PEOPLE'S PERCEPTION AND ACTION IN IMMERSIVE VIRTUAL ENVIRONMENTS (IVES)

By

Qiufeng Lin

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Computer Science

May, 2015

Nashville, TN

Approved:

Robert E Bodenheimer
John J Rieser
William Thompson
Benoit Dawant
Bennett A Landman

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my advisor Bobby Bodenheimer, who has guided me patiently through my research and encouraged me all the time. Without his guidance and persistent help this dissertation would not have been possible. I also thank him for his insightful thoughts and useful advice in my research and life during the past six years, which help me accommodate my life in a foreign country, build research skills, and prepare for my future career. Bobby, thank you very much for everything I have learned from you.

I would like to thank Professor John Rieser for providing deep psychological insights and discussions during my research, which helped me avoid many mistakes in my research experiments. Discussions on the details of my research problems with him always offer me deeper understanding about my research.

I would like to thank Benoit Dawant, Bennett Landman, and Bill Thompson, for their valuable advice and guidance through my research and dissertation. Thank Alan Peters for plotting Figure IV.6 for me.

Last, I would like to thank the members of the Learning in Virtual Environments (LIVE) lab, Xianshi Xie, Haojie Wu, Erin McManus, Stephen Bailey, Gayathri Narasimham, Tim McNamara, and Aysu Erdemir for their help on valuable discussions and my experiments.

Last of all, I thank my husband, Xianshi Xie, for supporting and encouraging during many long hours. I also want to thank my parents and friends for their supporting during this journey.

TABLE OF CONTENTS

		Page
ACKNOWLEDGMENTS		**ii**
LIST OF TABLES		**vi**
LIST OF FIGURES		**vii**
I	**INTRODUCTION**	**1**
I.1	Overview	5
II	**BACKGROUND AND RELATED WORK**	**7**
II.1	Virtual Environments	7
II.2	Immersive Virtual Environment	7
II.3	Self-avatars in Immersive Virtual Environments	10
II.4	Human Distance Perception	12
II.4.1	Sensory Cues for Depth Perception	12
II.4.2	How Egocentric Distance is Perceived in the Real World	15
II.4.3	How to Measure Egocentric Distance Perception	16
II.5	Distance Perception in Immersive Virtual Environments	19
II.6	Affordances	21
III	**DISTANCE PERCEPTION**	**24**
III.1	Experiment One: Replicate Wu et al. (2004) in the real world	26
III.1.1	Participants	27
III.1.2	Method	27
III.1.3	Results and Discussion	27
III.2	Experiment Two: Distance Estimation with Indirect Walking	29
III.2.1	Participants	29
III.2.2	Method	29
III.2.3	Results and Discussion	30
III.3	Experiment Three: Distance Estimation in Our laboratory	31
III.3.1	Participants	31
III.3.2	Method	32
III.3.3	Results and Discussion	32
III.4	Experiment Four: Distance Estimation in Immersive Virtual Environment	33
III.4.1	Participants	34
III.4.2	Materials and Apparatus	34

III.4.3 Method 34
III.4.4 Results and Discussion 39
III.5 Experiment Five: Near-to-far Scanning versus No Scanning in Immersive Virtual Environment 39
III.5.1 Participants 41
III.5.2 Materials and Apparatus 41
III.5.3 Method 41
III.5.4 Results and Discussion 42
III.6 Experiment Six: Training 43
III.6.1 Participants 45
III.6.2 Materials and Apparatus 45
III.6.3 Design and Procedure 45
III.6.4 Results and Discussion 47

IV AFFORDANCE EXPERIMENTS 49

IV.1 Experiment One: Pilot Study: Stepping Over or Ducking Under a Pole in the Real World 49
IV.2 Experiment Two: Stepping Over or Ducking Under a Pole in an IVE . . . 51
IV.2.1 Participants 53
IV.2.2 Materials and Apparatus 53
IV.2.3 Design and Procedure 55
IV.2.4 Results and Discussion 56
IV.3 Experiment Three: Stepping Over or Ducking Under a Pole in the Real World 58
IV.3.1 Participants 60
IV.3.2 Materials and Apparatus 60
IV.3.3 Design and Procedure 62
IV.3.4 Results and Discussion 62
IV.4 Experiment Four: Stepping Off a Ledge in an Immersive Virtual Environment 64
IV.4.1 Participants 65
IV.4.2 Materials and Apparatus 65
IV.4.3 Design and Procedure 65
IV.4.4 Results and Discussion 66
IV.5 Experiment Five: Stepping Off a Ledge in the Real World 68
IV.5.1 Participants 68
IV.5.2 Materials and Apparatus 68
IV.5.3 Design and Procedure 69
IV.5.4 Results and Discussion 70
IV.6 Experiment Six: Two Affordance Tasks to Explore the Effect of Changing the Size of the Self-avatar 70
IV.6.1 Participants 71
IV.6.2 Materials and Apparatus 71
IV.6.3 Design and Procedure 71

IV.6.4 Result and Discussion 73

V CONCLUSION AND FUTURE DIRECTION 77

V.1 Conclusion and Discussion 77
V.1.1 Distance Perception 77
V.1.2 Affordance Judgments 81
V.2 Listed Contributions 84
V.3 Future Directions 86
V.3.1 Avatars in Immersive Virtual Environments 86
V.3.2 Training in Immersive Virtual Environments 87
V.3.3 The Effect of Action 88
V.3.4 The User Experience in Immersive Virtual Environments 88
V.3.5 Physical Objects in Virtual World Experiments 88

BIBLIOGRAPHY 89

LIST OF TABLES

Table | Page

II.1 Spatial visual cues for distance estimation (adapted from Cutting and Vishton (1995)) 13

II.2 Indirect Measurement Methods for Distance Perception in the Real World 18

IV.1 Experimental Design 52

IV.2 Quantitative Results for Experiment One (Ducking Under/Stepping Over in the Immersive Virtual Environment) 58

LIST OF FIGURES

Figure Page

II.1 This figure shows how human determine target distance d using eye height h and angle of declination θ, d = hcotanθ. 16

III.1 This figure shows dimensions and geometry of the virtual environments laboratory and hallway outside. 26

III.2 This figure shows the constant errors as a function of distance in Experiment One. The black line shows the ground truth condition for comparison. The other four lines show the small aperture (13.6°) near-to-far scanning, small aperture far-to-near scanning, large aperture (21.1°) near-to-far scanning, and large aperture far-to-near scanning methods. Error bars indicate standard errors of the mean. 28

