed to be fairly intuitive, but it still requires a modest training time. FingerWorks does not recommend the TouchStream keyboard for older users who may have limited hand agility or users unwilling to spend time “relearning” how to type. With constant use, touch typists are expected to reach moderate speeds of 30-40 words per minute within a few days, with a return to full proficiency taking 3-4 weeks of practice [FingerWorks, 2002]. Average typing speed is 50-60 words per minute while accomplished typists reach speeds of 60-70 words per minute.

2.5 Task-Based Evaluation of Gesture Interactions

Gesture-based methods are not the most effective interaction scheme in all instances, and gesture recognition should not be used purely for its own sake. It cannot replace the precision of some interaction devices, and may be a less effective paradigm in some cases than more standard input techniques. However, gesture-based approaches do have unique advantages over other input technologies.

The following section examines the domain of VE tasks that are well suited to the use of gesture. To structure this classification, the testbed evaluation method is used to identify low-level tasks and their potential for implementation in gesture-based interfaces. Testbed evaluation describes a
process of empirically assessing interaction techniques for virtual environments in a generic context through the description of taxonomy of interactions for the tasks being evaluated [Bowman & Johnson, 2001; Bowman, Johnson, & Hodges, 1999; Poupyrev, et al, 1997]. In Bowman Johnson, and Hodges [2001], a set of universal interaction tasks for virtual environments are defined, including selection, manipulation, release and travel. These tasks are subsequently broken down into sets of separable subtasks. For example, the manipulation of an object might be composed of three subtasks: specifying the position of the object, specifying the orientation of the object, and providing feedback to the user. These subtasks may then be broken down into finer-grained descriptions including the type of interaction method implemented. Thus, the process of specifying the new location of the object in a manipulation task may be implemented with xyz sliders or by indicating a point in the 3D space with a wand or glove device. An interaction technique is made up of one technique component from each of the lowest-level subtasks.

2.6 Gesture in Selection

The task of selection encompasses the specification of an object or set of objects, most frequently for the purposes of manipulation or as the referent of a command. Selection is a frequently occurring task, and thus should be implemented to maximize efficiency. Gesture is an ideal interaction technique for selection tasks as it may be implemented in a way that closely mimics real-world interactions. Gesture-based selection is most frequently implemented via a virtual hand metaphor, pointing or arm extension, or through occlusion or framing of an object. The virtual hand metaphor is intended to simulate real-world interaction, allowing the user to reach out and “grab” or select objects in the VE in a natural way. Virtual hand implementations exhibit difficulty when objects are beyond the reach of the user; however, several “magic”
techniques have been developed to extend the user’s grip to remote locations in the VE. These arm extension techniques may be used to supplement gesture-based selection controls. A common example of an arm-extension implementation is the Go-Go technique [Poupyrev, et al. 1996], in which nonlinear mapping of the user’s hand position is applied when the physical hand exceeds a radius of proximity to the body. Thus, when the hand operates outside that radius, it may select distant objects using the same techniques that would be used to interact with objects closer to the body.

Ray-casting techniques make use of deictic or pointing gestures. A ray directed by the user’s hand and arm posture is used to indicate referent objects within the scene. Ray-casting saw its first application in Bolt’s “Put-That-There” interface (1980), in which gesture was used to disambiguate pronouns and allow for the specification of unknown or remote objects in the scene. Bowman and Hodges [1999] assert that ray-casting performance is more efficient than arm extension over a range of object distances, object sizes, and object densities. Their belief is that this difference is due to their reduction of Ray-casting to a 2D task through the elimination of changes in the roll of the wrist or hand. Ray-casting may be implemented as either a 2 or 3 degree-of-freedom task.

The final method of gesture-based selection is occlusion or framing selection. This selection method combines Ray-casting with gaze direction. In these implementations, a ray emanates from the user’s eye point and passes through his/her hand position to indicate the referent object [Pierce et al., 1997]. Through the combination of these two techniques, framing selection attempts to provide more accurate selection and pointing control by ensuring a line-of-sight casting.
2.7 Gesture in Manipulation

Manipulation is the second task classification in the taxonomy of VE interaction tasks, and most frequently follows selection. Manipulation refers to the user’s ability to change the properties or positions of objects in the VE. VE manipulation may include changes to the orientation or scale of objects as well as changes to other attributes such as shape, color and texture. It should be noted that command and system control tasks such as the interaction with menus or control panels do not fall under the manipulation heading, instead being classified as selection techniques.

Gesture-based manipulation tasks may take the form of direct object manipulation, indirect object manipulation, relative manipulation, and direct viewpoint manipulation. Direct object manipulation is an extension of the virtual hand metaphor for manipulation, and is a simple and intuitive means of controlling the position and orientation of a selected object. In these cases, the selected object is attached to the virtual hand and the actions of the physical hand are used to transform the object [Poupyrev et al., 1996]. Kinematic constraints limit the number of positions a user can achieve, however, so care must be taken to implement an approach that allows the user to quickly deselect the object and cease manipulation once a constraint is reached. Direct manipulation has been shown to perform more efficiently and provide a higher level of user satisfaction than techniques that involve tool-use or indirect manipulation such as Ray-casting [Bowman & Hodges, 1999]. The combination of Ray-casting for selection and direct manipulation for object transformations is the basis of the HOMER technique [Bowman & Hodges, 1997] and Sticky Finger [Pierce et al, 1997].

Relative manipulation is a class of techniques that make use of two-handed or body-relative interaction. Many researchers have described the benefits of implementing two-handed input for
interactive applications. The use of two hands increases the flexibility and expressiveness of the available command set by increasing the number of possible commands and physical mnemonics. Remote manipulation of objects becomes simple and intuitive with relative positional commands. The non-dominant hand may be used to provide a frame of reference, while the dominant hand is used to specify more precise relative transformations.

While manipulation is most frequently applied to objects in the scene, it may also be used to update view orientation, if this is not specified via the orientation of the user’s head. Direct object manipulation operates via a “camera in hand” or “scene in hand” metaphor in which the selected object remains fixed in the environment and the user manipulates the viewpoint using gesture. In general, this approach has shown to be less effective than viewpoint updates specified via head orientation as it limits the user’s understanding the spatial structure of the complete environment [Chance et al.1998]. Manipulations of this sort are commonly implemented and highly effective in design, prototyping and simulation applications.

2.8 Gesture in Travel

The most common VE interaction task is travel. Travel is the process of viewpoint control via movement and way finding. The taxonomy of passive movement interaction techniques developed by [Bowman et al., 1997] partitions travel into three subtasks: direction/target selection, velocity/acceleration selection, and input conditions. Arns [2002] offered a travel taxonomy defined by two major components, translation and rotation, which were broken down into physical and virtual implementations. Gesture-based implementations for travel include direct steering techniques, target-based techniques, relative motions tasks, and gesture-designated locomotion.
Gesture input is well suited for control of direct steering and target-based travel techniques. The former can be implemented in a variety of ways, including gestures which mimic steering of bicycle handlebars or the wheel of a car or simple mapping of the direction of travel to the posture of the user’s arm or hand. Steering techniques allow the user to look at objects of interest while moving, and take advantage of body-centered cues for direction. Target-based travel techniques provide the simplest metaphor for navigation, in which the user specifies a target in the environment to which the application should initiate travel. This method assumes that the goal location is known in advance and is visible by the user, though the technique may be implemented in a general manner that operates similar to steering techniques in the absence of a target. While target-based techniques have been found to be easily understood by novice users, steering techniques afford a higher degree of control and provide greater levels of spatial orientation [Bowman et al. 1999].

Navigation via relative motion of the body may be implemented in ways similar to relative manipulation and viewpoint techniques. Two-handed interactions may be utilized in flying or driving paradigms, in which the positions of the hands relative to one another control not only the direction of travel but also affect speed control via hand separation. Mine et al., [1997] recommended the use of head and hand posture to enable users to quickly switch between close, local views and more global distant views using a technique called “head-butt zoom.”

The final application of gesture to travel techniques is the use of gesture-designated locomotion. At its simplest level, this category includes physical rotation and translation [Arns, 2002]. The choice of physical navigation over virtual navigation is highly system-dependant. Six-sided and untethered display/tracking scenarios are ideal environments in which to take advantage of the spatial orientation afforded by physical rotation, whereas single-screen systems require virtual
translation and rotation to maintain the user’s focus on the display. Locomotion interfaces are frequently used in military simulation and training applications. It has been found that locomotion calibrates distance judgments and may contribute to an increased sense of presence and task transfer [Hollerbach, 2002]. Locomotion interfaces may be used to simulate walking, running and climbing activities and may be triggered via gestural knee actions, classifications based on head bobbing, and detection of hand gestures to indicate ascending or descending a ladder [Slater et al, 1995]. The combination of gesture and mechanical locomotion solutions shows promise in aiding in the development of advanced interaction techniques for training and simulation.

In all selection, manipulation and travel tasks, gesture has the potential to supplement and enhance existing interaction techniques by providing a redundant input method. Individual differences dictate that users can find some interaction techniques are more intuitive or easy to learn than others [Bowman, 2002]. By providing multiple interaction techniques for each task or a subset of tasks, developers can mitigate these ease-of-use and ease-of-learning challenges. In addition, the availability of multiple levels of interaction allow for the individual user to tailor the VR experience to suit the task, training and skill level, cognitive and motor physiology capabilities, and physical comfort.
2.9 Gesture in Traditional User Interfaces

Relatively little work has been attempted to use gesture in a traditional user interface. Of that which has been done, the most common approach is to use gesture as a direct mouse replacement. This has most often been done using indirect positioning. In other words the user moves their hand within some control space to move the cursor about the screen in an analogous fashion, rather than pointing directly at the screen to indicate exactly where they want the cursor to go. Both absolute positioning, where the location of the hand or an extended finger within the control space is mapped directly to a screen location, and relative positioning, as is done with a mouse, has been used. Typically some action, such as a change in pose or a key press by the other hand is used to simulate a mouse click.

Nesi and Bimbo [1995] used two cameras to position the mouse in 3-space. The hand is observed against a black background in a workspace, presumably to the side of the keyboard. The motion of the hand is smoothed using a predictive polynomial filter. To take the place of mouse buttons, three hand poses are used: palm down with the fingers extended and together, rotated from that 90 degrees so the palm faces sideways and the thumb is up, and palm down with the fingers curled in a fist. The poses are differentiated by taking the ratio of the sides of the bounding box of the hand when viewed from above.

Quek [1995] describes a system called Finger Mouse. The system is designed to allow the user to switch from typing to moving the mouse simply by assuming a pointing pose with their hand above the keyboard. The camera looks straight down on the hand from above. A finite state machine examines the shape of the segmented hand to determine when a finger is extended. When it is, the mouse is tracked by the location of the fingertip in the plane above the keyboard. Pressing the shift key while pointing triggers Mouse clicks. The system has been tested by
having users fill out on-screen forms, using pointing to select the field to type in. Some work has been directed at designing a workstation to make greater use of gesture.

Maggioni [1995] describes several additions to a conventional workstation that allows it to use both hand gestures and head movements. One camera images the user's face, another looks down on a region to the side of the keyboard to image the user's hand. When the hand is on the desk, its position is used to position the cursor like a conventional mouse. When it rises off the desk it enters a 3D mode where movement in the center of the imaged volume positions a 3D cursor. When the hand nears the edge of the control volume it moves the observer's viewpoint of the virtual space. Maggioni describes several hand poses that can be differentiated, but does not suggest how they might be used.

Wellner [1993] describes his work on an automated desktop that lets the user seamlessly combine physical and digital media. As part of the interface, a camera positioned above the desk observes the user's hands as he/she interacts with both paper and digital data projected onto the desktop. The system uses motion to find the hand, and segment it from the background. One can then determine the location that the user's finger or a stylus to which the finger is pointing. Wellner makes no attempt to classify the pose. Several novel interaction modes are suggested, such as using the finger to draw a circle around a graphic on a sheet of paper, then pointing to where it should be placed in a digital document. The system would then digitize that portion of the paper and place the digital graphic where the user indicated.
2.10 Guidelines for the Design of Gesture Command Sets

As a means to collect the advice of others, guidelines have been collected that provide some wisdom to the virtual environment’s designers and programmers. These guidelines consist of nine principles that help to provide better user experiences in the virtual environment. These nine are:

1) Provide guidelines to the user as to how the gestures have been implemented in the virtual environment. Users may not share the associations between gesture and command that are clear to the developer, making the gesture set non-intuitive and non-self revealing. For example, in order to travel within the VE, a user may march in place to simulate walking. If the developer implemented a leaning gesture as the only way to travel, the user’s command would fail. This conflict between expectation and implementation can cause breaks in presence and frustrate the user.

2) Provide continuous feedback to the user. Continuous feedback reinforces user confidence in the system and assures the user that a command has been recognized. When coupled with proprioceptive stimuli via gesture, feedback may encourage cross-modal transfers and enhance presence [Biocca, Kim & Choi, 2001].

3) Initiate a gesture with tension and conclude a gesture with a release from tension. Gesture-based interaction is, at its heart, a dialogue between the human and the computer. In order for the system to distinguish one word or phrase from the next, communication patterns must be consistent and direct. Tension emphasizes the structure of this dialogue, indicating the intention to issue a command. Correspondingly, relaxation of the muscles—a release from tension—may be used to denote the completion of the command [Baudel & Beudon-Lafon, 1993]. This behavior may be observed in mouse-based interactions. The selection of text begins with a click
of the mouse button. The finger remains tense as the user drags the mouse to select the text. Once the text is selected, the command is executed or completed by releasing the finger.

4) Provide the user the ability to cancel unintended actions. The principles of direct manipulation [Shneiderman, 1983] stress that the user should always be able to “undo” the previously issued command. Granting users this ability insures a feeling of control and reduces fear of causing irreparable damage.

Undo commands are of high importance in gesture-based systems, as the system may detect unintended gestures resulting from the user communication with other input devices or additional users. Additionally, users may require learning time to overcome errors and strengthen gesture command mnemonics.

5) Select gestures that can be executed quickly and easily. Quick and easy gestures reduce fatigue and favor ease of learning. In addition, gesture may be supplemented with additional input modalities to avoid requirements for frequent or awkward posturing.

6) Select gestures that do not require a high degree of precision. Precise gestures may be difficult for the user to execute consistently, and reliance on a high degree of precision by the system further hinders repeatability. Gesture provides limited tactile feedback beyond proprioception. Without a button to aim for or a visible path to follow, users must rely on more broad, body-based cues. Precision is further reduced by the lack of external devices that could help in steadying the body.

7) Select gestures that are distinctive. Gestures that are too similar may be difficult for both the system and the user to differentiate. Distinctive gestures help reduce confusion. Particular care should be taken in the use of non-symbolic gestures, as these gestures are not grounded in a familiar representation and could be more easily confused than their symbolic counterparts.
8) Assign symbolic gestures whenever possible, particularly to the most common commands. Symbolic gestures are the most intuitive and easy to learn, allowing novice and infrequent users to quickly interact with the system in a natural way. By assigning more complex or non-symbolic gestures to advanced commands, the user can increase their command set as they increase proficiency with the interface [Baudel et al., 1993].

9) Consider supplementing the gesture set with non-gesture methods to issue commands. Command and control tasks require non-symbolic gestures, and may be more effectively executed via menu selection or speech input. In addition, the use of multimodal systems could increase productivity, serve a broader range of skill-levels, and accommodate a greater number of active participants.

Using these guidelines, the developer of virtual environments can ensure that systems developed using gestures are effective and useful. Further, using the stereoscopic system hardware discussed in the previous section with input-output devices operating in real time, virtual environments can be created and maintained to have a sense of immersion in the virtual world that is created for the user to experience.
2.11 Objective of the Research

The aim of this research is to prove that the state of the art perceptual user interfaces (PUI) are better and efficient than the traditional graphical user interfaces (GUI). This is the age of Human Computer Interactions (HCI) and PUI. Data processing is done by computers through the body language of the humans.

Traditional GUI input devices like keyboard and mouse are being replaced by a built in camera and voice activated input within computers. The union of HCI with virtual reality has revolutionized this information age.

The recent work by Butler and Satter (2005) on competitive usability studies of virtual environment has opened doors for researchers to further investigate this interesting field of PUI versus GUI. Continuing on the same lines, an attempt has been made in this work to apply usability analysis for a virtual environment (CAVE™ with wand as navigation tool) to prove the superiority of PUI over GUI.
3. PROJECT DESCRIPTION

3.1 Background

This research project is an immediate outgrowth of a research collaboration including Maxwell (2001), Butler, and the Naval Research Laboratory (Maxwell et al., 2001), and work completed by Satter and Butler (2003), Butler et al. (2003 and 2004), and Satter et al. (2004) with approval from the UNO Institutional Review Board, Human Subjects Committee. It relies on other research efforts conducted at the Naval Research Laboratory for context in addition to the work of Maxwell and Butler and the recent work of Satter and Butler. In Maxwell’s DSVE (Design Synthesis Virtual Environment) system, the designer is provided with an immersive environment in which the engineer or naval architect uses a wand to select and manipulate three-dimensional geometry. The immersive environment uses two technologies to create an “artificial synthetic world.” They are: 1) real time tracking of head and/or limbs and 2) providing perspective rendering (generating images similar to photographs) to both left and right eyes. The wand is programmed to “read and respond” to human gestures. The 3D stereoscopic presentations are displayed in a CAVE™ immersive environment. With the work of Satter and Butler (2003), Butler et al. (2003, 2004), and Satter et al. (2004) competitive usability tests are employed to quantify the level of improvement from one interface to another interface.

Immersion is the sense of presence that is created when the reality of the virtual environment approaches that of the real world. Immersive technology gives the user the psychological experience of being surrounded by a virtual (computer generated) environment (van Dam et al., 2000). This experience is created with essential elements such as stereoscopic perspective vision, a display that permits the user to look in any direction, passive head and hand tracking,
and graphic computing power sufficient to achieve an adequate update rate. In such a system, the user perceives “real world” 3D existence. Typically, immersion comes in several forms: Head-mounted displays (HMDs), responsive workbenches, and CAVETM systems. HMDs present the user with small display screens positioned in front of the eyes. The responsive workbench provides an immersive environment based on projection through a single workbench screen, and CAVETM systems are specially constructed rooms with projectors on multiple surfaces, including walls, and possibly floor and/or ceiling.

For Maxwell’s DSVE, a multimodal interface is used. Multi-modal interfaces accept input to the computer using more than one paradigm. DSVE allows standard keyboard and mouse input; however, it also accepts gestural input with a hand held wand. It is this gestural input that allows designers to perform design activities intuitively by using a more natural gestural interface within the design environment, allowing more rapid, efficient actions to create and/or modify product geometry.

The DSVE was developed at the Naval Research Laboratory for use in the GROTTO. The GROTTO (Graphical Room for Orientation, Training, and Tactical Observation) is a CAVETM system developed by the Mechdyne Corporation, now a part of Fakespace, Inc. The GROTTO is a four-screen system that includes back projection (projection from behind the screen) on the left, front, and right walls and direct projection on the floor. The CAVETM system has been acquired by UNO through DoD property disposal, and it is available for use at UNO.

In the CAVETM, the adaptable, Crystal Eye glasses are used to produce a stereoscopic view, and an Ascension Flock of Birds tracking system is employed for head tracking of the user’s position and position of a wand in 3D space. Additionally, the wand is a modified PC joystick with an
interface box for control-button mapping. The wand is one of the primary means of interacting with the immersive virtual reality system.

In Maxwell’s DSVE, synthesized geometry is presented to the user as if it was suspended in mid-air, approximately one meter in front of the GROTTO’s front wall. This creates a “design volume” that exists roughly in the middle of the GROTTO. The significance of this display is that the user is not presented with a transformed 2D to 3D view of the objective model, but a true 3D representation. The immersive stereoscopic view is, therefore, different from a traditional computer aided design (CAD) presentation at a workstation, because the real time, stereoscopic rendering provides an artificial sense of depth. With the immersive virtual environment, the user creates, “walks around,” and interacts with geometry displayed within this volume. The left wall of the GROTTO contains a calculator and an open computer window, allowing designer communication with the computer system. The right wall includes a tool cabinet that represents the “tools” used to interact with the virtual environment. Figures 3.1 and 3.2 show gestural interaction in Maxwell’s DSVE.

It is argued that with the DSVE, the use of gestures and hand movements provide a more natural interface with the design environment, allowing a more rapid, efficient set of actions to create and/or modify product geometry. The Gestural interfaces are used to provide an improved means of interaction. The DSVE also provides stereoscopic perspective vision, an advantage in performing engineering design tasks.
In the work reported by Satter and Butler (2003), Butler et al. (2003 and 2004), and Satter et al. (2004) the workbench was employed as a test platform for interface evaluation. The goal was to use competitive usability to evaluate differences in interface methods. For example, conventional CAD systems with keyboard and mouse are assessed against gestural interfaces with wand or with gloves and voice. These interfaces are shown in Figures 3.3 and 3.4. In Satter’s test regimen (2005), three different Benchmarks were used to evaluate navigation, detection and repair of design errors, and spatial awareness. The focus by Satter and Butler (2003) has been on the shipbuilding industry. Their effort has included Northrop Grumman Ship Systems, the largest manufacturing employer in Louisiana. However, in the proposed research, the investigation is more general in nature. In this dissertation the focus is broadened to encompass environments that are not specific to shipbuilding, such as industrial plants, civil works, and buildings.
Figure 3.3 – Wand Gestural Interface with the Adaptable

Figure 3.4 – Voice & Glove Gestural Interface with the Adaptable
3.2 User Centered Development

Over the last several years, the Naval Research Laboratory (NRL) has developed a battlefield visualization system using a User-Centered Design and Evaluation process. This process involves evaluating and improving the user interface for a tactical display on a responsive workbench that employed immersive technology almost identical to Maxwell’s Design Synthesis in a Virtual Environment system that was GROTTO based virtual environment system developed at NRL. This battlefield visualization system included an extensive effort for user interface design. As a part of that project, Gabbard, Hix, and Swan (1999) found that “comparatively little effort has gone into user interaction components of VEs… Subsequent work by Satter and Butler (2005) has been a part of a research trend to reverse these circumstances.

The methods used at NRL and advocated by Hix et al. (1999) include a three-step evaluation process. This three-step process involves 1) heuristic evaluation, 2) formative evaluation, and 3) summative evaluation. Each of the usability evaluations provides input to the next stage of evaluation, and each evaluation technique has a successively higher evaluation cost, as shown in Figure 5. The reported research project uses these successive evaluation types because of the successful prior implementation at NRL, thereby avoiding the pitfalls of many virtual systems described by Gabbard, Hix, and Swan where, “many visually compelling VEs are difficult to use and thus unproductive.”

As the first process, an independent expert or experts undertake heuristic evaluation of the user interface. The interface is examined in a two-pass approach [Nielsen, 1994]. For the two-pass evaluation each expert first gains a general understanding of the flow of
interaction and then repeats the review process to identify specific interaction components and conflicts [Gabbard, Hix, & Swan, 1999].

The second evaluation process is formative evaluation in which users are employed to evaluate the virtual environment interface. There is a usability specialist to proctor the process in which users perform tasks as evaluators collect data. The formative evaluation includes five steps that are conducted iteratively. These steps include: 1) development of task scenarios, 2) representative users perform the scenarios, 3) evaluations collect data, 4) VE designers and evaluations suggest improvements, and 5) VE designers and evaluators refine task scenarios. Typically, critical incidents occur in which quantitative and qualitative data is developed. The quantitative data shows that a problem occurred, and the qualitative data tends to indicate where the problem occurred.

As the most expensive form of evaluation, summative evaluation is used to statistically compare final forms of VE design. Scenarios, developed during the formative evaluation process, are refined for use in evaluating final virtual environment interfaces. The results are a specific and quantitative answer to questions regarding which interface performs better [Gabbard, Hix, & Swan, 1999].

It is apparent that the science of user centered interface design for virtual environments (VE) is developing. As a part of this development Hix and Gabbard have authored a taxonomy for usability characteristics as their contribution to the science of VE interface design. They argue that the day of “let’s build it and see what happens” is over, and future research should be focused on the use of user centered design methods described in this research.
Others that have studied interface design for virtual environments have argued for user-centered design of the interfaces (Padlke, 2000). Theoretical models have also been developed to support design of virtual environments (Kaur et al., 1999), and competitive studies have been conducted (Evans et al., 1999) similar to the competitive study offered in this research. However, all possible manipulation tasks cannot be assessed. It is, therefore, important to identify a small and representative set of tasks from which to assess the system design (Poupyrev et al., 1997), and this argues for the methods developed by (Gabbard, Hix, and Swan, 1999) in which scenarios are constructed for typical cases as part of their user centered methods.
3.3 Gestural, Voice, and Multimodal Virtual Interfaces

In immersive virtual reality systems, interaction with the three-dimensional synthetic world was especially crucial, and the focus was initially on new devices, then research into higher-level techniques for interaction followed. One such development was the use of Three-Dimensional Widgets (Conner et al., 1992) in which an augmented transition network (ATN) was used for management of geometry. Handles and snapping was also employed. In a system developed by Deering (1995), a wand was used to sketch three-dimensional input, and the SKETCH system was developed at Brown University (Zeleznik et al., 1996) for processing wand strokes in two dimensions. This approach was extended by Bimber (1999). Mine (1997), in his explanation of the ISSAC system, discussed gestural interfaces in which there is action at a distance, worlds in miniature, and menus interaction.

In order to use gestural input, the gestural motion must be recognized. To aid in gesture recognition, Wexelblat (1995) reported on a hierarchical analysis of hand motion using features to capture gestural input. Other approaches used a rule-based interpretation (Lee et al., 1998) or neural network based systems (Nishino et al., 1997). In contrast, Kallmann and Thalmann (1999), employed smart objects that behaved in a manner consistent with the physical constraints normally associated with an object. The subject of gestural input recognition was surveyed by Hand (1997).

In addition to gestural interface design, solid free form models have been generated by a haptic interface (providing force feedback to the user) as demonstrated by Leu et al.
Other uses of haptic input are reported by Volkov and Vance (2001) in which force feedback is used for the design of an automotive hand brake. Use of a 3D track ball for solid modeling is reported by Stork and Maidhof (1997), wherein features and topological context are used to assist with the modeling effort. It should be noted that their techniques were not applied in a fully immersive environment. Other systems employ voice for interaction with the virtual environment. For example, Gao et al. (2000) used voice commands and constraint propagation. Other systems used both gestural input and voice commands (Bolt and Herranz, 1992; Chu et al., 1998).

As an important adjunct to the development of virtual environment interface paradigms, the use of testbeds has been suggested. Poupyrev et al. (1997) argue for in depth experimental studies because “there is still insufficient understanding of the essential characteristics and parameters of VR manipulation.” Their VRM A T system allows in depth studies, but it does not consider voice, gesture, and gaze. In a similar testbed, Boman et al. (1999) advocates evaluation of interaction techniques based on detailed empirical studies. They contrast this evaluation process with usability studies, and it is clear that information from a testbed study would be helpful in the development of virtual environments. However, the use of a usability approach for interface design focuses on generation of a system for a specific purpose, such as Design Synthesis. Therefore, testbed studies are useful, but do not seem to fit the current research due to their limited scope.

In related work, some researchers have focused on collaborative virtual environments (or CVEs). These collaborative virtual environments allow individuals at separate locations...
to interact using virtual reality systems. The MASSIVE system (Greenhalgh and Benford, 1995) was designed to allow multiple users to communicate using an arbitrary array of audio, graphics, and text over local and wide area networks. MASSIVE was further studied for user interaction by Tromp and Snowden (1997). In other work, four users communicated and interacted through four avatars (computer generated model of human body or body parts for visual realism) while navigating through a shopping mall (Schwartz et al, 1998). In another demonstration, Kiyokawa et al. (2000) provided interaction between two users with head mounted displays (HMDs) and reactive objects. Collaboration in the design of ceramic artifacts was shown by Nishino et al. (1999) where the artifacts are represented through parameters, and in the DVDS system reported by Arangarasan and Gadh (2000), collaboration is maintained through data sharing where data is maintained in a commercial CAD system. Through experience with the CVEs some insight is obtained regarding the development of virtual environments for engineering design.
3.4 General Experimental Procedure

The procedures in this study are modeled directly from procedures used in a similar study at the Naval Research Laboratory (NRL), and used previously at UNO. Dr. Edward Swan collaborated with Dr. Deborah Hix and Dr. Joseph Gabbard of Virginia Tech (VPI) in a project to improve the user interface for a Marine Corps tactical display system that employed a virtual reality, responsive workbench. Our methods are derived from their efforts.

The previous NRL and VPI project used an Informed Consent Form to provide information to the subjects of the circumstances, conditions, and risk of the interface testing procedures. Each subject signed this consent form. Procedures regarding the use of consent forms in this study are discussed in section 8 below. A copy of the UNO Consent Form is provided in Appendix 2.

In the NRL and VPI project, each study participant was also provided with a one-page description of the testing procedures and processes. The subjects were given some time to “play around” with the Virtual Environment for familiarity before they undertook a specific exercise. The subjects were then given a list of tasks to be completed, and the subjects were told to work quickly without sacrificing accuracy. The researchers timed each task, counted errors, recorded activity on an evaluation sheet, and completed videotaping of the subject’s performance. At the end of each task, the subject was asked to sketch the geography in the problem. This provided an additional test of the interface’s capacity to communicate effectively with the human user.
Similar methods have been used successfully here at UNO for an existing project approved by UNO’s Human Subjects Institutional Review Board (IRB). All subjects (testers) were provided with a consent form, the experiments in competitive usability were explained, time provided for familiarity, and the experiments conducted with NO NEGATIVE EFFECTS. This study was approved by the UNO Institutional Review Board (IRB), and was conducted successfully. A copy of the approval form is provided as Appendix 3.

Usability Benchmark studies involve tasks and metrics for specific aspects of the experiments with human subjects. These Benchmark studies involve timing the movements of a user as he or she navigates through a congested space, evaluating the number of errors detected in a specific scene in a given amount of time, determining the amount of time needed to correct a specific number of errors in a scene, and measuring the perceived location of objects (such as brightly colored and specifically shaped markers or icons) in comparison to the known placement of those objects. In this research, this includes what is termed Benchmark 1, 2, 3 and 4, with testing using two different interface paradigms: (a) CAVETM with Wand Interface, and (b) Traditional Computer Aided Design Workstation with Keyboard and Mouse.

Benchmark 1 involves a navigation scenario that required users to start at the entryway into the virtual factory space as a complex and crowded environment; then, locate pre-defined equipment within the space and, return to the entryway utilizing the interface tools of each of the environments under test. The measure is elapsed time (in seconds) to perform the navigation task. Each user located, identified, and catalogued (noted the equipment and position) four distinct parts within the space.
Benchmark 2 is similar to Benchmark 1 in that elapsed time from entry into the factory space, movement within the space, and return to the starting point is the prime measurement. However, with Benchmark 2, errors in construction of the space are presented (two errors per test), and the identification and correction of those errors (unknown to participants) are required by the user.

Benchmark 3 is designed to measure the ability of the environments to aid user spatial awareness. Users are presented with a totally foreign space into which two readily recognizable icons are randomly placed. Using each of the interface environments, users are asked to navigate through the space to locate each marker or icon. While the elapsed time required to locate each icon is recorded, the primary measure provided by this Benchmark is the users placement offset from the actual location of the icon. The user’s perceived placement of each icon is recorded on a 2D diagram of the space and the offset from the actual placement is measured in millimeters.

Benchmark 4 is designed to measure the ability of the participant to find as many of ten pre-set errors (unknown to participants) as possible in a span of four minutes.

Each hands-on test was followed with a questionnaire/survey designed to elicit the subjective evaluation of the interface environment from each user. Satter’s survey instrument is found in Appendix 4.

Each test is represented by an execution of scenarios with Benchmark 1, 2, 3 or 4 and took approximately 15 to 20 minutes for each subject. In this study, each Benchmark was administered to each subject three times to determine the effects of learning and familiarity with the interfaces, and each Benchmark was conducted with two different interfaces. Each subject executed four Benchmarks with two different interfaces,
repeated three times for a total of twenty-four 15 to 20 minute tests. In order to gauge the effects of retain-ability of learned activity, each test was conducted at different times on different days.

This study on virtual environments and interfaces involves testing of interaction for gestural interface usability and effectiveness of the interface software and hardware. Accordingly, the survey forms are adapted for tasks that must be tested in the context of a virtual environments system. With this in mind, the format and procedures of the earlier studies are retained, but the subject matter of the forms and tasks are modified to fit the current UNO and Dillard based investigation with gestural interfaces.

C. DATA COLLECTION

Participants

This research utilized a population of thirty students from Dillard University who volunteered to participate as test subjects (testers).

Record Keeping and Recordings

Survey information obtained and used in this study is retained for study purposes under lock and key. It is published only in the aggregate without disclosing the identities of the participants.

D. FUNDING SOURCE

There is no funding for this research and the test subjects have participated voluntarily.

E. RISKS TO PARTICIPANTS

Anonymity, Confidentiality and Handling of Data
The names and other personal information of the subjects in this study will not be disclosed, in publications and other documents reporting on research findings. Only a number with recorded results identifies each participant, and records that relate the numbers to an individual’s identity are not disclosed outside of the study team.

This study is conducted at the UNO – Northrop Grumman Maritime Technology Center of Excellence at 5100 River Road, Avondale, LA 70094. The building containing the experimental equipment, laboratory space, and offices of the investigators is under tight security. In order to be admitted to the building, an individual must either have a security badge or be admitted by someone already in the building. A security camera views everyone that enters the building through the UNO entrance, and security guards are positioned at several locations within the building. The security guards are on duty 24 hours per day, 7 days per week.