III.3 This figure shows the constant error as a function of distance in Experiment Two. The black line shows the ground truth condition for comparison. The other four lines show the small aperture (13.6°) near-to-far scanning and large aperture (21.1°) near-to-far scanning for each walking method, direct and indirect. Error bars indicate standard errors of the mean. 31

III.4 This figure shows the constant error as a function of distance in Experiment Three. The black line shows the ground truth condition for comparison. The other four lines show the small aperture (13.6°) near-to-far scanning and large aperture (21.1°) near-to-far scanning for each walking method, direct and indirect. Error bars indicate standard errors of the mean. 33

III.5 This figure a person wearing apparatus for generating a calibrated self-avatar. 35

III.6 This figure shows the equipment for generating a calibrated self-avatar. . 35

III.7 This figure shows the female avatar in the static avatar pose viewing herself in the mirror. The stepping stones are also shown. 37

III.8 This figure shows a partial rendering of the laboratory environment where distance estimation was done. 38

III.9 This figure shows the constant error as a function of distance in Experiment Four. The black line shows the ground truth condition for comparison. The other six lines show the no avatar, still avatar and animated avatar for each walking method, direct and indirect. Error bars indicate standard errors of the mean. 40

III.10 This figure shows the collar worn around subjects' neck. 43

III.11 This figure shows the constant error as a function of distance in the immersive virtual environment experiment. Error bars indicate standard errors of the mean. 44

III.12 This figure shows the constant error as a function of distance in the avatar experiment. Error bars indicate standard errors of the mean. 47

IV.1 The physical pole apparatus used for the pilot and following experiments. 50

IV.2 First person view used in judging whether to step over or duck under the pole. 50

IV.3 Results from pilot experiment (N=6). Error bars are standard errors of the mean. The minimum ratio for ducking under was 0.52 (0.02) and the maximum ratio for stepping over was 0.53 (0.02). 51

IV.4 This figure shows a person wearing the HMD and motion capture components (left), motion capture components along (top right) and the avatar models used for the self-avatar (bottom right). 54

IV.5 This figure shows the example of the maximum-likelihood procedure in the stepping over or under a pole in no avatar condition. The Y axis is the percentage of the pole height to eye height. A red cross means the participant selected to step under the pole. A black circle means the participant indicated to step over the pole. The final threshold is 0.55. . . 57

IV.6 Threshold magnitudes in the task of stepping over or under a pole. The y-axis is the ratio of the pole height to subject's eye height. The error bars show standard errors of the mean. In the figure, "No Avatar" indicates no self-avatar was present, whereas "Avatar" indicates the presence of a self-avatar; "No Action" indicates subjects only perform a judgment, whereas "Action" indicates subjects step over or duck under the pole; "Without Pole" indicates no physical pole was present, whereas the shaded "With Pole" indicates conditions where there a physical pole was collocated with the virtual pole. The colors differentiate each group of participants, as explained in the text. There were 48 total participants, with 12 participants indicated by each color. 59

IV.7 The pole by action interaction. The y-axis shows the ratio of the pole height to the subject's eye height. The blue line indicates the no action condition, and the red line indicates the action condition. 60

IV.8 The avatar by action interaction. The y-axis shows the ratio of the pole height to the subject's eye height. The blue line indicates the no action condition, and the red line indicates the action condition. 61

IV.9 This figure shows the threshold values in stepping over or under a pole in the real world with and without action. Error bars indicate standard errors of the mean. 63

IV.10 A view that a subject might see in the immersive virtual environment for the stepping off the ledge experiment in the *no avatar* condition (left) and *avatar* condition (right). 66

IV.11 Threshold magnitudes in the two experiments involving stepping off the ledge. The Y-axis is the proportion of the ledge height to subject's eye height. The error bars show standard errors of the mean. The left two error bars (red and blue) show Experiment Four, done in a virtual environment (VE), where no avatar and avatar indicate the state of the *avatar* condition, with a self-avatar calibrated to the size of the subject. The right two error bars (magenta and black) show the results of Experiment Five, done in the real world (RW), where no action and action indicate the state of the *action* condition. 67

IV.12 The ledge used for the stepping study in the real world. 69

IV.13 Threshold magnitudes in stepping over or under a pole in the three conditions. The Y axis is the percentage of the pole height to subject's eye height. The error bars show standard errors of the mean. In the figure, no avatar means no avatar condition, avatar 1 means the avatar's size is the same as the subject's size, avatar 1.15 means an avatar with 15% longer leg length. 74

IV.14 Threshold magnitudes in walking straight or stooping under a door in the three conditions. The Y axis is the percentage of the door height to subject's eye height. The error bars show standard errors of the mean. The no avatar, avatar 1 and avatar 1.15 have the same meaning as those in Figure IV.13. 75

CHAPTER I

INTRODUCTION

This thesis examines the fidelity of head-mounted display (HMD) immersive virtual environments through two aspects, the effects of users' digital representations on space perception and perception-action judgments. We explore this fidelity by comparing the coupling between perception and action, in both immersive virtual environments and the real world.

Immersive virtual environments are used in a wide variety of situations from education (Rizzo et al., 2006), to simulation (van Wyk and de Villiers, 2009), to clinical therapy (Rothbaum and Hodges, 1999, 2000). Fundamental questions that arise in all of these tasks are how good the virtual environments are at conveying the real world situations they are intended to convey, and how can this "goodness" or fidelity be measured? For many virtual environments, it is likely sufficient that people behave similarly in the virtual environment to the way that they behave in the real world. There are many situations, for example, where it is known that people do not behave in a similar manner. For example, a large body of literature exists to show that people judge egocentric distances differently in the real-world than in virtual environments presented with head-mounted displays (Loomis and Knapp, 2003; Thompson et al., 2004; Bodenheimer et al., 2007; Jones et al., 2008; Ries et al., 2009; Grechkin et al., 2010); there are also differences in completion time and behavior when performing other simple tasks in the real world and virtual environments (Williams et al., 2007; Streuber et al., 2009; McManus et al., 2011). Williams et al. (2007) called their measure of the difference between the real world and a virtual environment "functional similarity," (Ferwerda, 2003) and it bears a resemblance to the ideas of "correlational presence" proposed by Slater et al. (2009b).