In addition to the UNO – Northrop Grumman Maritime Technology Center of Excellence security procedures, all experimental equipment, laboratory space, and offices are maintained behind locked doors. Typically, these spaces are either occupied or locked, and a receptionist is usually on duty during normal working hours when many of the experiments may be conducted. Data on the subjects, their participation, and the research results shall be maintained in these secure spaces. Further, this data is segregated from other documents, and maintained as information confidential to the participants.

**Impact on Student Academic Records**

In order to allow students to choose to participate or not participate in these studies of traditional CAD systems compared with virtual environments, the students had an alternative to participation, in the event that they choose to not participate. This alternate
assignment consisted of a brief term paper or the writing of a short computer program that would involve human computer interface activity or to be a test subject in this research. The goal of this provision is to allow students to have a comfortable set of choices between participation as a test subject or an alternative that does not involve participation as a test subject.

**Risk and Risk Mitigation**

This study is conducted using immersive virtual computer environments or other artificial environments, as well as traditional computer aided design workstations. Individuals with neural disorders such as epilepsy have experienced problems remaining in a virtual environment, but the symptoms tend to decline rapidly when the subject is removed from the environment.

A check list, reviewed by Kimberly Rask, MD, PhD who is an Associate Professor of Medicine at Emory University is employed to detect potential medical problems (Rask, 2004). This check list is provided in Appendix 5. Any test subjects that reply “Yes” to any of the questions on the list was excluded from participation in this study.

Some people experience discomfort and motion sickness in a virtual environment. The symptoms typically exist only while the participant is in the virtual environment, and the symptoms disappear shortly after the subject is removed from the virtual environment. If an individual has discomfort or motion sickness difficulty with the virtual environment, they were removed as a subject in the study.
Recruitment Procedures

Additionally, subjects were limited to individuals over 18 years of age. The test population was recruited from Dillard University. The budget had no funds to pay the students for participation. We used flyers for advertisement. A copy of a Dillard University flyer is provided in Appendix 6. All participants were volunteers providing evidence of written consent to participate.

F. INFORMED CONSENT

Prior to obtaining informed, written consent, test subjects were provided with a short overview briefing on the goals of the study, the methods for obtaining data, and the nature of a typical test session. The subjects were provided with two copies of the consent form. One copy was signed by the participant and retained by the research team. The participants were also provided with a second copy for their use.

Subjects in this study were informed about the nature, procedures, and content of the study through recruiting materials and as a part of the informed consent procedures. Questions about procedures and methods were answered at that time and at any time that a subject expressed a question about methods and procedures.

A copy of the consent form was provided to the subjects. This form contained the contact information for the Principal Investigator and Co-Investigator including address, email address, and phone numbers.
G. DATA USE

The data collected in this study is used in the aggregate to reach conclusions about the benefits and limitations of traditional CAD system when compared to virtual environment using a wand. In all forms of publications, the subjects are identified only by letters and numbers that cannot be linked to subject identity without access to project files maintained under lock and key and not published. The anticipated forms of publications include conference papers, dissertation and thesis publication, journal paper publication, presentations at conferences, presentations to sponsors and other interested parties, classroom presentation, and presentations internal to both universities.
4 BENCHMARK 1 (NAVIGATION)

4.1 Description

The Benchmark 1 scenario was designed to test the user’s ability to utilize the two environments/interfaces to navigate through the study space locating each of four distinct items/parts within the space. The common measure recorded was simply the elapsed time to navigate the space (from a common starting point), locate each required item/part, and return to the starting point. Each user performed this Benchmark three times in each of the two environments. The analysis of the final pass results of these Benchmark 1 tests by the users is presented in the following sections. The other pass results are given in Appendix A. Pass 3 results represent each user’s final exposure to each environment within each scenario (Benchmark). Therefore, pass 3 results show the user’s ability to perform the required tasks. Each environment/interface (Non Stereo workstation and Stereo CAVE™) is represented in a distinct chart.

4.2 Pass-to-Pass Improvements in Elapsed Times

Figure 4.1 shows user elapsed times for pass 3 of the navigation Benchmark tests in the two environments. A preliminary investigation of the chart data show that the users performed navigation tasks faster using the CAVE™ stereoscopic (wand) interface over the non-stereoscopic environment.
4.3 B1-Pass-to-Pass Comparison of Elapsed Times Analysis:

Table 4.1 presents the improvements in navigation times for users with each successive exposure to each of the two test environments. For B1, the elapsed timings improved for both CAVE™ and Workstation from Pass-to-Pass. Comparing CAVE™ and Workstation, the elapsed timings appear to have improved more for CAVE™ with a higher percentage from Pass to Pass than Workstation. Note that there appears to be 34% improvement in CAVE™ from pass 1 to pass 3 against only 24% improvement in workstation from pass 1 to pass 3. This means that stereoscopic environment resulted in sharper decreases in navigation times than for non-stereoscopic environment.
4.4 Elapsed Times Detailed Statistical Analysis

All statistical analyses of the test data were performed using Number Cruncher Statistical Systems (NCSS) software. Considerable assistance in interpreting the results was gained from NCSS. NCSS software provides both descriptive statistics on the data and a T-test that aids in selecting the proper tests based on the distribution of the test data.

The descriptive statistics tests are performed to determine if the sets of environment data are normally distributed (Gaussian distribution). Such testing (Normality Testing) quantifies and reports the discrepancy between the distribution of the data and the ideal Gaussian distribution. NCSS uses the Kolmogorov-Smirnov (K.S) test for calculating this value; the KS statistic. A larger KS statistic value denotes a higher discrepancy and is used to compute a traditional statistic P value. The results presented here are based on the means and standard deviations of each set of Benchmark, environment, and test pass sample results.

The P value from the normality test answers the question: “In a random sample from a Gaussian distribution, what is the probability (P value) of obtaining a sample that deviates as much from a Gaussian distribution (or more so) than the given sample. Stated differently, the P value answers the question: If the population is Gaussian, what is the chance (as measured by probability) that a randomly selected sample of this size would
have a KS statistic as smaller giving a higher $P > 0.10$ value for a normal distribution?" [NCSS, 2004]

Since the sample sizes for this study are relatively small (30 users), a large $P$ value only means that the data is consistent with a Gaussian (normal) population. This does not exclude the possibility of a non-Gaussian population.

There are two hypotheses in this case. The first is the null hypothesis ($H_0$) that states that there is no difference between the two environments. The second is the alternative hypothesis ($H_a$) that states that the environment with the smaller (faster) elapsed time is “better”. (In this particular case, CAVE$^{TM}$ has a lower mean than workstation; meaning that the user had faster times for CAVE$^{TM}$ than Workstation.)

In either parametric (normal distribution) or nonparametric testing, it is sufficient to test the null hypothesis of equal means for normal distribution and the null hypothesis of equal medians for non-normal distribution:

Null Hypothesis: ($H_0$): $m_1 - m_2 = 0$.

Should $H_0$ prove true, the means of the navigation times (or any other variable) for the two environments being compared are equal (at the 90% confidence level) and thus there is no statistical difference in the compared environments. However, should the test fail, statistical credence can be given to the alternative hypothesis:

Alternative Hypothesis: ($H_a$): $m_1 - m_2 \neq 0$.

$H_a$ true indicates that there is a 90% confidence that the means are not equal and thus navigation in the two environments are statistically different and by analysis, the
environment producing lower elapsed times is “better.” This constitutes a statistically significant proof of different means for the data.

4.5 Mann-Whitney Test

The Mann-Whitney Test is used when there is a non-normal distribution and the normality test fails or when the data is non-variant. It uses the median to compare differences between the two groups.

The median is used for non-normal comparisons because the median is unaffected by the non-normal distribution of the data. The mean, since its calculation involves all the data, is skewed by the non-normality of the data. Therefore, the mean is an unreliable measure to use in tests. Hence the median is used instead of the mean (T-test).

4.6 Pass 3 Statistics

Table 4.2 (Benchmark 1 Pass 3 Elapsed Timings/B1p3Tim) presents the descriptive statistics test results (normality testing) of the K.S. test followed by the results of Levene’s test for equal variance of the data. The null hypothesis (H₀) and alternative hypothesis (H₁) discussed above applies here.

<table>
<thead>
<tr>
<th>B1P3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>219.3</td>
<td>48.9</td>
<td>134</td>
<td>354.00</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>22%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>424.4</td>
<td>93.4</td>
<td>260</td>
<td>735.00</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>22%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>Equal Var?</td>
</tr>
<tr>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>6.77</td>
</tr>
</tbody>
</table>

Mann-Whitney test used because of unequal variance

**Table 4.2: B1p3Tim Elapsed Times**

For Table 4.2, the K.S. test is used to test for normality of data. Since the P value is greater than 0.1, the data are normal. Levene’s test is used to test for equal variance.
Since the P value is less than 0.1 the data have unequal variances. In this case, since the data has unequal variance, the Mann-Whitney test is used. With the Mann-Whitney test, the P value is less than 0.1. That indicates that the medians are unequal for CAVE™ and workstation. Examination of these results shows that for the two environments, the differences are statistically significant. The conclusion then is that at the 90% confidence level, there is significant evidence to support the alternative hypothesis ($H_a$). Thus, since the CAVE™ demonstrates shorter elapsed find/repair times, this environment is statistically “better” than non-stereoscopic workstation environment for Benchmark 1 during pass 3 elapsed timings.

Figure 4.1 provides a summary of Benchmark 1; pass 3 elapsed navigation times for all user timings in the two environments under test. As shown in the chart, CAVE™ environment resulted in somewhat lower navigation times. It should also be noted that as a group, all users performed better using the stereoscopic environment (CAVE™ and Wand) over the non-stereoscopic environment (workstation).

### 4.7 User Subjective Overall Environment Ratings

After completion of each pass of each Benchmark test in each environment users provided their subjective views of their experience by completing the 22-question Usability Survey (see Appendix 4) rating the environment on a scale of 1 to 5 (very poor to very good).

The questions were grouped into 4 areas (navigation, locating, movement, and general). Following is a presentation of user overall impressions ratings of the interfaces for performing Benchmark 1 tasks (navigation) at the completion of the 3rd pass as a
representation of user’s final evaluations of each interface. The results of the impressions ratings for all other passes are presented in Appendix A.

As discussed above, each user was asked to rate his/her experience via the Usability Survey at the completion of each pass of each Benchmark test. The Figure 4.2 (Benchmark 1 Pass 3 Overall Impressions Ratings/B1p3Ovr) shown below presents the overall impressions ratings of the users at the completion of the 3rd pass of the Benchmark 1 scenario. As such, this represents each user’s final impression of the navigational capabilities of each environment.

A further examination of the results detailed in Figure 4.2 show that upon completion of the Benchmark tests, users preferred the stereoscopic wand interface over traditional CAD workstation interface.

![Figure 4.2: B1p3Ovr Overall Impressions Ratings](image-url)
Table 4.3: B1p3Ovr Overall Impressions Ratings

For Table 4.3, the K.S. test is used to test for normality of data. Since the P value is less than 0.1 for CAVE™, the data are not normal. Levene’s test is used to test for equal variance; since the P value is less than 0.1 the data have unequal variance. In this case, since the data is not normal, the Mann-Whitney test is used. With the Mann-Whitney test, P value is greater than 0.1, which indicates that the medians are equal for CAVE™ and workstation. Examination of these results shows that for the two environments, the differences are not statistically significant. The conclusion then is that at the 90% confidence level, there is significant evidence to support the null hypothesis (H₀). This proves that none of the two environments are statistically better than each other for Benchmark1 pass3 overall impressions subjective ratings.

Thus, since the CAVE™ demonstrates shorter elapsed find/repair times, this environment is statistically “better” than non-stereoscopic workstation environment for Benchmark 1 during pass 3 elapsed timings.
4.8 B1-Pass to Pass Comparison of Overall Impressions Ratings Analysis:

Table 4.4 shows the pass-to-pass improvements in user overall impression ratings for each of the environments. Note that with each successive exposure (pass-to-pass) the user’s overall impressions of the interfaces improved. Examination of the pass-to-pass analysis of improvements noted in Table 4.4 shows that for Benchmark 1 overall impressions subjective ratings, the ratings improved for both CAVE™ and Workstation from pass-to-pass. In comparing the CAVE™ and workstation, the ratings appear to have improved more for the CAVE™ environment with a higher percentage from pass-to-pass than for the workstation.

<table>
<thead>
<tr>
<th></th>
<th>Pass1 to Pass2</th>
<th>Pass2 to Pass3</th>
<th>Pass1 to Pass3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>-0.65</td>
<td>-0.3</td>
<td>-0.95</td>
</tr>
<tr>
<td>%</td>
<td>-19%</td>
<td>-7%</td>
<td>-27%</td>
</tr>
<tr>
<td>W/S</td>
<td>-0.23</td>
<td>-0.69</td>
<td>-0.92</td>
</tr>
<tr>
<td>%</td>
<td>-7%</td>
<td>-19%</td>
<td>-27%</td>
</tr>
</tbody>
</table>

Table 4.4: B1- pass-to-pass Comparison of Overall Impressions Ratings

The negative values in Table 4.4 show that pass 1 ratings were lower than pass 2 and pass 2 ratings were lower than pass 3. This means that user’s preference improved from pass to pass. For example a value of -27% for the CAVE™ (pass 1 to pass 3) is calculated as (3.5-4.45)/3.5, where 3.5 and 4.45 represent the means of Benchmark 1 over impressions ratings for pass 1 and pass 3 respectively.
5 BENCHMARK 2 (FIND AND REPAIR MANIPULATION)

5.1 Description

Using the same virtual factory space as used for Benchmark 1, in Benchmark 2 users were required to navigate through the space looking for “errors” that had been injected into the design. Typical “errors” were a screen, turbine or fan, eyewash or conveyor belt, or cyclone separator all placed at a different place from their original place. Users were then required to “fix” the error. The “fix” required the user to utilize the interface (environment) under test (CAVE™, workstation), typically, re-positioning the part to a more suitable location/orientation. Elapsed times were noted for each activity. The elapsed time recorded was the time required to locate and identify the 1st error; the time to “fix” the 1st error; the time to locate and identify the 2nd error; the time to “fix” the 2nd error; and the time to return to the starting position within the space.

The find/repair exercise (Benchmark 2) was repeated three times (three passes) for each of the thirty users in each of the two environments under test and the User Survey was administered to each user after each pass in each environment. As with the Benchmark 1 testing, sequencing of the testers through the two environments was randomized so that not all of the users were testing the same interface at the same time.
5.2 Benchmark 2 – pass 3 Elapsed timing analysis:

Figure 5.1 (Benchmark 2 Pass 3 Elapsed Timings / B2p3Tim) presents a representation of the elapsed times required by users to perform a typical set of find/repair operations as defined in the Benchmark 2 scenarios. The results presented are for the last (3rd) execution of the test. All other results are presented in Appendix B. These times should, and do, represent the “best/fastest” execution times for the group. It should be noted that stereoscopic interface resulted in shorter execution times (as compared to the non-stereoscopic interface). This proves that CAVE™ is faster, efficient and better environment workstation.

![Benchmark 2 - Pass 3 Elapsed Times](image-url)

**Figure 5.1: B2p3Tim Elapsed Times**
5.3 Pass-to-Pass Comparison of Elapsed Times Analysis:

<table>
<thead>
<tr>
<th>B2 Pass to Pass Comparison</th>
<th>Pass1 to Pass2</th>
<th>Pass2 to Pass3</th>
<th>Pass1 to Pass3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>Diff</td>
<td>%</td>
<td>Diff</td>
</tr>
<tr>
<td>69.43</td>
<td>21%</td>
<td>66.6</td>
<td>26%</td>
</tr>
<tr>
<td>W/S</td>
<td>79.1</td>
<td>18%</td>
<td>76.9</td>
</tr>
</tbody>
</table>

Table 5.1: Pass-to-Pass Comparison of Elapsed Times

Table 5.1 presents the improvement in find/repair (manipulation) times for users with each successive exposure to each of the two test environments. Note that there appears to be about a 42% improvement in CAVE™ from pass 1 to pass 3 against 36% improvement in workstation from pass 1 to pass 3. This means that the stereoscopic interface appeared to produce reduced find/repair elapsed times over the non-stereoscopic interface.

5.4 Detailed Statistical Analysis

As described for the Benchmark 1 testing, all statistical analyses of the test data were performed using NCSS. The K.S. normality testing was performed on the Benchmark 2 results. Levene’s test was used to test for equal variance of the data. The null hypothesis ($H_0$) and alternative hypothesis ($H_a$) as discussed in section 4.4 for Benchmark 1 statistical analysis testing applies here (Benchmark 2) as well.

5.5 Pass 3 Statistical Analysis

Table 5.2 presents the descriptive statistics test results normality testing and variance test results of each Benchmark 2 pass 3 dataset by environment. All other results are presented in Appendix B. In this analysis, it is important to note the results of the Kolmogorov-Smirnov test (KS test statistic) for normal (Gaussian) distribution. In this
case, note that the pass 3 B2 datasets for the non-stereoscopic environment fail the KS statistic test for normal distribution of the data. Thus the NCSS software performs a nonparametric, Levene’s test to test for equal variance.

<table>
<thead>
<tr>
<th>B2P3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>187.5</td>
<td>42.7</td>
<td>111</td>
<td>260</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>22%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>272.6</td>
<td>59.3</td>
<td>173</td>
<td>408</td>
<td>&lt;0.10</td>
<td>No</td>
<td>22%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene’s Test</td>
<td>Mann-Whitney Test</td>
</tr>
<tr>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table 5.2: B2p3Tim Elapsed Times Statistics

For table 5.2, the K.S. test is used to test for normality of data. Since the P value is less than 0.1 for workstation, the data are not normal. Levene’s test is used to test for equal variance. With that stipulation, since the P value is greater than 0.1, the data have equal variance. Thus, since the data is not normal, Mann Whitney test is used. However, with a Mann Whitney test P value less than 0.1, which indicates that the medians are unequal for the CAVE™ – Workstation comparison. Examination of these results shows that for the two environments, the differences are statistically significant. The conclusion then is that at the 90% confidence level, there is significant evidence to support the alternative hypothesis (H_a). Thus, since the stereoscopic wand environment demonstrates shorter elapsed find/repair times, this environment is statistically “better” than non-stereoscopic workstation environment for Benchmark 2 during pass 3.
5.6 User Subjective Overall Environment Ratings

After completion of each pass of each Benchmark test in each environment, users provided their subjective views of their experience by completing the 22-question Usability Survey (see Appendix 4) rating the environment on a scale of 1 to 5 (very poor to very good). The questions were grouped into 4 areas (navigation, locating, movement, and general). Following is a presentation of user overall impressions ratings of the interfaces for performing Benchmark 2 tasks (find/repair) at the completion of the 3rd pass as a representation of user final evaluations of each interface. All other results are presented in Appendix B.

5.7 Benchmark 2 – pass 3 Overall Impressions ratings analysis:

As discussed above, each user was asked to rate his/her experience via the Usability Survey at the completion of each pass of each Benchmark test. Figure 5.2 (Benchmark 2 pass 3 Overall Impressions ratings /B2p3Ovr) presents the overall impressions ratings of the users at the completion of the 3rd pass of the Benchmark 2 scenario. As such, this represents each user’s final impression of the find/repair capabilities of each environment. For Benchmark 2 pass 3 overall impressions ratings, figure 5.2 shows that user’s preferred CAVE™ over workstation.
Figure 5.2: B2p3Ovr Overall Impressions Ratings

<table>
<thead>
<tr>
<th>B2OP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.45</td>
<td>0.45</td>
<td>4</td>
<td>4.95</td>
<td>&lt;0.10</td>
<td>No</td>
<td>10.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.28</td>
<td>0.18</td>
<td>3.95</td>
<td>4.75</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
</tbody>
</table>

Table 5.3: B2p3Ovr Overall Impressions Ratings

For Table 5.3, the K.S. test is used to test for normality of data. Since the P value is less than 0.1 for the CAVE™, the data are not normal and Levene’s test is used to test for equal variance. Since the P value is less than 0.1 the data have unequal variance. Furthermore, since the data are not normal, the Mann Whitney test is used. With a Mann Whitney test, P value greater than 0.1, which indicates that the medians are equal for the CAVE™ and workstation. Examination of these results shows that for the two environments, the differences are not statistically significant. The conclusion then is that at the 90% confidence level, there is significant evidence to support the null hypothesis.
(H₀). This proves that neither of the two environments is statistically better than each other for Benchmark 2 pass 3 overall impressions subjective ratings.

5.8 B2-Pass to Pass Comparison Overall Impressions Ratings Analysis:

Table 5.4 shows the pass-to-pass improvements in user overall impression ratings for each of the environments. Note that with each successive exposure (pass-to-pass) the user’s overall impressions of the interfaces improved. Examination of the pass-to-pass analysis of improvements noted in Table 5.4 shows that for Benchmark 2, overall impressions subjective ratings, the ratings improved for both CAVE™ and Workstation from pass-to-pass. Comparing CAVE™ and Workstation, the ratings appear to have improved more for CAVE™ with a higher percentage from pass-to-pass than Workstation. Hence, the CAVE™ environment is barely preferred over Workstation for B2 Overall impressions subjective ratings.

<table>
<thead>
<tr>
<th></th>
<th>Pass1 to Pass2</th>
<th>Pass2 to Pass 3</th>
<th>Pass1 to Pass3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>-0.7</td>
<td>-0.18</td>
<td>-0.88</td>
</tr>
<tr>
<td>W/S</td>
<td>-0.26</td>
<td>-0.65</td>
<td>-0.91</td>
</tr>
</tbody>
</table>

Table 5.4: B2-Pass to Pass Comparison Overall Impressions Ratings

The negative values in Table 5.4 show that pass 1 ratings were lower than pass 2 and pass 2 ratings were lower than pass 3. This proves that user’s subjective ratings improved from pass to pass. For example a value of -27% for Workstation (pass 1 to pass 3) is calculated as (3.37-4.28)/3.37, where 3.37 and 4.28 represent the means of Benchmark 2 overall impressions ratings for pass 1 and pass 3 respectively.
6 BENCHMARK 3 (SPATIAL AWARENESS)

6.1 Description

In order to evaluate the ability of each environment/interface to aid users in their awareness of a design space, a unique space, totally unknown to the users, was created. For the test, the space created was a virtual factory space and a machine shop. Into this space the test administrators were able to inject an obelisk icon (an elongated, gray-white, pyramid topped by a sphere as shown below in figure 6.0) that is not normally found in any factory space. Two such icons were randomly placed into the new space for each pass of the test. From a common starting point, users were required to navigate through the space looking for the icons within the space. The time required each user to locate each icon was recorded and the users were asked to note the location for each (placement within the space).
Upon completion of the test each user was shown a 2-dimensional, 8.5” x 11”, plan-view of the space and asked to note the placement of each of the two icons. The test administrators then recorded the offset (in mm) between user placement and the actual location of the icons.

This exercise (Benchmark 3) was repeated in each of the 2 environments and the User Survey administered to each user after each pass in each environment. As with the other Benchmark testing, sequencing of the testers through the two environments was randomized so that not all of the users were testing the same interface in the same order. This randomization was used to eliminate bias in the testing.
6.2 Benchmark 3, Pass 3, Part 1 & 2 Placement Offsets Analysis:

Following is a presentation of the Benchmark 3, pass 3, part 1 and part 2 placement offsets for all the users. Pass 3 results are presented here as representative of user best-final case results. All other results are presented in Appendix C.

Figure 6.1 (Benchmark 3 pass 3 Icon 1 Offsets / B3p3-1off) presents user placement of the first icon within the new space. The results clearly indicate a higher spatial awareness using the stereoscopic CAVE™ environment. Using the stereoscopic interface, users on average located the icon within 11 mm of its actual location. User’s placement of the icon using the workstation non-stereoscopic environments was within 12.83 mm of its actual location.

Inspection of the standard deviation values of table 6.1 for the location of icon 1 shows a high variance in offset for the stereoscopic interface and shows low variance for the non-stereoscopic interface. This is an indication of the consistency of the non-stereoscopic method in spatial recognition efforts. Users were able to locate the icons better in workstation (2-dimensional non-stereoscopic environment) on a 2-dimensional, 8.5” x 11” paper than in a CAVE™.
Figure 6.1: B3p3-1off Pass 3-Icon 1 Offsets

Table 6.1: B3p3-1off Pass 3-Icon 1 Offsets

Figure 6.2 (Benchmark 3 pass 3 Icon 2 Offsets / B3p3-2off) presents user placement of the second icon within the new space. The results clearly indicate a higher spatial awareness using the stereoscopic the CAVE™ environment. Using the stereoscopic interface users, on average, located the icon within 7.77 mm of its actual location. User placement of the icon using the workstation non-stereoscopic environments was within 13.8 mm of its actual location.
Inspection of the standard deviation values of table 6.2 for the location of icon 2 shows a high variance in offset for the non-stereoscopic interface and shows low variance for the stereoscopic interface. This is an indication of the consistency of the stereoscopic method in spatial recognition efforts. Users were able to locate the icons much better in the CAVE™ environment than in a workstation. This proves that users performed better after practice in the CAVE™ environment proving the significance usability analysis.

![Benchmark 3 - Pass 3 - Part 2 Offsets](image)

**Figure 6.2: B3p3-2off Pass 3-Icon 2 Offsets**
6.3 Detailed Statistical Analysis

The following sections present a detailed statistical analysis of the Benchmark 3 results of the user group in a manner similar to the previous Benchmarks. As discussed above, the NCSS software package was used to perform each analysis. Each set of user icon 2 placement offsets is first examined to determine if the data is normally distributed (Gaussian distribution) using the KS statistic. The descriptive statistics test results are presented in tabular form followed by the results of Levene’s test for equal variance of the data. The null hypothesis (H₀) and alternative hypothesis (Hₐ) discussed for Benchmark 1 statistical analysis testing applies here (Benchmark 3) as well.

6.4 Benchmark 3 Pass 3 Statistics

Benchmark 3, pass 3, icon 2 offsets represent each user’s view of the placement of the required device in a foreign space. As such, the results of this pass/icon placement represent a reasonable characterization of the user’s spatial awareness within each environment.

6.5 B3p3-2off – Benchmark 3 Pass 3 Descriptive Statistics

Table 6.2 presents the results of the descriptive statistics analysis of user’s pass 3 location of icon 2 in the test environment. All other results are presented in Appendix C. The K.S. test is used to test for normality of data. Since the P value is less than 0.1, the data is not normal. Next Levene’s test is then applied to test for equal variance. Since the P value is greater than 0.1 the data has equal variance. Since the data is not normal, Mann Whitney test is used. A Mann Whitney test P value less than 0.1 indicates that medians are
unequal for the CAVE™ and workstation. Examination of these results shows that for the two environments, the differences are statistically significant. The conclusion then is that at the 90% confidence level, there is significant evidence to support the alternative hypothesis (Hₐ). Thus, since the stereoscopic wand environment demonstrates shorter offset distances, the CAVE™ environment is statistically “better” than non-stereoscopic workstation environment for Benchmark 3 during pass 3 for Icon 2 placements.

<table>
<thead>
<tr>
<th>B3Part2P3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>7.77</td>
<td>7.16</td>
<td>0</td>
<td>28</td>
<td>&lt;0.10</td>
<td>No</td>
<td>92%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>13.8</td>
<td>16.53</td>
<td>0</td>
<td>91</td>
<td>&lt;0.10</td>
<td>No</td>
<td>120%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene’s Test</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>F-Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
</tr>
</tbody>
</table>

*Table 6.2: B3p3-2off Pass 3-Icon 2 Offsets*
6.6 Benchmark 3 Pass 3 Overall Impressions Ratings Analysis:

Figure 6.3: B3p3Ovr Pass 3 Overall Impressions Ratings

Figure 6.3 (Benchmark 3 pass 3 Overall Impressions Ratings / B3p3Ovr) graphically presents comparisons of the Benchmark 3 (spatial awareness) pass 3 overall ratings of the two environments. Inspection of the average ratings shows that users preferred the stereoscopic environment (CAVE™) over the non-stereoscopic environment (workstation).
6.7 Detailed Statistical Analysis

The following sections present a detailed statistical analysis of user overall impressions ratings of the two test environments following their 3rd and final pass of the Benchmark 3 scenario. All other results are presented in Appendix C. The statistical analysis of these ratings provides insight into the final opinions of the users. As discussed above, the NCSS software package was used to perform each analysis. Each set of user overall impressions ratings is first examined to determine if the data are normally distributed (Gaussian distribution) using the KS statistic. The descriptive statistics test results are presented in tabular form followed by the results of Levene’s test for equal variance of the data. The null hypothesis (H₀) and alternative hypothesis (H₁) discussed for Benchmark 1 statistical analysis testing applies here (Benchmark 3) as well.

6.8 Benchmark 3 Pass 3 Overall Impressions Ratings Statistics

As noted, Benchmark 3, pass 3, overall impressions ratings represent each user’s view of the placement of the required device in a foreign space. As such, these ratings represent a reasonable characterization of the user’s overall impressions of the interfaces after the use of each to determine his/her spatial awareness of a previously unknown environment.

Table 6.3 presents the results of the descriptive statistics analysis of user’s Benchmark 3 pass 3 overall impressions of the interface. The K.S. test is used to test for normality of data. Since the P value is less than 0.1 for the CAVE\textsuperscript{TM}, the data are not normal. Levene’s test is used to test for equal variance and since the P value is greater than 0.1 the data have equal variance. Since the data is not normal, Mann Whitney test is used. With the Mann-Whitney test P value less than 0.1, which indicates that the medians are
unequal for the CAVE™ and workstation. Examination of these results shows that for the two environments, the differences are statistically significant. The conclusion then is that at the 90% confidence level, there is significant evidence to support the alternative hypothesis (H_a). This proves that the CAVE™ environment is preferred over workstation for Benchmark 3 pass 3 overall impressions subjective ratings.

<table>
<thead>
<tr>
<th>B3OP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.69</td>
<td>0.2</td>
<td>4.2</td>
<td>4.90</td>
<td>&lt;0.10</td>
<td>No</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.33</td>
<td>0.16</td>
<td>4</td>
<td>4.75</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
</tbody>
</table>

**Table 6.3: B3p3Ovr Pass 3 Overall Impressions Ratings Descriptive Statistics**
6.9 Benchmark 3 Pass-to-Pass Comparison Analysis:

Table 6.4: B3I1 pass-to-pass Comparison of Offset distances

Table 6.4 presents pass-to-pass comparison of Benchmark 3 part 1/Icon 1 offsets. The positive values in table 6.4 prove that pass 1 offsets were greater than pass 2 and pass 2 offsets were greater than pass 3. This proves that user’s placement of the icon on the paper improved from pass-to-pass with respect to the icon’s exact location in the two test environments. For example a value of 57% for Workstation (pass 1 to pass 3) is calculated as (30.1-12.83)/30.1, where 30.1 and 12.83 represent the means of Benchmark 3 part 1/Icon 1 offsets for pass 1 and pass 3 respectively. From table 6.4 one can conclude that user’s showed more improvement from pass to pass in workstation than in CAVE™. This is due to the fact that users were able to place the icons better in workstation (2-dimensional non-stereoscopic environment) on a 2-dimensional, 8.5” x 11” paper than in a CAVE™.

Table 6.5: B3I2 pass-to-pass Comparison of Offset distances

Table 6.5 (Benchmark 3 Icon 2 or part 2 pass-to-pass comparison / B3I2) presents pass-to-pass comparison of Benchmark 3 part 2/Icon 2 offsets. The positive values in table 6.5
prove that pass 1 offsets were greater than pass 2 and pass 2 offsets were greater than pass 3. For example a value of 51% for Workstation (pass 1 to pass 3) is calculated as 
\[(28.53-13.8)/28.53,\] where 28.53 and 13.8 represent the means of Benchmark 3 part 2/Icon 2 offsets for pass 1 and pass 3 respectively. From table 6.5 one can conclude that user’s showed more improvement from pass to pass in workstation than in the CAVE™. This is due to the fact that users were able to place the icons better in workstation (2-dimensional non-stereoscopic environment) on a 2-dimensional, 8.5” x 11” paper than in a CAVE™.

<table>
<thead>
<tr>
<th></th>
<th>Pass1 to Pass2</th>
<th>Pass2 to Pass 3</th>
<th>Pass1 to Pass3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>-0.82</td>
<td>-23%</td>
<td>-1.15</td>
</tr>
<tr>
<td>W/S</td>
<td>-0.24</td>
<td>7%</td>
<td>-0.95</td>
</tr>
</tbody>
</table>

Table 6.6: B3 Overall Impressions Ratings pass to pass Comparison

Table 6.6 presents pass-to-pass comparison of Benchmark 3 overall impressions subjective ratings. The negative values in table 6.6 prove that pass 1 ratings were lower than pass 2 and pass 2 ratings were lower than pass 3. For example a value of -28% for Workstation (pass 1 to pass 3) is calculated as 
\[(3.38-4.33)/3.38,\] where 3.38 and 4.33 represent the means of Benchmark 3 overall impressions ratings for pass 1 and pass 3 respectively. From table 6.6 one can conclude that the CAVE™ is preferred over workstation.
7. BENCHMARK 4 (FAULT IDENTIFICATION)

7.1 Description

In a typical design review process, a design space is presented to the reviewer(s) who examine the space for design flaws (faults). The purpose of this study is to help determine the applicability/usability of various user interfaces (both stereoscopic and non-stereoscopic) in improving this process. Based on the preliminary results of the previous Benchmark testing, a fourth Benchmark scenario was prepared to use the stereoscopic CAVE™ environment for the location and identification of faults within a design space. The scenario implemented and reported here is built upon the operations and scenarios developed for Benchmarks 1, 2, and 3.