Digital representations of people are called avatars, and when they represent ourselves, self-avatars. We are particularly interested in self-avatars because of the potential issue of "fit" between a person's physical body and the virtual body. What is this fit? Requiring a body-scaled match between the person's real, physical body and a simulated body might be reasonable — although limiting — but most virtual environments do not provide any representation of a person's body in the virtual environment. In the real world, motor learning and development involve exploring the fit between the body and the environment (Thelen, 1995), and if virtual environments provide no "body" with which to explore, then it is reasonable to ask if learning is necessarily limited. However, providing a matched-size self-avatar in a virtual environment is non-trivial, requiring extra equipment and extra time on the part of the users of the virtual environment. Thus, knowing whether a matched self-avatar is needed or not for the purposes of fidelity, even in basic cases, is an important area for investigation.

The coupling between perception and action was emphasized in the work of Gibson (1979). Gibson theorized that there were properties of the environment that represented possibilities for action; he called these *affordances*. An affordance is independent of a person's ability to perceive it, but if a person perceives an affordance, then action is possible depending on the fit connecting the individual's body and the environment. Consider, for example, the decision as to whether one can step over a fence or needs to duck under it. Young children routinely crawl under low fences that most adults routinely step over. And consider the decision whether one can walk straight through the doorway in a children's playhouse or needs to duck to pass through. In each case the decisions to step over, duck under, to walk straight through or to duck under depends on the relation of the environment (the fence's height or the doorway's height) and the actor's body. Taller people can step over higher fences and need to duck under higher doorways than shorter people. In the physical world wrong decisions have consequences: people trip when stepping over fences or hit their heads when passing through doorways. People typically seem to make decisions with little or no reflection. Their decisions show that perception is generally body-scaled, that is, the dimensions of the environment are perceived in terms of their body's dimensions and what they can do (Gibson, 1966). The concept of affordance is now central in perception (Michaels, 2003) and human-computer interaction (Norman, 2002).

Immersive virtual environments provide an easy way to simulate real world scenarios and examine perception-action judgments with or without actual action, which require people to observe the virtual space in terms of their perception and capability of actions. Simulation of these judgments allow us to assess people's space perception by recording their intention to act without the actual actions. And it is easy to change size and environments in immersive virtual environments, which allow us to examine the effect of the body scale on different tasks. Currently, HMD-based immersive virtual environments are still used at the research level, but they hold the promise of widespread use in various field such as architecture, education, and training, especially with the the development of commodity level displays, such as Oculus Rift. Understanding the underlying constraints of these environments and knowing the effect of using avatars inside virtual environments should enable more effective design and use of virtual worlds. There are interesting issues that arise as the concept of affordances is applied to immersive virtual environments. In its most fundamental form, an issue arises from the fact that immersive virtual environments are constructed, not resulting directly from natural selection or other biological processes. However, as many immersive virtual environments seek to mimic or replicate the real world, then possibilities for action in the virtual environment should, insofar as possible, duplicate those in the real world. The fidelity of the virtual environment could then be measured by how well the possibilities for action in the virtual environment match the possibilities for action in the real world. The way this measurement might be accomplished will be discussed below.

Affordances are often measured for individuals in terms of critical thresholds that divide a possible action from an impossible one (Warren, 1984; Warren and Whang, 1987). This method of measuring them provides a basis for comparing a virtual environment to the real world. A modern view of affordance suggests that because of individual differences and variations, the operationalization of an affordance is best done by characterizing them as probabilistic functions that represent a person's likelihood of successful performance of an action.The advantage of the modern approach is that it allows standard procedures of psychophysics and signal detection, e.g., Green and Swets (1966), to be used to estimate the affordance function, or, if desired, the critical threshold.

This thesis first examines egocentric distance perception in the real world and in immersive virtual environments. Distance perception is a great task for perception-action coupling because of the way it is measured, through a form of blind walking. While there are other ways to look a distance perception that do not involve an action output, e.g. verbal estimates (Gilinsky, 1951; Harway, 1963). They are not considered here. We first attempt to replicate a standard distance perception experiment performed in Wu et al. (2004) to establish a baseline of people's performance using near-to-far scanning. We developed and calibrated a new form of blind walking suited to our laboratory to measure egocentric distance. The size of our laboratory only allows direct blind walking up to 5 meters. We also explored the effect of the self-avatar, the effect of the near-to-far scanning, and the effect of different training methods on distance perception in our immersive virtual environment setup.

This thesis studies perception-action affordances, and self-avatars more deeply in a different context. We study the affordances of stepping over or ducking under a pole, that of stepping straight off a ledge, and that of walking under a doorway. We examined these affordances both with and without a self-avatar in the immersive virtual environment. We performed both a perception task and a perception with action task, in both the real world and in a virtual environment for the first two affordance tasks. In the perception task individuals were asked to say whether or not they could step over or duck under the pole and step off a ledge. In the perception with action task, they were asked physically to step over or under the pole and physically to step down off the ledge. Threshold values were obtained for both affordances for both types of task. We found that there is a significant difference in individuals' perception of the threshold at which they could perform an action when they were asked to perform it rather than when they were simply asked whether they could perform it. Most importantly for virtual environments, we found that the presence of a matched-size self-avatar affects the affordance judgment, making it consistent with the real world, and thus providing critical information for people deciding what they can and cannot do in the virtual environments. To study the size effect of the self-avatar, we explored two affordances, that of stepping over or ducking under a pole, and that of walking through a doorway. We changed the leg length of the self-avatar to 15% longer and to evaluate whether change their behavior when the size of the self-avatar is changed in the two affordance tasks.

The work presented here is important because it tries to seek important factors that might affect people's interaction with virtual words. The ultimate goal of immersive virtual environments is to provide high fidelity interactions between human and virtual worlds so that people can have experiences in immersive virtual environments as if they are real. Knowing people's performance in spacial perception and perception-action judgments will enable better design of future immersive virtual environments. Examining the effect of the self-avatar on different tasks also provides us more knowledge on how to design the virtual environments. As mentioned, providing a matched-size self-avatar in a virtual environment is non-trivial, requiring extra equipment and extra time on the part of the users of the virtual environment. My work here assesses different factors in immersive virtual environments that might affect people and provide useful information in the use of self-avatars for future immersive virtual environments.

I.1 Overview

The remainder of this thesis is organized as follows. Chapter II provides background and previous research to the current work. Chapter III examines distance estimation in the real world and in immersive virtual environments. Chapter IV explores the three affordance task in the real world and in immersive virtual environments. Chapter V summarizes the contributions of the work and indicates opportunities for future research.

CHAPTER II

BACKGROUND AND RELATED WORK

II.1 Virtual Environments

Virtual reality (VR) is a computer-created environment that can simulate the real world environment. These simulations are also called virtual environments (VEs). Most current virtual environments are experienced visually, displayed on computer screens or viewed through stereoscopic displays. With the development of technology, some virtual environment simulations employ additional information, such as audio cues (Lokki and Grohn, 2005) or haptic information, like touch, smell, and so on (Sherman and Craig, 2003). People use some human computer interfaces to interact with virtual environment systems. During these interactions, virtual environments are first created by computer rendering. To interact with these environments, people can use a device like a mouse, keyboard, joysticks, motion controllers, etc., or they can interact naturalistically through gesture and movement.

II.2 Immersive Virtual Environment

When people are in virtual environments, we call them immersive virtual environments. Immersive virtual environments are different from computer-screen-based virtual environments because immersive virtual environments are designed to associate with the feeling of presence or immersion inside the environments. It is difficult to define immersion. Heim (1998) described immersion as a psychological effect that is generated from "device that isolate the senses sufficiently enough to make a person feel transported to another place." Immersion was also defined as human's perception and interaction to a virtual environment (Witmer and Singer, 1998). In that paper, immersion is "a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences." The authors also suggested several factors that might affect people's immersion inside a virtual environment:

"isolation from the physical environment, perception of self-inclusion in the virtual environments, natural modes of interaction and control, and the perception of self-movement". Slater and Wilbur (1997) defined immersion as the extent to which virtual environments are capable of delivering an inclusive and vivid illusion of reality to human's senses. Although there is no settled definition on immersion, the concept of immersion describes how people receive the sensory information, which is created by computer technology and generates the illusion of being in a real world environment.

Immersive virtual environments have been used in a variety of real world applications from education, to simulation, to clinical therapy. It is important that these environments convey the situations they are intended to convey so that people have similar experiences in the virtual environments compared to those in the real world. An important issue is how this similarity, "goodness" or fidelity is measured. For many virtual environments, it is likely that people behave similarly in the virtual environment to the way that they behave in the real world. However, there are many situations where it is known that people do not behave in a similar manner. For example, a large body of literature exists to show that people judge egocentric distances differently in the real-world than in virtual environments presented with head-mounted displays (HMDs) (Loomis and Knapp, 2003; Thompson et al., 2004; Bodenheimer et al., 2007; Jones et al., 2008; Ries et al., 2009; Grechkin et al., 2010); there are also differences in completion time and behavior when performing other simple tasks in the real world and virtual environments (Williams et al., 2007; Streuber et al., 2009; McManus et al., 2011). Williams et al. (2007) called their measure of the difference between the real world and a virtual environment "functional similarity," and it bears a resemblance to the ideas of "correlational presence" proposed by Slater et al. (2009b).

There are two common types of immersive virtual environments: large-screen projection immersive virtual environments and HMD based immersive virtual environments. Large-screen projection virtual environments use large screens around the user to generate the effect of immersion. Different large-screen projection virtual environments vary in the number, shape, and size of the large screens around the user. The Cave Automatic Virtual Environment (CAVE) is a common large-screen projection virtual environment, where there are at least four large screens around the user. The first CAVE system was introduced by Cruz-Nera et al. (1992), in which a large screen was curved around the user to guarantee that the user had equal distances from the entire screen. Another type of immersive virtual environments is produced by wearing a head-mounted display (HMD). An HMD is a helmet equipped with two small screens , one for each eye, which allows users to view rendered images stereoscopically. The HMD is usually connected to a rendering machine by cables transmitting the real-time rendering images. To enable the stereo effect, the rendering machine generates different graphics for left and right eyes simultaneously and sends them to the two small screens. The HMD is also equipped with position and orientation tracking device to update user's position and orientation inside the virtual environments.

To interact with CAVE systems, people can use a bicycle (Plumert et al., 2004), a treadmill (Mohler et al., 2004) or some haptic device, like a joystick. People can also locomote in a limited space in CAVE systems. People interact with HMD based immersive virtual environments naturally (Xie et al., 2010), use haptic device like a joystick (Xie et al., 2010), and use treadmills, either linear or omnidirectional, e.g. the cyberwalk (Schwaiger et al., 2007), etc. As mentioned before, interactions with immersive virtual environments can occur in many forms. Interaction with an IVE other than vision is not a significant part of this thesis, so we have only briefly mentioned that.

Both CAVE systems and HMD based immersive virtual environments have advantages and disadvantages. CAVE systems do not restrict people's field of view. They create a complete sense of presence in immersive virtual environments. CAVEs also allows multiple users experience virtual worlds at the same time without extra effort. However, CAVE systems are inherently large and cumbersome, costly and complex to build and maintain. Usually CAVE systems only have a single walking direction, which constrains people's exploration inside the immersive virtual environment. Recently, the omnidirectional treadmill system was introduced by Schwaiger et al. (2007), in which 25 conventional treadmills are linked together and allow people to walk or jog near naturally. The HMD based immersive virtual environments meet or exceed the performance of CAVE systems in several aspects. The cost of an HMD is much less, in addition to obvious advantages in size, complexity and portability of an HMD based immersive virtual environments. The image quality is much better. Historically, HMDs are bulky, which weighted around two pounds, and have very limited field of view. But with the development of technology, new HMDs are lighter, for example the Oculus Rift, weights only 390g and has larger field of view (Young et al., 2014). The HMD based immersive virtual environments have some drawbacks. HMDs are usually connected to cables to transmit the images. So it is not safe for people to jump or run when people are wearing an HMD. Second, the limited field of view (FOV) makes people's experience different in immersive virtual environments and in the real world. Using HMDs to experience immersive virtual environments might also cause motion sickness when people wear the display and immersive in the virtual environments for a certain period of time. This thesis focuses on HMD based immersive virtual environments.

II.3 Self-avatars in Immersive Virtual Environments

Avatars are defined as the digital representation of humans in online media or virtual environments (Bailenson and Blascovich, 2004). The first-person representations of users, which we call self-avatars, are common in virtual worlds and modern 3-D video games. But users complain about the lack of natural control on the avatars, like the body movements, locomotion, expressions and so on. Animated self-avatars were rarely used in immersive virtual environments because of the expensive equipment acquired for tracking the movements of human beings and the complexity of building such a system. The development of affordable full body motion capture systems, like the Vicon motion capture system, and powerful graphics cards which highly increase the rendering power, allow real-time tracking of people and adoption of animated self-avatars in immersive virtual environments.