Using the same virtual factory space as used for Benchmark 1, ten distinct design faults were injected into this space similar to those prepared for Benchmark 2 (find/repair). However, the Benchmark 4 testing requires only that the users utilize the interface to locate and identify as many of these faults as possible in four minutes. As with the previous testing, each user searches the faults utilizing the traditional CAD workstation (non-stereoscopic interface) and the stereoscopic wand interface in the CAVE™ environment. The two scenario sequences were randomized (non-stereoscopic vs. CAVE™) and users were randomly assigned to start with either the non-stereoscopic interface or in the CAVE™ environment.

As each user progressed through the active scenario/environment locating and identifying faults, the specific fault and the elapsed time was recorded for the analysis. Although this
method provides a significant quantity of data, for Benchmark 4, the key metric for comparison was the total number of faults found in each environment.

This exercise (Benchmark 4) was repeated in each of the two environments under test and the User Survey administered to each user after each pass in each environment. As with the other Benchmark testing, sequencing of the testers through the two environments was randomized so that not all of the users were testing the same interface in the same order. This randomization was used to eliminate bias in the testing.

7.2 Benchmark 4, Pass 3, faults count Analysis:

The following is a presentation of the Benchmark 4, pass 3; faults count analysis for all the users. Pass 3 results are presented here as representative of user best-final case results. All other results are presented in Appendix D

Figure 7.1 presents the user’s ability to find faults in a span of four minutes in each of the two environments. The results clearly indicate a higher fault count using the stereoscopic CAVE™ environment. In CAVE™, users on an average located 9.17 or 9 out of 10 faults in a span of 4 minutes. On the other hand, in workstation, users on an average located 7.1 or 7 out of 10 faults in a span of 4 minutes.
Figure 7.1: B4p3 Faults Count

Table 7.1: B4p3 Faults Count Statistics

<table>
<thead>
<tr>
<th>B4P3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>9.17</td>
<td>0.7</td>
<td>8</td>
<td>10</td>
<td>&lt;0.10</td>
<td>No</td>
<td>1%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>7.1</td>
<td>0.66</td>
<td>6</td>
<td>8</td>
<td>&lt;0.10</td>
<td>No</td>
<td>1%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th></th>
<th>Levene's Test</th>
<th>Mann-Whitney Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F-Value</td>
<td>Pr &gt; F Var?</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>0.26</td>
<td>0.61</td>
</tr>
</tbody>
</table>
### 7.3 B4p3- Benchmark 4 Pass 3 Descriptive Statistics

Table 7.1 (Benchmark 4 pass 3 faults count / B4p3) presents the results of the descriptive statistics analysis of user’s pass 3 faults count in the two-test environment. The K.S. test is used to test for normality of data. Since the P value is less than 0.1, the data are not normal. The Levene’s test to test for equal variance was then used. Since the P value is greater than 0.1 the data have equal variance. Since the data are not normal, Mann Whitney test is used. With the Mann Whitney test, P value is less than 0.1, which indicates that medians are unequal for CAVE™ and workstation. Examination of these results shows that for the two environments, the differences are statistically significant. The conclusion then is that at the 90% confidence level, there is significant evidence to support the alternative hypothesis (Hₐ). Thus, since the stereoscopic wand environment demonstrates faster faults count, CAVE™ is statistically “better” than non-stereoscopic workstation environment for Benchmark 4 during pass 3.

### 7.4 Benchmark 3 pass 3 Overall Impressions Ratings Analysis:

Figure 7.2 (Benchmark 4 pass 3 Overall Impressions Ratings / B4p3Ovr) graphically presents comparisons of the Benchmark 4 (faults count) pass 3 overall ratings of the two environments. Inspection of the average ratings shows that users preferred the stereoscopic environment (CAVE™) over the non-stereoscopic environment (workstation).
7.5 Detailed Statistical Analysis

The following sections present a detailed statistical analysis of user overall impressions ratings of the two test environments following their 3rd and final pass of the Benchmark 4 scenario. All other results are presented in Appendix D. The statistical analysis of these ratings provides insight into the final opinions of the users. As discussed before, the NCSS software package was used to perform each analysis. Each set of user overall impressions ratings is first examined to determine if the data are normally distributed (Gaussian distribution) using the KS statistic. The descriptive statistics test results are presented in tabular form followed by the results of Levene’s test for equal variance of the data. The null hypothesis (H₀) and alternative hypothesis (Hₐ) discussed for Benchmark 1 statistical analysis testing applies here (Benchmark 4) as well.

7.6 Benchmark 4 Pass 3 Overall Impressions Ratings Statistics

Table 7.2 presents the results of the descriptive statistics analysis of user’s Benchmark 4 pass 3 overall impressions of the interface. The K.S. test is used to test for normality of data. Since the P value is less than 0.1 for workstation and the CAVE™, the data are not normal. Levene’s test is used to test for equal variance; since the P value is greater than 0.1 the data have equal variance. Since the data are not normal, Mann Whitney test is used. But with Mann Whitney test, P value is less than 0.1, which indicates that medians are unequal for the CAVE™ and workstation. Examination of these results shows that for the two environments, the differences are statistically significant. The conclusion then is that at the 90% confidence level, there is significant evidence to support the alternative
hypothesis (H₀). This proves that the CAVETM environment is preferred over workstation environment in Benchmark 4 pass 3 overall impressions subjective ratings.

![Figure 7.2: B4p3Ovr Overall Impressions Ratings](image)

<table>
<thead>
<tr>
<th>B4OP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.65</td>
<td>0.2</td>
<td>4</td>
<td>5.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.36</td>
<td>0.23</td>
<td>3.8</td>
<td>4.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Mann-Whitney Test</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave vs W/S</td>
<td>0.01</td>
<td>0.99</td>
<td>Yes</td>
<td>-4.69</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Table 7.2: B4p3Ovr Overall Impressions Ratings Statistics**

<table>
<thead>
<tr>
<th>B4 Pass to Pass Comparison</th>
<th>Pass1 to Pass2</th>
<th>Pass2 to Pass 3</th>
<th>Pass1 to Pass3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diff</td>
<td>%</td>
<td>Diff</td>
<td>%</td>
</tr>
<tr>
<td>Cave</td>
<td>-0.93</td>
<td>-12%</td>
<td>-0.77</td>
</tr>
<tr>
<td>W/S</td>
<td>-0.93</td>
<td>-16%</td>
<td>-0.37</td>
</tr>
</tbody>
</table>

**Table 7.3: B4 Pass-to-Pass Comparison of Faults Count**

Table 7.3 presents pass-to-pass comparison of Benchmark 4 (Faults Count). The negative values in table 7.3 prove that pass 1 faults count was less than pass 2 and pass 2 faults
count was less than pass 3. For example a value of -22% for Workstation (pass 1 to pass 3) is calculated as (5.8-7.1)/5.8, where 5.8 and 7.1 represent the means of Benchmark 4 for pass 1 and pass 3 respectively. From table 7.3 one can conclude that user’s showed more improvement from pass-to-pass in the CAVE™ than in workstation. This is due to the fact that users found the faults easily in a four screen CAVE™ than on a single screen traditional CAD workstation.

<table>
<thead>
<tr>
<th>B4 Overall Ratings Pass to Pass Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Diff</td>
</tr>
<tr>
<td>Cave</td>
</tr>
<tr>
<td>W/S</td>
</tr>
</tbody>
</table>

**Table 7.4: B4 Overall Impressions Ratings Pass to Pass Comparison**

Table 7.4 presents pass-to-pass comparison of Benchmark 4 overall impressions subjective ratings. The negative values in table 7.4 prove that pass 1 ratings were lower than pass 2 and pass 2 ratings were lower than pass 3. For example a value of -32% for CAVE™ (pass 1 to pass 3) is calculated as (3.52-4.65)/3.52, where 3.52 and 4.65 represent the means of Benchmark 4 overall impressions ratings for pass 1 and pass 3 respectively. From table 7.4 one can conclude that the CAVE™ environment is preferred over workstation.
8. CONCLUSIONS AND RECOMMENDATIONS

The following is a brief description of all the Benchmarks to be tested in the future:

**Benchmark 1** – Navigation in each of the environments
3 passes; 30 users – using virtual factory model

**Benchmark 2** – Find/Repair in each of the environments
3 passes – 2 errors each, 30 users – using virtual factory model

**Benchmark 3** – Spatial Awareness
3 passes – 2 icons each, 30 users – using virtual Machine Shop

**Benchmark 4** – Individual Shopping List in each of the environments
3 passes – 4 (minutes) errors each; 30 users individually – using virtual factory model

**Benchmark 5** – Collaborative Shopping List in each of the environments
Groups of 3, each tester “driving” a pass – 4 (minutes) errors each – using virtual factory model

Generalizing the two environments comparisons presented in chapters 4, 5, 6, and 7; one can infer the following conclusions for all users:

For Benchmark 1 (Navigation) the statistics shows better results (lower timings and higher subjective ratings) for the CAVETM in both objective and subjective measures than the workstation, except for Benchmark 1 pass 3 Navigation and Overall ratings in which the subjective ratings do not suggest which of the two environments are better.

For Benchmark 2 (Find/Repair) the statistics shows better results (lower timings and higher subjective ratings) for the CAVETM in both objective and subjective measures than the workstation, except for Benchmark 2 pass 3 Location ratings, General, and Overall ratings in which the subjective ratings do not suggest which of the two environments are better.

For Benchmark 3 (spatial awareness) the statistics shows better results (lower offset distances and higher subjective ratings) for the CAVETM in both objective and subjective measures than the workstation.
For Benchmark 4 (shopping list), the statistics shows better results (lower timings and higher subjective ratings) for the CAVE™ in both objective and subjective measures than the workstation.

The results presented below prove the objective of this research that the state of the art Perceptual User Interface or PUI (CAVE™ and wand) are much better, efficient, faster environment than the traditional Graphical User Interface GUI (Workstation and mouse),

- 94% of the results were in favor of CAVE™ in both objective and subjective measures.
- 2/3 of the results for pass-to-pass improvement were better for the CAVE™ for both objective and subjective measures.

8.1 Enhancements for Further Study

During the course of the testing as documented in this report, users/testers, test administrators, and test developers often suggested possible enhancements to the interfaces that warrant further testing and evaluation. These enhancements ran the gambit from simply expanding the interface to include a “you-are-here” tracking map in one corner of the BOOM device or workstation or ImmersaDesk display to the multi-screen, immersive, CAVE™ environment. Although the research team would prefer to provide test data and evidence for each potential improvement, rigorous and professional usability testing required that the interfaces remain static during the actual summative tests. To minimize the impact of these constraints, a less formal initial test of some interfaces might be performed.

8.2 Tracking Map

The ‘you-are-here” tracking map enhancement was suggested by the user groups as an aid to navigation through a new space. As an initial implementation, the developers under Satter’s (2005) test regime attempted to insert such a map, in 3D, into the existing CAVE™, Workstation, ImmersaDesk environment. This initial implementation proved
to place too large a computing burden on the system and slowed user response times to what was considered to be an unusable level. The enhancement was removed due to the significant latency.

However, after considerable discussions with the users, administrators, and developers, the group came to the conclusion that the map enhancement might be implemented in a less burdensome manner. It is suggested that this enhancement be changed to a callable map activated/deactivated by the user via voice command. Thus, the compute load is not constant and the user requests “you-are-here” help only on-demand. Finally these maps should also be inserted in BOOM device for competitive usability testing in the future.

8.3 Future Work

Each of the Benchmarks needs to be tested in the future for the below mentioned scenarios with interfaces such as Voice/Data Glove, mouse, touch pad mouse and Wand (Wire & wireless type):

1) 3 sided CAVETM (UNO) vs. 4 sided CAVETM (Avondale) (Individually no Map)
2) 3 sided CAVETM (UNO) vs. 4 sided CAVETM (Avondale) (Collaborative no Map)
3) 3 sided CAVETM (UNO) vs. 4 sided CAVETM (Avondale) (Individually with Map)
4) 3 sided CAVETM (UNO) vs. 4 sided CAVETM (Collaborative with Map)
5) 21” Workstation vs. 3 sided CAVETM (UNO) (Individually without Map)
6) 21” Workstation vs. 3 sided CAVETM (UNO) (Collaborative without Map)
7) 21” Workstation vs. 3 sided CAVETM (UNO) (Individually with Map)
8) 21” Workstation vs. 3 sided CAVETM (UNO) (Collaborative with Map)
9) 21” Workstation vs. 4 sided CAVETM (Individually without Map) (Done)
10)21” Workstation vs. 4 sided CAVETM (Collaborative without Map)
11)21” Workstation vs. 4 sided CAVETM (Individually with Map)
12) 21” Workstation vs. 4 sided CAVE™ (Collaborative with Map)

13) 21” Workstation vs. 86” workbench Non Stereo (Individually without Map)

14) 21” Workstation vs. 86” workbench Non Stereo (Collaborative without Map)

15) 21” Workstation vs. 86” workbench Non Stereo (Individually with Map)

16) 21” Workstation vs. 86” workbench Non Stereo (Collaborative with Map)

17) 21” Workstation vs. 86” workbench Stereo (Individually without Map)

18) 21” Workstation vs. 86” workbench Stereo (Collaborative without Map)

19) 21” Workstation vs. 86” workbench Stereo (Individually with Map)

20) 21” Workstation vs. 86” workbench Stereo (Collaborative with Map)

21) 86” workbench Non Stereo vs. 3 sided CAVE™ (UNO) (Individually no Map)

22) 86” workbench Non Stereo vs. 3 sided CAVE™ (UNO) (Collaborative no Map)

23) 86” workbench Non Stereo vs. 3 sided CAVE™ (UNO) (Individually with Map)

24) 86” workbench Non Stereo vs. 3 sided CAVE™ (Collaborative with Map)

25) 86” workbench Stereo vs. 3 sided CAVE™ (UNO) (Individually without Map)

26) 86” workbench Stereo vs. 3 sided CAVE™ (UNO) (Collaborative without Map)

27) 86” workbench Stereo vs. 3 sided CAVE™ (UNO) (Individually with Map)

28) 86” workbench Stereo vs. 3 sided CAVE™ (UNO) (Collaborative with Map)

29) 86” workbench Non Stereo vs. 4 sided CAVE™ (Individually without Map)

30) 86” workbench Non Stereo vs. 4 sided CAVE™ (Collaborative without Map)

31) 86” workbench Non Stereo vs. 4 sided CAVE™ (Individually with Map)

32) 86” workbench Non Stereo vs. 4 sided CAVE™ (Collaborative with Map)

33) 86” workbench Stereo vs. 4 sided CAVE™ (Individually without Map)

34) 86” workbench Stereo vs. 4 sided CAVE™ (Collaborative without Map)
35) 86” workbench Stereo vs. 4 sided CAVE™ (Individually with Map)
36) 86” workbench Stereo vs. 4 sided CAVE™ (Collaborative with Map)
37) 86” workbench Non Stereo vs. 86” workbench Stereo (Individually no Map)
38) 86” workbench Non Stereo vs. 86” workbench Stereo (Collaborative no Map)
39) 86” workbench Non Stereo vs. 86” workbench Stereo (Individually with Map)
40) 86” workbench Non Stereo vs. 86” workbench Stereo (Collaborative with Map)
41) Finally all of the above mentioned Benchmarks can also be compared against a traditional BOOM device, five sided CAVE™ and six sided CAVE™.
References


64. Number Cruncher Statistical System (NCSS software, 2004).


68. Polhemus, (2005), „First in the Third Dimension,“. http://www.polhemus.com


of the ACM Conference on Virtual Reality Software and Technology, Lausanne, Switzerland, pp. 37-44.


Appendix A
Benchmark 1 (Navigation) Detail

Benchmark1 - Pass 1 Elapsed Times

<table>
<thead>
<tr>
<th>User #</th>
<th>Elapsed Time (in Sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1</td>
<td>600</td>
</tr>
<tr>
<td>U2</td>
<td>535</td>
</tr>
<tr>
<td>U3</td>
<td>538</td>
</tr>
<tr>
<td>U4</td>
<td>553</td>
</tr>
<tr>
<td>U5</td>
<td>563</td>
</tr>
<tr>
<td>U6</td>
<td>536</td>
</tr>
<tr>
<td>U7</td>
<td>539</td>
</tr>
<tr>
<td>U8</td>
<td>525</td>
</tr>
<tr>
<td>U9</td>
<td>520</td>
</tr>
<tr>
<td>U10</td>
<td>551</td>
</tr>
<tr>
<td>U11</td>
<td>550</td>
</tr>
<tr>
<td>U12</td>
<td>550</td>
</tr>
<tr>
<td>U13</td>
<td>550</td>
</tr>
<tr>
<td>U14</td>
<td>550</td>
</tr>
<tr>
<td>U15</td>
<td>550</td>
</tr>
<tr>
<td>U16</td>
<td>550</td>
</tr>
<tr>
<td>U17</td>
<td>550</td>
</tr>
<tr>
<td>U18</td>
<td>550</td>
</tr>
<tr>
<td>U19</td>
<td>550</td>
</tr>
<tr>
<td>U20</td>
<td>550</td>
</tr>
<tr>
<td>U21</td>
<td>550</td>
</tr>
<tr>
<td>U22</td>
<td>550</td>
</tr>
<tr>
<td>U23</td>
<td>550</td>
</tr>
<tr>
<td>U24</td>
<td>550</td>
</tr>
<tr>
<td>U25</td>
<td>550</td>
</tr>
<tr>
<td>U26</td>
<td>550</td>
</tr>
<tr>
<td>U27</td>
<td>550</td>
</tr>
<tr>
<td>U28</td>
<td>550</td>
</tr>
<tr>
<td>U29</td>
<td>550</td>
</tr>
<tr>
<td>U30</td>
<td>550</td>
</tr>
</tbody>
</table>

Figure A-1: B1p1Tim Elapsed Times

Table A-1: B1p1Tim Elapsed Times Statistics

<table>
<thead>
<tr>
<th>B1P1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>330.8</td>
<td>58.8</td>
<td>229</td>
<td>453.00</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>18%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>560.8</td>
<td>57.4</td>
<td>425</td>
<td>725.00</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>18%</td>
</tr>
</tbody>
</table>

Levene used to check for equal variance

Levene’s Test: F-Value = 1.13, P Value = 0.29

Homogeneity of Variance Test for Differences

- Levene’s Test: F-Value = 1.13, P Value = 0.29
- T-Test: Value = -15.33, P Value = < 0.001

Cave vs W/S

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene used to check for equal variance</td>
<td>Mean</td>
<td></td>
</tr>
</tbody>
</table>
Figure A-2: B1p1Nav Navigation Ratings

Table A-2: B1p1Nav Navigation Ratings Statistics
Figure A-3: B1p1Loc Locating Ratings

<table>
<thead>
<tr>
<th>B1LP1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.51</td>
<td>0.22</td>
<td>3</td>
<td>4.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>0.06%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.33</td>
<td>0.17</td>
<td>3</td>
<td>3.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>0.05%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal</th>
<th>Mann-Whitney Test</th>
<th>Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.256</td>
<td>0.62</td>
<td>Yes</td>
<td>-3.45</td>
<td>0.0003</td>
<td>No</td>
<td>Cave</td>
<td></td>
</tr>
</tbody>
</table>

Table A-3: B1p1Loc Locating Ratings Statistics

119
Benchmark 1 - Pass 1 - Movement Ratings

Figure A-4: B1p1Mov Movement Ratings

Table A-4: B1p1Mov Movement Ratings Statistics
Figure A-5: B1p1Gen General Impressions Ratings

Table A-5: B1p1Gen General Impressions Ratings Statistics
Benchmark 1 - Pass 1 - Overall Impressions Ratings

Figure A-6: B1p1Ovr Overall Impressions Ratings

Table A-6: B1p1Ovr Overall Impressions Ratings Statistics
Figure A-7: B1p2Tim Elapsed Times

Table A-7: B1p2Tim Elapsed Times Statistics

<table>
<thead>
<tr>
<th>B1P2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>290.7</td>
<td>54.5</td>
<td>207</td>
<td>424</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>19%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>495.9</td>
<td>88.2</td>
<td>350</td>
<td>785</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>18%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>Equal Var?</td>
</tr>
<tr>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>2.46</td>
</tr>
</tbody>
</table>
Benchmark 1 - Pass 2 - Navigation Ratings

Figure A-8: B1p2Nav Navigation Ratings

Table A-8: B1p2Nav Navigation Ratings Statistics
Figure A-9: B1p2Loc Locating Ratings

<table>
<thead>
<tr>
<th>B1LP2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.09</td>
<td>0.33</td>
<td>3.6</td>
<td>4.8</td>
<td>&lt;0.10</td>
<td>No</td>
<td>0.08%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.59</td>
<td>0.21</td>
<td>3.2</td>
<td>4.0</td>
<td>&lt;0.10</td>
<td>No</td>
<td>0.06%</td>
</tr>
</tbody>
</table>

### Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Levene's Test</th>
<th>Equal?</th>
<th>Mann-Whitney Test</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-Value</td>
<td>P Value</td>
<td>Value</td>
<td>P Value</td>
<td></td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>1.24</td>
<td>0.27</td>
<td>-5.61</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table A-9: B1p2Loc Locating Ratings Statistics
Figure A-10: B1p2Mov Movement Ratings

Table A-10: B1p2Mov Movement Ratings Statistics
Figure A-11: B1p2Gen General Impressions Ratings

Table A-11: B1p2Gen General Impressions Ratings Statistics
Figure A-12: B1p2Ovr Overall Impressions Ratings

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U1</td>
<td>U2</td>
<td>U3</td>
<td>U4</td>
<td>U5</td>
<td>U6</td>
<td>U7</td>
<td>U8</td>
<td>U9</td>
<td>U10</td>
<td>U11</td>
<td>U12</td>
<td>U13</td>
<td>U14</td>
<td>U15</td>
<td>U16</td>
<td>U17</td>
<td>U18</td>
<td>U19</td>
<td>U20</td>
<td>U21</td>
<td>U22</td>
<td>U23</td>
<td>U24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A-12: B1p2Ovr Overall Impressions Ratings Statistics
Figure A-13: B1p3Tim Elapsed Times

Table A-13: B1p3Tim Elapsed Times Statistics
Figure A-14: B1p3Nav Navigation Ratings

Table A-14: B1p3Nav Navigation Ratings Statistics
Figure A-15: B1p3Loc Locating Ratings

Table A-15: B1p3Loc Locating Ratings Statistics
Benchmark 1 - Pass 3 - Movement Ratings

![Bar chart showing movement ratings for different users and conditions.]

**Figure A-16: B1p3Mov Movement Ratings**

<table>
<thead>
<tr>
<th>B1MP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.45</td>
<td>0.37</td>
<td>4</td>
<td>5.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>0.08%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.21</td>
<td>0.25</td>
<td>3.6</td>
<td>4.60</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>0.07%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>Equal Man-Whitney Test</td>
</tr>
<tr>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>17.74</td>
</tr>
</tbody>
</table>

**Table A-16: B1p3Mov Movement Ratings Statistics**
**Figure A-17: B1p3Gen General Impressions Ratings**

<table>
<thead>
<tr>
<th>B1GP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.45</td>
<td>0.37</td>
<td>4</td>
<td>5.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>0.08%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.32</td>
<td>0.24</td>
<td>3.8</td>
<td>4.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.50%</td>
</tr>
</tbody>
</table>

**Homogeneity of Variance**

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;0.001</td>
<td>No</td>
<td>-1.37</td>
<td>Cave</td>
</tr>
</tbody>
</table>

**Table A-17: B1p3Gen General Impressions Ratings Statistics**
Benchmark 1 - Pass 3 - Overall Impressions Ratings

Figure A-18: B1p3Ovr Overall Impressions Ratings

<table>
<thead>
<tr>
<th>B1OP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.45</td>
<td>0.38</td>
<td>4</td>
<td>4.90</td>
<td>&lt;0.10</td>
<td>No</td>
<td>0.09%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.27</td>
<td>0.18</td>
<td>3.95</td>
<td>4.75</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>0.04%</td>
</tr>
</tbody>
</table>

Table A-18: B1p3Ovr Overall Impressions Ratings Statistics
Figure A-19: B1-3pAvgTim Elapsed Times

<table>
<thead>
<tr>
<th>B1PA</th>
<th># Users</th>
<th>Mean (in Sec.)</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>280.3</td>
<td>47.2</td>
<td>202</td>
<td>393</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>17%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>493.7</td>
<td>73.5</td>
<td>395</td>
<td>748</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>15%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene’s Test</td>
</tr>
<tr>
<td>F-Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
</tr>
</tbody>
</table>

Table A-19: B1-3pAvgTim Elapsed Times Statistics
Figure A-20: B1-3pAvgNav Navigation Ratings

Table A-20: B1-3pAvgNav Navigation Ratings Statistics
Benchmark 1 - 3 Pass Average - Locating Ratings

Figure A-21: B1-3pAvgLoc Locating Ratings

Table A-21: B1-3pAvgLoc Locating Ratings Statistics
Benchmark 1 - 3 Pass Average - Movement Ratings

Figure A-22: B1-3pAvgMov Movement Ratings

<table>
<thead>
<tr>
<th>B1M3PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.03</td>
<td>0.23</td>
<td>3.67</td>
<td>4.47</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>6.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.72</td>
<td>0.14</td>
<td>3.47</td>
<td>4.00</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
</tbody>
</table>

Table A-22: B1-3pAvgMov Movement Ratings Statistics

<table>
<thead>
<tr>
<th>Test for Differences</th>
<th>Levene's Test</th>
<th>Equal Variance</th>
<th>Mann-Whitney Test</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave vs W/S</td>
<td>F-Value</td>
<td>P Value</td>
<td>Value</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Figure A-23: B1-3pAvgGen General Impressions Ratings

Table A-23: B1-3pAvgGen General Impressions Ratings Statistics
Figure A-24: B1-3pAvgOvr Overall Impressions Ratings

<table>
<thead>
<tr>
<th>B1O3PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.03</td>
<td>0.2</td>
<td>3.7</td>
<td>4.43</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>0.05%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.73</td>
<td>0.14</td>
<td>3.52</td>
<td>4.10</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>0.04%</td>
</tr>
</tbody>
</table>

Table A-24: B1-3pAvgOvr Overall Impressions Ratings Statistics

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.65</td>
<td>0.02</td>
<td>No</td>
<td>-5.21</td>
<td>&lt;0.001</td>
<td>No</td>
<td>Cave</td>
</tr>
</tbody>
</table>
Appendix B
Benchmark 2 (Find and Repair Manipulation) Detail

**Benchmark 2 - Pass 1 Elapsed Times**

![Bar chart showing elapsed times for different users.]

**Figure B-1: B2p1Tim Elapsed Times**

<table>
<thead>
<tr>
<th>B2P1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>323.6</td>
<td>67.44</td>
<td>195</td>
<td>450</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>21%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>428.6</td>
<td>65.65</td>
<td>297</td>
<td>593</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>15%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test F-Value</td>
<td>T-Test Equal? Significant?</td>
</tr>
<tr>
<td>P Value</td>
<td>Equal?</td>
</tr>
</tbody>
</table>

| Cave vs W/S | 0.3  | 0.58 | Yes | -6.11 | <0.0001 | No | Cave |

**Table B-1: B2p1Tim Elapsed Times Statistics**
Figure B-2: B2p1Nav Navigation Ratings

Table B-2: B2p1Nav Navigation Ratings Statistics
### Figure B-3: B2p1Loc Locating Ratings

### Table B-3: B2p1Loc Locating Ratings Statistics
Figure B-4: B2p1Mov Movement Ratings

<table>
<thead>
<tr>
<th>B2MP1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.58</td>
<td>0.28</td>
<td>3</td>
<td>4.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>8.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.39</td>
<td>0.23</td>
<td>3</td>
<td>4.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>7.00%</td>
</tr>
</tbody>
</table>

Table B-4: B2p1Mov Movement Ratings Statistics
Figure B-5: B2p1Gen General Impressions Ratings

<table>
<thead>
<tr>
<th>B2GP1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.65</td>
<td>0.28</td>
<td>3</td>
<td>4.40</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>5.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.39</td>
<td>0.21</td>
<td>3</td>
<td>3.80</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>8.00%</td>
</tr>
</tbody>
</table>

Table B-5: B2p1Gen General Impressions Ratings Statistics
Figure B-6: B2p1Ovr Overall Impressions Ratings

Table B-6: B2p1Ovr Overall Impressions Ratings Statistics
Figure B-7: B2p2Tim Elapsed Times

Table B-7: B2p2Tim Elapsed Times Statistics
Figure B-8: B2p2Nav Navigation Ratings

<table>
<thead>
<tr>
<th>B2NP2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.26</td>
<td>0.35</td>
<td>3.8</td>
<td>4.80</td>
<td>&lt;0.10</td>
<td>No</td>
<td>8.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.66</td>
<td>0.26</td>
<td>3.2</td>
<td>4.40</td>
<td>&lt;0.10</td>
<td>No</td>
<td>7.00%</td>
</tr>
</tbody>
</table>

Table B-8: B2p2Nav Navigation Ratings Statistics
Figure B-9: B2p2Loc Locating Ratings

Table B-9: B2p2Loc Locating Ratings Statistics
Benchmark 2 - Pass 2 - Movement Ratings

Figure B-10: B2p2Mov Movement Ratings

Table B-10: B2p2Mov Movement Ratings Statistics
Figure B-11: B2p2Gen General Impression Ratings

Table B-11: B2p2Gen General Impression Ratings Statistics
Figure B-12: B2p2Ovr Overall Impression Ratings

<table>
<thead>
<tr>
<th>B2OP2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.27</td>
<td>0.35</td>
<td>3.9</td>
<td>4.85</td>
<td>&lt;0.10</td>
<td>No</td>
<td>8.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.63</td>
<td>0.27</td>
<td>3.15</td>
<td>4.10</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>7.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
</tr>
<tr>
<td>Equal Var?</td>
</tr>
<tr>
<td>F-Value</td>
</tr>
<tr>
<td>P Value</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>0.51</td>
<td>0.48</td>
<td>Yes</td>
<td>Cave</td>
</tr>
</tbody>
</table>

Table B-12: B2p2Ovr Overall Impression Ratings Statistics
Benchmark 2 - Pass 3 Elapsed Times

Figure B-13: B2p3Tim Elapsed Times

<table>
<thead>
<tr>
<th>B2P3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>187.5</td>
<td>42.7</td>
<td>111</td>
<td>260</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>22%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>272.6</td>
<td>59.3</td>
<td>173</td>
<td>408</td>
<td>&lt;0.10</td>
<td>No</td>
<td>22%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Levene's Test</th>
<th>Equal</th>
<th>Mann-Whitney Test</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-Value</td>
<td>P Value</td>
<td>Value</td>
<td>P Value</td>
<td></td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>0.94</td>
<td>0.36</td>
<td>-5.20</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Table B-13: B2p3Tim Elapsed Times Statistics
Figure B-14: B2p3Nav Navigation Ratings

Table B-14: B2p3Nav Navigation Ratings Statistics
Figure B-15: B2p3Loc Locating Ratings

Table B-15: B2p3Loc Locating Ratings Statistics
Figure B-16: B2p3Mov Movement Ratings

<table>
<thead>
<tr>
<th>B2MP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.43</td>
<td>0.45</td>
<td>4</td>
<td>5.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>10.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.21</td>
<td>0.26</td>
<td>3.6</td>
<td>4.60</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>6.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance
Levene's Test Equal Var?
F-Value P Value Cave vs W/S <0.001 No

Test for Differences
Mann-Whitney Test Equal? Significant?
Value P Value Cave -1.48 0.07 No Cave

Table B-16: B2p3Mov Movement Ratings Statistics
Figure B-17: B2p3Gen General Impressions Ratings

Table B-17: B2p3Gen General Impressions Ratings Statistics
Figure B-18: B2p3Ovr Overall Impressions Ratings

Table B-18: B2p3Ovr Overall Impressions Ratings Statistics
Benchmark 2 - 3 Pass Average Elapsed Times

Figure B-19: B2-3pAvgTim Elapsed Times

Table B-19: B2-3pAvgTim Elapsed Times Statistics

<table>
<thead>
<tr>
<th>B2PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>255.1</td>
<td>47.1</td>
<td>168</td>
<td>334</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>18%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>350.23</td>
<td>52.9</td>
<td>267</td>
<td>485</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>15%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance

<table>
<thead>
<tr>
<th>Levene's Test</th>
<th>Equal T-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Figure B-20: B2-3pAvgNav Navigation Ratings

Table B-20: B2-3pAvgNav Navigation Ratings Statistics
Figure B-21: B2-3pAvgLoc Locating Ratings

<table>
<thead>
<tr>
<th>B2L3PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.09</td>
<td>0.2</td>
<td>3.8</td>
<td>4.53</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>5.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.76</td>
<td>0.15</td>
<td>3.4</td>
<td>4.20</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance

<table>
<thead>
<tr>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
</tr>
<tr>
<td>F-Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
</tr>
</tbody>
</table>

Mann-Whitney Test

<table>
<thead>
<tr>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>P Value</td>
</tr>
<tr>
<td>-5.39</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table B-21: B2-3pAvgLoc Locating Ratings Statistics
Figure B-22: B2-3pAvgMov Movement Ratings

Table B-22: B2-3pAvgMov Movement Ratings Statistics
**Figure B-23: B2-3pAvgGen General Impressions Ratings**