Most previous research on avatars in virtual environments paid attention on social interaction and presence. This research showed that users interact with avatars representing other animate entities in a similar way to their interactions with real human beings. Personal space, which is the minimum distance a person is comfortable with when another person approach them, is one of the most popular measures of interaction explored (Hall, 1966; Hayduk, 1983). In immersive virtual environments using an HMD, social interactions become more salient because avatars can react to exhibit behaviors that correspond to user's actions appropriately (Bailenson et al., 2001; Blascovich et al., 2002; Garau et al., 2003).

The importance of viewing one's body in virtual environments has been recognized for some time. Part of this research examined the co-located avatars, which were rendered to correspond in both positions and movements to the users. Slater et al. (1995) showed that the rating of presence in the virtual environment was enhanced in some situations if the subjects reported that they had a subjective association with a virtual body that was rendered as part of the simulation. Usoh et al. (1999) showed that the presence correlated highly with the degree of association with the virtual body. They also argued that tracking all limbs in the virtual environment could increase the presence gains. The significance of the fidelity of avatars was recognized in Lok et al. (2003). They conducted a task that involved manipulating real objects while viewing a virtual simulation of the same objects. And they varied avatar's hands in the virtual environment. The results on task performance and subjective presence showed that a "believable" avatar played a more important role than the visual fidelity of the hands. User social interactions to avatars have also been examined in some recent work (Bailenson et al., 2008; Bailenson and Segovia, 2010). Other recent work involves interactions with avatars in group and dyadic situations (Hodgins et al., 2010; Ennis et al., 2010), and the ability of avatars to invoke out-of-body and phantom limbs experiences (Lenggenhager et al., 2007, 2009; Slater et al., 2008, 2009b). Besides self-avatars have been used to explore emotional and physical presence in interactions in immersive virtual environments (Dodds et al., 2010; Slater and Usoh, 1994).

II.4 Human Distance Perception

The first half of this thesis deals with distance perception, mainly focusing on distance perception in an immersive virtual environment. Distance perception is fundamental in the real world. How people perceive the distance and how to measure people's accuracy on distance perception are interesting topics for research. Perceived distances can be either egocentric or exocentric. Egocentric distance is the distance between an observer and an external object. Exocentric distance is the distance between two objects external to the observer. We first give an overview of how depth is perceived by the visual system, an enormous topic of current research in its own right. Our discussion is based on Cutting and Vishton (1995) and Thompson et al. (2004). We then discuss the body of work dealing with egocentric distance perception in the real world, a topic of direct relevance to some of the experiments performed in Chapter III. We then focus on distance perception in immersive virtual environments, where the main contribution of our work lies.

II.4.1 Sensory Cues for Depth Perception

There are spatial cues, which provide information for people to judge the depth of objects in the real world. Cutting and Vishton (1995) described a detailed information of the visual spatial cues used for space perception.

In general, there are three different kinds of cues for distance estimation: absolute distance cues, relative distance cues and ordinal distance cues. Absolute distance cues provide information on distance based on a known unit distance, like feet or meters, between two target objects. Relative distance cues provide distance information based on some ratio of distances between objects. The ordinal cues indicate the order of two objects in depth.

Visual spatial cues can be obtained through one eye or two eyes, which we call monocular viewing and binocular viewing, respectively. In binocular viewing, people get the same retinal stimulus in the two eyes. Studies showed that for distances more than about 10

Table II.1: Spatial visual cues for distance estimation (adapted from Cutting and Vishton (1995))

Visual Cue Source	Limitation
binocular disparity	limited range
binocular convergence	limited range
accommodation	very limited range
motion parallax	need to know viewpoint velocity
relative motion	need to know viewpoint velocity
height-in-field	require eye height information
linear perspective	
texture gradients	
occlusion	all distances
aerial perspective	adaptation to local conditions
familiar size	
shading	

meters, people receive the same cues in binocular viewing and monocular viewing.

There are two binocular cues, both of which are ocular-motor cues. The first is binocular disparity. Binocular disparity uses the two retinal images of the same scene from slightly different angles due to horizontal inter-pupillary distance between the two eyes. People are able to triangulate the distance to an object with a high degree of accuracy. If an object is far away from the observer, the disparity falling on the two images is small so people receive the same cues when the distance is long. If the object is near, the disparity is large. The other binocular cue is called convergence. This cue is produced by people's two eyes focusing on the same object, then they converge. Cutting and Vishton (1995) showed that convergence cues only provide absolute information for objects within a few meters.

There are other monocular cues which provide depth information for distance. The first is called accommodation. This cue is generated when people need to bring an object into focus and it provides distance information based on the amount of focusing during the process. Accommodation is only effective for distances less than 2 meters (Palmer, 1999). Relative motion and motion parallax are two additional monocular cues, which are also referred to as dynamic cues because they are related to motion. Relative motion means the relative positions of two targets provides relative distance information when the observer is moving. Motion parallax is the relative motion of the target against the background provides information bout distance when the observer or the target is moving. This can be observed clearly when driving a car. Nearby things move much faster than far off objects. The velocity of the movement is required for estimating distances when experiencing motion parallax. But it has been shown to be a weak cue for egocentric distance perception (Beall et al., 1995).

Other monocular cues are referred to as pictorial cues. These cues include linear perspective, height in visual field, texture gradient, lighting and shading, etc. Linear perspective indicates the property of parallel lines appearing to converge with the increasing distance and eventually reaching a vanishing point. Height in visual field is the vertical displacement of an object from the horizontal plane. Texture gradient is shown to provide information about the relative information for distance estimation and the shape of objects (Gibson, 1950). These three cues also belong to perspective cues. Perspective cues are really helpful to recover absolute distance when the eye-height is known (Ooi et al., 2001). In this situation, the distance can be get from the eye-height and the angular declination. Angular declination is the angle between a target object and the horizon.

Lighting and shading also give relative depth information in the form of shadows. Aerial perspective refer to the effect of distant objects have lower luminance contrast and lower color saturation due to light scattering by the atmosphere. Due to this, objects further away seems blurry than the nearby objects from a person's point of view. Relative size is another monocular cue, which provides information about the relative depth of two objects if two objects are known to be the same size but their absolute size is not known. Familiar size can be used to determine the absolute depth of an object if people have previous knowledge of the object's size. Occlusion means one object will seem closer if it is in front of the other object when the first object overlaps or partly overlaps with the other object.

Besides the visual information, there might be other information for people to judge absolute distance. Auditory cues are used to investigate egocentric distance in Loomis et al. (1998, 2002).

II.4.2 How Egocentric Distance is Perceived in the Real World

How human determine the absolute distance of objects in the real world is still an open question. A number of studies (e.g., Rieser et al. (1990); Loomis et al. (1996)) have shown that people are accurate at judging egocentric distances in the real world out to approximately 20m using blind walking. In these studies, participants first viewed a target on the floor and then blindly walked to the remembered distance without vision. One explanation of why people can perceive the egocentric distance so accurately in the real world is that human is able to use eye height information to scale the target distance (see Figure II.1 on page 16). Here, eye height and angle of declination (θ in Figure II.1 on page 16): the angle between a person's eye level and the viewing direction to the target, play important roles. If people can acquire their eye height and the angle of declination, they can easily get the distance with equation $d = h\mathrm{cotan}\theta$, where d is the distance to the target, h is the eye height and θ is the angle of declination. Although we are not sure about how people recover the absolute distance, there might be several factors contributing to determine the eye height and angle of declination during the process of visual perception: the visualization of one's feet and the target, the horizontal information got from seeing the target on the floor, people's proprioception of the horizon and gravitational vertical, people's experience in the real world. Sedgwick (1983) reported that the angle of declination provided people with enough information to judge egocentric distances if eye height information is known. And there are several studies showing that manipulating the angle of declination information can change people's judgments of distance (Ooi et al., 2001; Andre and Rogers, 2006). Ooi et al. (2001) explored the effect of angle of declination using prisms. They showed that increasing the angle of declination decreased distance estimations when people perceived distances using blind walking. The importance of integrating the near ground information to faciliate the distance estimation has been recoginized in Wu et al. (2004). In our work we conjectured that a self-avatar would help distance estimation by providing a better estimate of eye height. This conjecture appears to be false, as discussed later.

Some studies (Gibson, 1950; Sinai et al., 1998) showed that a continuous surface of the ground is important for estimating absolute distances. Disrupting the homogeneity of the ground surface by manipulating the texture gradient or introducing a gap in Sinai et al. (1998) reduced people's accuracy on distance perception in the real world. Wu et al. (2004) have shown the salience of near surface cues in estimating longer distance and suggest that ground surface texture integration was a key factor in this estimation. Their reasoning requires observers to know their eye height and then estimate the angle of declination. In their experiments, people were able to estimate distances accurately with a small field of view using the scanning method. This method integrated the ground information from their feet out to the target.

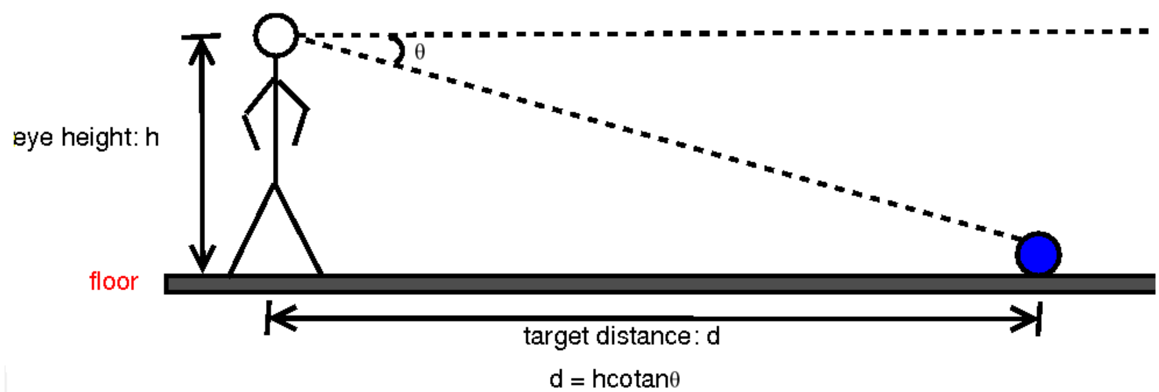

Figure II.1: This figure shows how human determine target distance d using eye height h and angle of declination θ, d = hcotanθ.

II.4.3 How to Measure Egocentric Distance Perception

Due to the complexity and cognitive nature of human's perception, there is no way to measure human's perception directly. Fortunately, we can measure people's response after the perception and use the response information to infer people's perception indirectly. This process involves controlling the sensory information perceived by a person, then measuring the observer's response, and finally inferring people's performance on the specific perception. These indirect methods need careful consideration to make sure that what is been measured really reflects what people perceive. For the response, we can have action- based or non-action-based measurements to record people's response.

For distance perception, there are several commonly used indirect measurement methods. The first is non-action-based, in which the observer is simply asked to report the distances they have perceived in a familiar unit, like feet or meters. This method is called verbal reporting (Gilinsky, 1951; Harway, 1963). Other than verbal reporting, there are several action-based methods, in which the observer is asked to do a corresponding action after perceiving the target distance. The most commonly used action-based method is called blind walking, in which people are asked to walk to or turn to the remembered target after perceiving the target distance (Rieser et al., 1990). For blind walking, there are direct and indirect methods. In direct blind walking, people are asked to walk to the remembered target distance blindly and directly after perceiving. In indirect blind walking, people walk some distance before they turn to face the target or people are led to a new starting position before they start walking. Our method is another kind of indirect blind walking method and will be described in more detail later. Some other action based methods include blind throwing (Witt et al., 2004), and pointing (Fukusima et al., 1997). A brief description of these methods are listed in Table II.2 on page 18.

Verbal reporting has been used in some studies to measure people's distance perception accuracy in the real world (Gilinsky, 1951; Harway, 1963) and they showed that distances were underestimated by around 20%. In these experiments, verbal reporting is considered to be more influenced by experimental instructions and more variable than other methods, like direct blind walking.