<table>
<thead>
<tr>
<th>B2G3PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.11</td>
<td>0.2</td>
<td>3.8</td>
<td>4.47</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>5.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.81</td>
<td>0.19</td>
<td>3.4</td>
<td>4.13</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>T-Test</th>
<th>Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>0.34</td>
<td>0.56</td>
<td>Yes</td>
<td>-5.85</td>
<td>&lt;0.001</td>
<td>No</td>
<td>Cave</td>
<td></td>
</tr>
</tbody>
</table>

**Table B-23: B2-3pAvgGen General Impressions Ratings Statistics**
Table B-24: B2-3pAvgOvr Overall Impressions Ratings Statistics
Appendix C
**Benchmark 3 (Spatial Awareness)**

**Figure C-1: B3p1-1off Pass 1- Icon 1 Offsets**

<table>
<thead>
<tr>
<th>B3Part1P1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>24.53</td>
<td>10.23</td>
<td>7</td>
<td>44</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>42%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>30.1</td>
<td>8.78</td>
<td>14</td>
<td>44</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>29%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>Equal T-Test</td>
</tr>
<tr>
<td>F-Value</td>
<td>Pr &gt; F</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>0.42</td>
</tr>
</tbody>
</table>

**Table C-1: B3p1-1off Pass 1- Icon 1 Offsets Statistics**
Figure C-2: B3p1-2off Pass 1- Icon 2 Offsets

Table C-2: B3p1-2off Pass 1- Icon 2 Offsets Statistics
Figure C-3: B3p1Nav- Pass 1- Navigation Ratings

<table>
<thead>
<tr>
<th>B3NP1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.49</td>
<td>0.17</td>
<td>3.14</td>
<td>3.86</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.38</td>
<td>0.11</td>
<td>3.14</td>
<td>3.57</td>
<td>&lt;0.10</td>
<td>No</td>
<td>3.00%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>Mann-Whitney Test</td>
</tr>
<tr>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>2.73</td>
</tr>
</tbody>
</table>

Table C-3: B3p1Nav- Pass 1- Navigation Ratings Statistics
### Figure C-4: B3p1Loc- Pass 1- Locating Ratings

<table>
<thead>
<tr>
<th>User #</th>
<th>Rating</th>
<th>Cave</th>
<th>Wk Sh</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.40</td>
<td>3.60</td>
<td>3.20</td>
<td>3.40</td>
</tr>
<tr>
<td>2</td>
<td>3.40</td>
<td>3.60</td>
<td>3.20</td>
<td>3.40</td>
</tr>
<tr>
<td>3</td>
<td>3.40</td>
<td>3.40</td>
<td>3.20</td>
<td>3.40</td>
</tr>
<tr>
<td>4</td>
<td>3.40</td>
<td>3.40</td>
<td>3.20</td>
<td>3.40</td>
</tr>
</tbody>
</table>

**Table C-4: B3p1Loc- Pass 1- Locating Ratings Statistics**

<table>
<thead>
<tr>
<th>B3LP1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.54</td>
<td>0.2</td>
<td>3</td>
<td>4.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>6.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.35</td>
<td>0.17</td>
<td>3</td>
<td>3.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

### Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Test</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>0.13</td>
<td>0.72</td>
</tr>
</tbody>
</table>

170
### Benchmark 3 - Pass 1 - Movement Ratings

**Figure C-5: B3p1Mov- Pass 1- Movement Ratings**

**Table C-5: B3p1Mov- Pass 1- Movement Ratings Statistics**

<table>
<thead>
<tr>
<th>B3MP1</th>
<th>Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.56</td>
<td>0.24</td>
<td>3.2</td>
<td>4.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>7.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.38</td>
<td>0.2</td>
<td>3</td>
<td>3.80</td>
<td>&lt;0.10</td>
<td>No</td>
<td>6.00%</td>
</tr>
</tbody>
</table>

**Homogeneity of Variance Test for Differences**

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal Var?</th>
<th>Mann-WhitneyTest</th>
<th>Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.87</td>
<td>0.18</td>
<td>Yes</td>
<td>-2.72</td>
<td>0.003</td>
<td>No</td>
<td>Cave</td>
<td></td>
</tr>
</tbody>
</table>
Figure C-6: B3p1Gen- Pass 1- General Impressions Ratings

Table C-6: B3p1Gen- Pass 1- General Impressions Ratings Statistics
Figure C-7: B3p1Ovr- Pass 1 Overall Impressions Ratings

Table C-7: B3p1Ovr- Pass 1 Overall Impressions Ratings Statistics
**Figure C-8: B3p2-1off- Pass 2-Icon 1 Offsets**

**Table C-8: B3p2-1off- Pass 2-Icon 1 Offsets Statistics**
Figure C-9: B3p2-2off- Pass 2-Icon 2 Offsets

Table C-9: B3p2-2off- Pass 2-Icon 2 Offsets Statistics
Figure C-10: B3p2Nav Pass 2-Navigation Ratings

<table>
<thead>
<tr>
<th>B3NP2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.33</td>
<td>0.22</td>
<td>3.86</td>
<td>4.57</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.67</td>
<td>0.19</td>
<td>3.43</td>
<td>4.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

Table C-10: B3p2Nav Pass 2-Navigation Ratings Statistics
Figure C-11: B3p2Loc Pass 2 Locating Ratings

Table C-11: B3p2Loc Pass 2 Locating Ratings Statistics
Benchmark 3 - Pass 2 - Movement Ratings

Figure C-12: B3p2Mov Pass 2 Movement Ratings

Table C-12: B3p2Mov Pass 2 Movement Ratings Statistics
Figure C-13: B3p2Gen Pass 2 General Impressions Ratings

Table C-13: B3p2Gen Pass 2 General Impressions Ratings Statistics
Benchmark 3 - Pass 2 - Overall Impressions Ratings

User # | Rating
--- | ---
Wk Sta | 3.70 3.65 3.25 3.45 3.25 3.25 3.60 3.40 3.40 3.75 3.70 3.25 3.45 3.75 3.75 3.85 3.35 3.90 3.90 3.65 3.65 4.05 3.80 3.60 3.90 3.55 3.50 3.50 3.80 4.00 3.62
Cave | 4.10 4.30 4.55 4.45 4.00 4.00 4.00 4.45 4.50 4.60 4.30 4.25 4.45 4.55 4.55 4.80 4.60 4.60 4.40 4.35 4.55 4.40 4.65 4.50 4.60 4.40 4.30

Figure C-14: B3p2Ovr Pass 2 Overall Impressions Ratings

<table>
<thead>
<tr>
<th>B3OP2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.36</td>
<td>0.26</td>
<td>3.8</td>
<td>4.80</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.62</td>
<td>0.24</td>
<td>3.25</td>
<td>4.05</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>6.00%</td>
</tr>
</tbody>
</table>

Table C-14: B3p2Ovr Pass 2 Overall Impressions Ratings Statistics

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>Equal T-Test</td>
</tr>
<tr>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>0.31</td>
</tr>
</tbody>
</table>
Figure C-15: B3p3-1off Pass 3-Icon 1 Offsets

Table C-15: B3p3-1off Pass 3-Icon 1 Offsets Statistics
Figure C-16: B3p3-2off Pass 3-Icon 2 Offsets

<table>
<thead>
<tr>
<th>B3Part2P3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>7.77</td>
<td>7.16</td>
<td>0</td>
<td>28</td>
<td>&lt;0.10</td>
<td>No</td>
<td>92%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>13.8</td>
<td>16.53</td>
<td>0</td>
<td>91</td>
<td>&lt;0.10</td>
<td>No</td>
<td>120%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>Equal Var?</td>
</tr>
<tr>
<td>F-Value</td>
<td>Pr &gt; F</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>1.07</td>
</tr>
</tbody>
</table>

Table C-16: B3p3-2off Pass 3-Icon 2 Offsets Statistics
Figure C-17: B3p3Nav Pass 3 Navigation Ratings

Table C-17: B3p3Nav Pass 3 Navigation Ratings Statistics
Benchmark 3 - Pass 3 - Locating Ratings

Figure C-18: B3p3Loc Pass 3 Locating Ratings

Table C-18: B3p3Loc Pass 3 Locating Ratings Statistics
Figure C-19: B3p3Mov Pass 3 Movement Ratings

<table>
<thead>
<tr>
<th>B3MP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.65</td>
<td>0.19</td>
<td>4.2</td>
<td>5.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.3</td>
<td>0.23</td>
<td>3.6</td>
<td>4.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

### Homogeneity of Variance

- **Levene's Test**: Equal
- **Mann-Whitney Test**: Value

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.02</td>
<td>0.88</td>
<td>Yes</td>
<td>No Cave</td>
</tr>
</tbody>
</table>

Table C-19: B3p3Mov Pass 3 Movement Ratings Statistics
Figure C-20: B3p3Gen Pass 3 General Impressions Ratings

Table C-20: B3p3Gen Pass 3 General Impressions Ratings Statistics
Figure C-21: B3p3Ovr Pass 3 Overall Impressions Ratings

Table C-21: B3p3Ovr Pass 3 Overall Impressions Ratings Statistics
**Figure C-22: B3-3pA-1Off 3 Pass Avg. Part 1 Offsets**

![Graph showing benchmark offsets](image)

**Table C-22: B3-3pA-1Off 3 Pass Avg. Part 1 Offsets Statistics**

<table>
<thead>
<tr>
<th>B3Part1PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>15.86</td>
<td>5.58</td>
<td>7</td>
<td>34</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>35%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>19.43</td>
<td>6.21</td>
<td>7.7</td>
<td>33</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>32%</td>
</tr>
</tbody>
</table>

**Homogeneity of Variance Test for Differences**

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>Pr &gt; F</th>
<th>Var?</th>
<th>T-Test</th>
<th>Value</th>
<th>Pr &gt; T</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.75</td>
<td>0.39</td>
<td>Yes</td>
<td>2.34</td>
<td>0.01</td>
<td>No</td>
<td>Cave</td>
<td></td>
</tr>
</tbody>
</table>
Figure C-23: B3-3pA-1Off 3 Pass Avg. Part 2 Offsets

Table C-23: B3-3pA-1Off 3 Pass Avg. Part 2 Offsets Statistics
### Benchmark 3 - 3 Pass Average. - Navigation Ratings

<table>
<thead>
<tr>
<th>User</th>
<th>Rating</th>
</tr>
</thead>
</table>

**Figure C-24: B3-3pAnav 3 Pass Avg. Navigation Ratings**

**Table C-24: B3-3pAnav 3 Pass Avg. Navigation Ratings Statistics**

<table>
<thead>
<tr>
<th>B3N3PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.14</td>
<td>0.12</td>
<td>3.9</td>
<td>4.38</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.8</td>
<td>0.1</td>
<td>3.62</td>
<td>4.00</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>3.00%</td>
</tr>
</tbody>
</table>

**Levene's Test for Differences**

- **F-Value**: Cave vs W/S = 1.13
- **P Value**: 0.29

**Equal Var?** Yes, T-Test:

- **Value**: -11.54
- **P Value**: <0.001

**Equal?** Yes, Cave
Figure C-25: B3-3pALoc 3 Pass Avg. Locating Ratings

Table C-25: B3-3pALoc 3 Pass Avg. Locating Ratings Statistics
Benchmark 3 - 3 Pass Average. - Movement Ratings

Figure C-26: B3-3pAmov 3 Pass Avg. Movement Ratings

Table C-26: B3-3pAmov 3 Pass Avg. Movement Ratings Statistics
Benchmark 3 - 3 Pass Average - General Impressions Ratings

Figure C-27: B3-3pAgen 3 Pass Avg. General Impressions Ratings

Table C-27: B3-3pAgen 3 Pass Avg. General Impressions Ratings Statistics
Figure C-28: B3-3pAOvr 3 Pass Avg. Overall Impressions Ratings

Table C-28: B3-3pAOvr 3 Pass Avg. Overall Impressions Ratings Statistics
Benchmark 4 Fault Identification Detail

![Benchmark 4 Faults Count - Pass 1](chart)

**Figure D-1: B4p1 Faults Count**

<table>
<thead>
<tr>
<th>B4P1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>7.47</td>
<td>0.94</td>
<td>6</td>
<td>9</td>
<td>&lt;0.10</td>
<td>No</td>
<td>13%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>5.8</td>
<td>0.76</td>
<td>4</td>
<td>7</td>
<td>&lt;0.10</td>
<td>No</td>
<td>13%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance

Levene's Test

<table>
<thead>
<tr>
<th>F-Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Var?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave vs W/S</td>
<td>3.92</td>
<td>0.05</td>
<td>No</td>
</tr>
</tbody>
</table>

Mann-Whitney

<table>
<thead>
<tr>
<th>Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave vs W/S</td>
<td>5.53</td>
<td>&lt;0.0001</td>
<td>No</td>
</tr>
</tbody>
</table>

Table D-1: B4p1 Faults Count Statistics
**Figure D-2: B4p1Nav Navigation Ratings**

**Table D-2: B4p1Nav Navigation Ratings Statistics**
### Figure D-3: B4p1Loc Locating Ratings

#### Table D-3: B4p1Loc Locating Ratings Statistics

<table>
<thead>
<tr>
<th>B4LP1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.49</td>
<td>0.17</td>
<td>3</td>
<td>3.80</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.33</td>
<td>0.18</td>
<td>3</td>
<td>3.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

#### Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Var?</th>
<th>Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.001</td>
<td>0.99</td>
<td>Yes</td>
<td>-3.35</td>
<td>0.0004</td>
<td>No</td>
<td>Cave</td>
</tr>
</tbody>
</table>
Benchmark 4 - Pass 1 - Movement Ratings

Figure D-4: B4p1Mov Movement Ratings

Table D-4: B4p1Mov Movement Ratings Statistics
Figure D-5: B4p1Gen General Impressions Ratings

<table>
<thead>
<tr>
<th>B4GP1</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.52</td>
<td>0.13</td>
<td>3.2</td>
<td>3.80</td>
<td>&lt;0.10</td>
<td>No</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.33</td>
<td>0.15</td>
<td>3</td>
<td>3.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th></th>
<th>Levene's Test</th>
<th>Mann-Whitney Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>0.21</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Table D-5: B4p1Gen General Impressions Ratings Statistics
Figure D-6: B4p1Ovr Overall Impressions Ratings

Table D-6: B4p1Ovr Overall Impressions Ratings Statistics
Figure D-7: B4p2 Faults Count

Table D-7: B4p2 Faults Count Statistics
Figure D-8: B4p2Nav Navigation Ratings

<table>
<thead>
<tr>
<th>B4NP2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.03</td>
<td>0.34</td>
<td>3.43</td>
<td>4.57</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>8.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.54</td>
<td>0.1</td>
<td>3.43</td>
<td>3.86</td>
<td>&lt;0.10</td>
<td>No</td>
<td>3.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Mann-Whitney Test</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>36.59</td>
<td>&lt;0.001</td>
<td>No</td>
<td>-6.38</td>
<td>Cave</td>
</tr>
</tbody>
</table>

Table D-8: B4p2Nav Navigation Ratings Statistics
Figure D-9: B4p2Loc Locating Ratings

Table D-9: B4p2Loc Locating Ratings Statistics
Benchmark 4 - Pass 2 - Movement Ratings

Figure D-10: B4p2Mov Movement Ratings

<table>
<thead>
<tr>
<th>B4MP2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.01</td>
<td>0.37</td>
<td>3.4</td>
<td>4.80</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>9.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.53</td>
<td>0.19</td>
<td>3</td>
<td>3.80</td>
<td>&lt;0.10</td>
<td>No</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

Table D-10: B4p2Mov Movement Ratings Statistics
Figure D-11: B4p2Gen General Impression Ratings

Table D-11: B4p2Gen General Impression Ratings Statistics
Figure D-12: B4p2Ovr Overall Impression Ratings

<table>
<thead>
<tr>
<th>B4OP2</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>3.96</td>
<td>0.29</td>
<td>3.6</td>
<td>4.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>7.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.53</td>
<td>0.21</td>
<td>3.2</td>
<td>4.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>6.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

Levene's Test
- F-Value: 3.97
- P Value: 0.05
- Equal Var?: No

Mann-Whitney Test
- Value: -5.23
- P Value: <0.001
- Equal?: No
- Significant?: Cave

Table D-12: B4p2Ovr Overall Impression Ratings Statistics
Figure D-13: B4p3 Faults Count

Table D-13: B4p3 Faults Count Statistics
Figure D-14: B4p3Nav Navigation Ratings

<table>
<thead>
<tr>
<th>B4NP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.59</td>
<td>0.17</td>
<td>4.43</td>
<td>4.86</td>
<td>&lt;0.10</td>
<td>No</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.33</td>
<td>0.12</td>
<td>4</td>
<td>4.57</td>
<td>&lt;0.10</td>
<td>No</td>
<td>3.00%</td>
</tr>
</tbody>
</table>