Direct blind walking has been used in a number of studies to determine human's distance perception accuracy in the real world. The studies showed that people are overall accurate when perceiving distances in the range of 2 to 24 meters indoor or outdoor (El-

Table II.2: Indirect Measurement Methods for Distance Perception in the Real World

Method	Description
Direct Blind Walking	Look at the target and remember the distance, then close the eyes and walk to the remembered target distance.
Triangulated Blind Walking	Look at the target, turn and walk. When told to turn, stop and turn to the target direction. Then walk or point to the target position.
Indirect Blind Walking (our method)	Our method, described in detail later. Look at the target, turn and walk for some distance to the starting position. Then walk to the target distance.
Verbal Reporting	Look at the target and judge the distance, report the distance in feet/meters.
Blind Throwing	Look at the target. Then close eyes and throw an object such as a bean bag to the remembered target location.

liott, 1987; Rieser et al., 1990; Loomis et al., 1992, 1996; Fukusima et al., 1997; Philbeck and Loomis, 1997; Loomis and Beall, 1998) in the real world. For example, Rieser et al. (1990) reported that people are able to judge the distances accurately with direct blind walking up to 24 meters in an outdoor environment.

There are several different methods in triangulated blind walking (Fukusima et al., 1997). Triangulated walking has been commonly used as an indirect measurements and has several variations. In some studies, participants view a target on the floor. Then they are asked to turn and walk forward until the experimenter asks them to turn toward the viewed target, and continue walking until they think they stand on the viewed target. Some other variations of triangulated walking involve that participants do not walk the rest of the way to the target. They are asked to point to the target location by either physically pointing or turning to face the target direction. Our work uses an indirect walking method that does not involve triangulation, in which subjects are led to a new starting position before they walk to the remembered target distance.

Size judgments is another indirect measurement method for distance perception. Gilinsky (1951) argued that perception of size and distance co-vary with each other. Then one can measure the size of the perceived target and then calculated the target distance based on that.

II.5 Distance Perception in Immersive Virtual Environments

Egocentric distance perceptions are quite accurate in the real world (Rieser et al., 1990; Loomis et al., 1996). But replicating the same experiments as discussed previously in the virtual environments shows different results. People typically underestimate the distances by 20-50% (Witmer and Sadowski, 1998; Thompson et al., 2004; Knapp and Loomis, 2004; Jones et al., 2008; Ries et al., 2009; Grechkin et al., 2010; Mohler et al., 2010). The complete reasons for the underestimations are still not clear. A number of individual factors, which might alter distance cues, have been rejected in isolation, e.g., image quality (Thompson et al., 2004; Messing and Durgin, 2005), display field of view (Creem-Regehr et al., 2005; Knapp and Loomis, 2004; Willemsen et al., 2009), stereoscopic (Willemsen et al., 2008), and the mass and inertia of HMDs (Willemsen et al., 2009). Other than research on distance cues, there is research on transitional environments, in which a virtual replica of the real world environment was reported to improve people's distance perception significantly compared to when people are placed in a completely new virtual environment (Interrante et al., 2007). There is another set of research that shows that walking through the virtual environment with continuous feedback causes dramatic improvement of post-interaction distance judgments in immersive virtual environments (Waller and Richardson, 2008; Mohler et al., 2006; Richardson and Waller, 2005, 2007; Kelly et al., 2013, 2014).

Additionally, the amount of compression reported by researchers varies widely, with some researchers experiencing very little distance compression (Jones et al., 2008) and some researchers finding no distance compression at all (Interrante et al., 2007). If the method of estimating distance used in virtual environment changes, then some of these factors, e.g., image quality, have an effect on distance judgments (Kunz et al., 2009). In the real world, Wu et al. (2004) showed the salience of near surface cues in estimating distance and suggest that ground surface texture integration was a key factor in this estimation. With a scanning method, which was called near-to-far scanning, people can judge the distances accurately. We replicated this experiment in the real world and our virtual environments, and the results are discussed in Chapter III. In the real world, we did not find the significance of scanning as that in Wu et al. (2004). In our virtual environment, people did not underestimate distance. Draper (1995) found equivocal results when investigated the effect of a self-avatar on distance perception and spatial orientation tasks. The presence of the avatar had no effect on a search and replace task, which might be caused by the overall high performance on these tasks. The most recent tasks involving using motion capture and high quality avatars has shown that the performance of distance perception in virtual environments was improved by being able to see a rendered virtual self-avatar (Ries et al., 2008, 2009; Mohler et al., 2010; Phillips et al., 2010), although other work has not found such an effect (McManus et al., 2011). But the results in McManus et al. (2011) showed that including character self-avatars or avatar animations before or during task execution was beneficial to performance on some common interaction tasks within the immersive virtual environments. And there is advantage of the third-person viewing of the self-avatar on distance perception (Mohler et al., 2010).

Moreover, people tend to have different behaviors when actual actions are involved. For example, people underestimated distances in the real world using imagined walking (Grechkin et al., 2010). But when people use blind walking, they can perceive the distances accurately. Milner and Goodale (2008) showed that people have two visual systems: vision for perception and vision for action. These two systems work differently and will affect people's behaviors in some tasks, like grasping. In their work, they showed that when people are not required to grasp objects, sometimes they cannot judge the size or dimension of the objects accurately. But when they are required to grasp, people calibrate their action to grasp and they can decide whether they can grasp it or not accurately. Learning the effect of the action will help us understand the effect of the self-avatar in immersive virtual environments.

II.6 Affordances

There is a large body of work on affordances in the psychology literature. In this section we discuss those papers that are particularly relevant to work we present in this thesis. The first empirical study of affordances was done by Warren (1984), who operationalized the concept in terms of body-based measurements, studying stair climbing and the affordance of climbability (climbable-unclimbable). Warren and Whang (1987) provided further evidence that affordances are based on body-scale information, studying the ability of people to pass through apertures. Of particular note in that paper is that Warren and Whang designed experiments with conditions in which subjects merely reported what they would do — whether they would walk through an aperture — versus actually walking through an aperture. They found that subjects were more conservative about the passability of an aperture when they actually walked through the aperture rather than when they had to report whether they would. These experiments are similar to the experimental conditions that we will have in our experiments.

Additional lines of research indicate different psychophysical functions when people are asked to look in order to make perceptual identifications or other judgments versus when they are asked to look in order to produce a visually guided action (Goodale and Milner, 1992; Milner and Goodale, 2008). The research shows that in the physical world that people tend to make more accurate judgments when they are asked to act on what they see relative to when they are asked simply to say what they see. Thus, this line of research suggests that a condition involving action would be desirable for comparing critical thresholds of affordances.