Table D-14: B4p3Nav Navigation Ratings Statistics

Homogeneity of Variance
Levene's Test
<table>
<thead>
<tr>
<th>F-Value</th>
<th>P Value</th>
<th>Equal Var?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave vs W/S</td>
<td>8.3</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Test for Differences
Mann-Whitney Test
<table>
<thead>
<tr>
<th>Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave vs W/S</td>
<td>-5.44</td>
<td>&lt;0.0001</td>
<td>No Cave</td>
</tr>
</tbody>
</table>
Figure D-15: B4p3Loc Locating Ratings

<table>
<thead>
<tr>
<th>B4LP3</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.63</td>
<td>0.2</td>
<td>4.3</td>
<td>5.00</td>
<td>&lt;0.10</td>
<td>No</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>4.34</td>
<td>0.16</td>
<td>3.8</td>
<td>4.60</td>
<td>&lt;0.10</td>
<td>No</td>
<td>4.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance Test for Differences

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
<th>F-Value</th>
<th>P Value</th>
<th>Var?</th>
<th>Value</th>
<th>P Value</th>
<th>Equal?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.98</td>
<td>0.09</td>
<td>No</td>
<td>-4.86</td>
<td>&lt;0.0001</td>
<td>No</td>
<td>Cave</td>
</tr>
</tbody>
</table>

Table D-15: B4p3Loc Locating Ratings Statistics
Figure D-16: B4p3Mov Movement Ratings

Table D-16: B4p3Mov Movement Ratings Statistics
Figure D-17: B4p3Gen General Impressions Ratings

Table D-17: B4p3Gen General Impressions Ratings Statistics
Figure D-18: B4p3Ovr Overall Impressions Ratings

Table D-18: B4p3Ovr Overall Impressions Ratings Statistics
Figure D-19: B4-3pAvg Faults Count

Table D-19: B4-3pAvg Faults Count Statistics
Figure D-20: B4-3pAvgNav Navigation Ratings

<table>
<thead>
<tr>
<th>B4N3PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.01</td>
<td>0.15</td>
<td>3.71</td>
<td>4.33</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.74</td>
<td>0.07</td>
<td>3.62</td>
<td>3.90</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>2.00%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Homogeneity of Variance</th>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
<td>Equal?</td>
</tr>
<tr>
<td>F-Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>No</td>
</tr>
<tr>
<td>10.46</td>
<td>0.002</td>
</tr>
<tr>
<td>Mann-Whitney Test</td>
<td>Equal?</td>
</tr>
<tr>
<td>Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Cave vs W/S</td>
<td>No</td>
</tr>
<tr>
<td>-5.96</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table D-20: B4-3pAvgNav Navigation Ratings Statistics
Figure D-21: B4-3pAvgLoc Locating Ratings

Table D-21: B4-3pAvgLoc Locating Ratings Statistics
Figure D-22: B4-3pAvgMov Movement Ratings

Table D-22: B4-3pAvgMov Movement Ratings Statistics
Figure D-23: B4-3pAvgGen General Impressions Ratings

<table>
<thead>
<tr>
<th>B4G3PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.04</td>
<td>0.13</td>
<td>3.8</td>
<td>4.30</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>3.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.74</td>
<td>0.14</td>
<td>3.47</td>
<td>4.00</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
</tbody>
</table>

Table D-23: B4-3pAvgGen General Impressions Ratings Statistics
Figure D-24: B4-3pAvgOvr Overall Impressions Ratings

Table D-24: B4-3pAvgOvr Overall Impressions Ratings Statistics

<table>
<thead>
<tr>
<th>B4O3PA</th>
<th># Users</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Low</th>
<th>High</th>
<th>P Value</th>
<th>Normal?</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave</td>
<td>30</td>
<td>4.04</td>
<td>0.13</td>
<td>3.8</td>
<td>4.33</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>3.00%</td>
</tr>
<tr>
<td>W/S</td>
<td>30</td>
<td>3.74</td>
<td>0.14</td>
<td>3.47</td>
<td>4.00</td>
<td>&gt;0.10</td>
<td>Yes</td>
<td>4.00%</td>
</tr>
</tbody>
</table>

Homogeneity of Variance

<table>
<thead>
<tr>
<th>Test for Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test</td>
</tr>
<tr>
<td>Equal Var?</td>
</tr>
<tr>
<td>Value</td>
</tr>
<tr>
<td>P Value</td>
</tr>
<tr>
<td>Equal?</td>
</tr>
<tr>
<td>Significant?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cave vs W/S</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-Value</td>
</tr>
<tr>
<td>P Value</td>
</tr>
<tr>
<td>Equal?</td>
</tr>
<tr>
<td>Significant?</td>
</tr>
</tbody>
</table>
Appendix 1 – Course Completion Certificates
Completion Certificate

This is to certify that

Kurt Satter

has completed the Human Participant Protection Education for Research Teams online course, sponsored by the National Institutes of Health (NIH), on 05/26/2006.

This course included the following:

- key historical events and current issues that impact guidelines and legislation on human participant protection in research.
- ethical principles and guidelines that should assist in resolving the ethical issues inherent in the conduct of research with human participants.
- the use of key ethical principles and federal regulations to protect human participants at various stages in the research process.
- a description of guidelines for the protection of special populations in research.
- a definition of informed consent and components necessary for a valid consent.
- a description of the role of the IRB in the research process.
- the roles, responsibilities, and interactions of federal agencies, institutions, and researchers in conducting research with human participants.
Completion Certificate

This is to certify that

Syed Ahmed

has completed the Human Participant Protection Education for Research Teams online course, sponsored by the National Institutes of Health (NIH), on 07/12/2005.

This course included the following:

- key historical events and current issues that impact guidelines and legislation on human participant protection in research.
- ethical principles and guidelines that should assist in resolving the ethical issues inherent in the conduct of research with human participants.
- the use of key ethical principles and federal regulations to protect human participants at various stages in the research process.
- a description of guidelines for the protection of special populations in research.
- a definition of informed consent and components necessary for a valid consent.
- a description of the role of the IRB in the research process.
- the roles, responsibilities, and interactions of federal agencies, institutions, and researchers in conducting research with human participants.
Appendix 2 – Consent Form
Consent Form

1. Title:
Usability Studies with Virtual and Traditional Computer Aided Design Environments

2. Investigators:
PI at UNO: Donald Barbe, PhD, PE, UNO, LA 70148 (504) 280-7062, debce@uno.edu
Dr. Kurt Michael Satter, UNO, LA 70148, (504) 610-8256, ksatter@uno.edu
Co-I at Dillard University: Professor Syed Adeel Ahmed, (504) 816-4506, FAX (504) 816-4724, sahmed@dillard.edu

3. Purpose of the Research
The research conducted in this project focuses on the use of gestural interfaces in immersive virtual environments when compared with traditional methods. Conventional methods of computing include the use of laptop and desktop workstations. Our purpose in this research is to try new interfaces to determine which interface is best for eventual use in immersive virtual environments. Essentially, we are testing for user friendliness of new human-computer interfaces.

4. Procedures for this Research
In order to complete the studies in this project, participants will use different versions of software and different pieces of hardware to determine the effectiveness of the various human-computer interaction methods. For example, gloves with which a user can “grab” a three dimensional image may be tested, voice commands may be employed, or a wand device may be used to point, select, and move three dimensional images. An ability to navigate through an artificial world or awareness of geometric relationships may also be tested. Some tests may involve traditional computer workstations and conventional keyboard and mouse interfaces. During these tests, recordings of activity may be used to document system effectiveness and use. This may include videotaping, voice recording, note taking by research staff, and completion of survey forms. Typically, a scenario for completing a design or interaction task is used to give you an opportunity to try the different versions of the hardware and software. Research staff will instruct you and give you an opportunity to ask questions. All testing with hardware and software will be done in cooperation with one or more members of the project research staff.

5. Potential Risks of Discomfort
For part of the evaluation, you will be wearing specialized liquid-crystal display (LCD) stereoscopic glasses. These glasses allow users to see three-dimensional images by turning each eye-piece lens on and off at a rate of 120 times per second. Although many virtual environment researchers and developers have observed no problems using the glasses for prolonged periods of time, you should be aware that there is evidence that flicker due to cycling can cause problems in individuals with certain brain abnormalities such as epilepsy. To eliminate the possibility of this type of problem we require that you answer questions reviewed by a physician to ensure your safety during testing. Please let the evaluators know if you feel that you may be at risk or if you have any questions regarding the safety of the stereographic glasses.

Also, you will be standing during the majority of session. When wearing the stereoscopic environment glasses, your visibility will be somewhat limited; however, you will be able to see things around you in the room or other test space --- much like wearing tinted sunglasses indoors. Since virtual environments have the potential to be visually and psychologically immersive, at some time, you may feel that you are not standing in a room operating a computer, but instead, part of some virtual world. We will have someone watching you at all times to ensure that you remain balanced and in a safe standing position. We will also be watching to ensure that the cables attached to the glasses do not become entwined in such a manner that may cause you to become unbalanced. Finally, we will be monitoring you for signs of eyestrain and/or dizziness. If at any time you feel any signs of dizziness, eyestrain, or lack of balance, stop what you are doing, let us know that you may be feeling uncomfortable, and if necessary, take the glasses off.

You will be asked to perform various tasks with this system, all of which require you to view objects in a virtual or synthetic environment (either with stereographic glasses, or without) and manipulate objects using various interfaces (possibly including a wand, gloves, voice, and other techniques). We are evaluating the ability of our system as a means to accomplish specific information retrieval, manipulation of geometry, navigation and other related and tasks, to make the virtual environment system as effective and usable as possible. We are not in any way evaluating you. We expect one testing session to last about one half hour or less, and you may be videotaped during the session. During this session, the virtual or synthetic environment and the input device(s) may be videotaped; your hands may be visible on the tape but your face should be covered by the LCD glasses. This videotape may be used only for purposes of evaluating and improving the virtual or synthetic environment interface techniques and will not be distributed nor viewed by anyone other than members of the research team. Your name will not be associated with any data that are collected during this evaluation session.
Your rights as a participant are as follows:

1. You have the right to withdraw from the session at any time for any reason.
2. At the conclusion of your session, you may see your data, if you so desire. If you decide to withdraw your data, please inform the evaluator immediately. Otherwise, identification of data might not be possible because of our efforts to ensure anonymity.
3. You are requested not to discuss this session with other people who might be in the group from which other participants could be drawn.

Finally, we greatly appreciate your time and effort for participating in the evaluation of the virtual and synthetic environment interfaces. Remember, you cannot fail any part of this session, and there is no right or wrong answers. The session is to identify usability problems with the system. Your signature below indicated that you have read this consent form in its entirety and that you voluntarily agree to participate.

If you wish to discuss these or any other discomforts you may experience, you may call the Project Director listed in #2 of this form.

6. Potential Benefits to You or Others
Through this study, you will have an opportunity to experiment with new hardware and methods for immersive virtual computing. You will have a chance to experience new human computer interface paradigms. This may help you to assess, use, and adapt to new technologies as they are commercialized. You will also have a chance to be involved directly in research in new technology, and this will help you to understand research procedures and methods.

Additionally, you will have a chance to participate in the development of new technology. This gives you a chance to work on “next generation” computer systems, and you will have a chance to contribute to technology which may revolutionize the practice of computing for decades to come. These new methods may allow American industry to improve computer utility, interface quality, improve computer interaction in new ways, and enhance the competitive posture of American companies in key industries.

7. Alternative Procedures
There are no alternate procedures. However, you may withdraw from this study at any time.

Your participation is entirely voluntary and you may withdraw consent and terminate participation at any time without consequence.
8. Protection of Confidentiality
As a participant in this study your identity and participation will remain confidential. Your data will be identified by a number that cannot be associated with you, only by information in the study’s files. These files remain in the custody of the research team and are protected from disclosure.

Records of the tests and results from this study will be kept in offices that are normally either occupied or locked, at the testing site. Further, the test location has a security system for entry into the University of New Orleans area. In order to be admitted to the building, an individual must either have a security badge or be admitted by someone already in the building. Everyone that enters the building through the UNO entrance is viewed by a security camera, and security guards are positioned at several locations within the building. The security guards are on duty 24 hours per day, 7 days per week.

9. Signatures

I have been fully informed of the above-described procedure with its possible benefits and risks and I have given permission for participation in this study.

<table>
<thead>
<tr>
<th>Signature of Subject</th>
<th>Name of Subject (Print)</th>
<th>Date</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Signature of Person Obtaining Consent</th>
<th>Name of Person Obtaining Consent (Print)</th>
<th>Date</th>
</tr>
</thead>
</table>
Appendix 3 – Approval for Human Subjects
University of New Orleans

Campus Correspondence

Dr. Donald Barbe, PI
Syed Ahmed, Co-investigator
Dr. Kurt Satter, Co-investigator

6/22/2006

RE: Usability studies with virtual and traditional computer aided design environments

IRB#: 01jul05

The IRB committee met on 6/21/06 to discuss the series of events which occurred regarding the change in PI from Dr. Alley Butler to Dr. Donald Barbee and the use of an unapproved consent form. The committee deemed that a violation has occurred, but that this violation was a minor violation. The committee concluded that the participants were not placed at additional risk from this violation and that approval will be maintained. In other words, the data may still be used and approval remains intact.

To minimize the likelihood that such an event will happen in the future, the committee would like a member of the IRB committee to be invited to give a talk about human subjects’ research to faculty and graduate students in the coming academic year.

Please remember that any additional changes to the procedures or protocols must be reviewed and approved by the IRB prior to implementation.

If an adverse, unforeseen event occurs (e.g., physical, social, or emotional harm), you are required to inform the IRB as soon as possible after the event.

Sincerely,

Laura Scaramella, Ph.D.
Chair, University Committee for the Protection of Human Subjects in Research

cc: Dr. Keith Wismer, IRB chair, Dillard University
Dr. William Lannes, Engineering Management Department, chair
Dr. Robert Cashner, Vice Chancellor for Research
Appendix 4 – Survey used by Satter (2005)
# Usability Survey

**User ID:**

**Environment:** 19" CAD 86" CAD Stereo Glove

**Pass:** 1 2 3

<table>
<thead>
<tr>
<th>Category</th>
<th>Very Good</th>
<th>Good</th>
<th>Neutral</th>
<th>Poor</th>
<th>Very Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Navigation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Initial impression of navigational modes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Gross control movement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Speed of cursor/pointer movement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Ability to make fine adjustments to the placement of the cursor/pointer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Ability to recover cursor/pointer movements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Ease of use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 After-test impression of the navigational modes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Initial impression of the interface in locating specific parts/equipment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Ease of identification of selected part/equipment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Ability to make fine adjustments in selecting specific parts/equipment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Ease of use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 After-test impression of the location/selection mechanism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Initial impression of the interface for relocating parts/equipment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Ease of movement across the three axis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Ability to make fine part/equipment movement adjustments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Ease of use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 After-test impression of the movement mechanism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Initial impression of the overall system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Ability to relate a 2D platform to the space as presented</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 &quot;Intuitiveness&quot; of the interface - do the controls follow expected use?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Overall ease of use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 After-test impression of the overall system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Comments**
Appendix 5 – Medical Condition Questionnaire
Medical Condition Questionnaire

Please answer the following questions by circling the correct answer to the question:

Yes  No  
Are you under the care of a physician or other health care provider for any neurological (brain) disorder?

Yes  No  
Have you ever been diagnosed with Epilepsy?

Yes  No  
Have you ever had an unexplained seizure?

Yes  No  
Have you ever had an unexplained loss of consciousness?

These answers have been provided truthfully to the best of my knowledge.

_________________________________________  ___________________________  __________
Signature of Subject                        Name of Subject (Print)                  Date

_________________________________________  ___________________________  __________
Signature of Person Providing Questionnaire  Name of Person Providing Questionnaire (Print)  Date
Appendix 6 – Dillard University Flyer
Opportunity for research in virtual environments.
Experience new human computer interface technology and participate in leading edge computer science and engineering research.

For more info please contact Professor Syed Adeel Ahmed (Computer Science/Physics & Pre-Engineering)
Dent Hall # 168 sahmed@dillard.edu (504)816-4506
Vita

Syed Adeel Ahmed was born in Hyderabad, India and attended Saint Paul’s High School and Saint Mary’s Junior College in Hyderabad. Upon the completion of his Bachelor of Science degree in Electronics & Communication Engineering from Osmania University Hyderabad, India in 1997. He worked for Hindustan Aeronautics Limited in Hyderabad for one year. After that he went on to became a Microsoft Certified Professional (MCP) in 1999.

In 2001 he received a Master of Science degree in Electrical Engineering from University of New Orleans, L.A. He followed that program with a second M.S. degree in Engineering Management from University of New Orleans, L.A.

Upon the completion of his first M.S. degree, Mr. Syed began teaching full time as an Assistant Professor of Computer Science/Physics-Engineering at the Dillard University, New Orleans, L.A.

He also teaches at Southern University at New Orleans (SUNO), L.A In August of 2003; Syed Adeel Ahmed was accepted into PhD Program at University of New Orleans, L.A, where he is being awarded his PhD degree on December 16th 2006.