Warren (1984) noted that size of relevant body parts was an important factor in affordance thresholds. Pufall and Dunbar (1992) studied stepping onto and stepping over for children, reporting that the maximum height of negotiable obstacles was proportional to leg length for children aged 6 to 10 years. Kretch and Adolph (2013) studied the behavior of young walkers at visual drop-offs and also found that affordance thresholds were related to leg length. Stefanucci and Geuss (2009, 2010) conducted a series of studies and found that perceptual properties are biased by perceived body-based affordances, e.g., in one of their experiments, broad-shouldered subjects underestimated the width of an aperture compared to narrow-shouldered subjects.

This area of work emphasizes the importance of body-based measures in affordance judgments. Whether these measurements are determined visually or kinesthetically (or by what proportion) is a question not easily settled through real-world experiments, where subjects are embodied in their own bodies. In virtual environments, the visual cues can be manipulated. Lin et al. (2012) looked at the affordance of stepping over or ducking under, with the size of the self-avatar was manipulated. That work found that the affordance judgment closely tracked the visual manipulation. However, that work was only done in the virtual environment, and important questions about behavior in the real world were left open. Work in the real world that approaches answers to these questions can be found in that of Mark and colleagues (Mark, 1987; Mark et al., 1990), who manipulate eye height in their studies of sitability and stair climbing. This work noted eye height's importance as a preferred body-based measure in determining whether an action was possible or not. Stefanucci and Geuss (2010) altered subjects' perceived heights by having them wear a helmet or shoes in a passability experiment similar to our own where subjects had to duck under a horizontal barrier. The work presented in this thesis does not manipulate the size of the self-avatar, however, but focuses on a comparison between the virtual environment and the real world, and on the the possible value added to perception and action judgments, questions left unanswered by Lin et al. (2012).

Affordances in immersive virtual environments have been less studied, although they have been recognized as a component of presence (Zahorik and Jenison, 1998), and their potential utility in the design process noted (Smets et al., 1995; Flach and Holden, 1998; Gross et al., 2005). Dalgarno and Lee (2010) studied learning affordances for 3D virtual learning environments. Lepecq et al. 2009 examined the affordance of walking through a virtual aperture of variable widths (cf., Warren and Whang (1987)), finding a similarity between the real and virtual worlds. Geuss et al. (2010) also examined perceived passability through two poles in both the real and virtual worlds, finding that the affordance judgments between the two were not significantly different. Regia-Corte et al. (2013) examined the affordance of standing on a slanted surface in a virtual environment. They found that subjects perceived the affordance of standing and were able to make a judgment as to whether they could or could not have an upright stance in such an environment. Lin et al. (2013) examined the affordance of stepping off a ledge in the virtual environment. With the exception of Lepecq et al. (2009), none of this work has had an action component, and none has compared subjects' performance to performance in the real world. Here, we do both, and extend our prior results for stepping off a ledge with a comparison to real world performance.

The task of stepping off of a ledge is modeled on the visual cliff of Gibson and Walk (1960). The visual cliff is one of the more compelling virtual environments in terms of the sense of presence that subjects feel in it (Slater et al., 1995; Usoh et al., 1999; Meehan et al., 2002), and has been used extensively in testing such things (Zimmons and Panter, 2003; Slater et al., 2009a).

CHAPTER III

DISTANCE PERCEPTION

The body of work in this chapter was motivated by a simple hypothesis. As we have seen, self-avatars improve distance estimation (Ries et al., 2008, 2009; Mohler et al., 2010; Phillips et al., 2010). One way they might do that is by giving people a better estimate of their eye height h (see Figure II.1) by providing a body size cue to reference from. A promising way to begin to test this hypothesis force people to look down at their self-avatar as they do a distance determination. A reasonable way for them to do this is in the near-to-far scanning procedure of Wu et al. (2004) used for their well-known distance estimation experiments. We thus designed a series of experiments to test this chain of reasoning. As we will see, the entire framework falls apart, but novel and important things are learned about distance judgments in immersive virtual environments.

The near-to-far scanning method used in Wu et al. (2004) is described as the following: subjects were asked to wear a pair of goggles, which had a monocular rectangular aperture in the right eye position that could be open and closed. When open, the goggles would have a specified vertical FOV of 21.1°. During the scanning, subjects were asked to look down first, open the goggle and scan out from their feet to the target. Then they would close the goggle, reposition their heads to the look-down position, open the goggle, scan out to the target again, and close the goggle.

We designed three experiments in the real world to provide a baseline of people's performance on distance perception using this near-to-far scanning. In Experiment One, we replicated the experiment four in Wu et al. (2004). To be consistent, we had two scanning methods (near-to-far and far-to-near) and two FOVs (13.6° and 21.1°). We also had the same number of subjects and target distances as that in the paper. In Experiment Two, we eliminated the far-to-near scanning method but kept the two vertical FOVs (13.6° and

21.1°). For the blind walking, we added a new method of blind-walking besides the *direct blind walking*, which we call *indirect blind walking*, motivated by the geometry of our IVE laboratory as depicted in Figure III.1 on page 26. The diagonal distance of our IVE laboratory permits direct blind-walking up to distances of about 5m. At longer distances, subjects began to remark about their fear of walking into the wall and these remarks likely bias their judgments. The hallway outside the IVE laboratory extends for more than 30m, and blind walking experiments can be conducted in the hallway. The first two experiments were conducted outside in a campus environment on a large grass field to compare the two scanning methods (near-to-far and far-to-near), the two FOV conditions (13.6° and 21.1°) and the two walking conditions (*direct blind walking* and *indirect blind walking*). We mapped the geometry of our IVE laboratory and the hallway to that in the outside environment. The third experiment repeated Experiment Two and was done in our IVE laboratory (not in an IVE). This repetition was to see the effect of changing the experimental environment, which has been shown to have significant impact on distance estimation in some studies (Lappin et al., 2006), but not in others (Bodenheimer et al., 2007).

After the real world experiment with the near-to-far scanning method by indirect walking, we performed these experiments in an immersive virtual environment, manipulating self-representation, to see what effect it has on distance perception. In the real world experiments, we knew that the larger field of view was better for people to perceive distance. So we eliminated the small field of view condition in our virtual environment experiment. In Experiment Four, we had two walking methods (*direct blind walking* and *indirect blind walking*). We also designed three self-avatar conditions (*no avatar*, *still avatar*, and *animated avatar*) in each of the walking methods.

As we will see, the results of these experiments are somewhat perplexing in that no significant distance underestimation occurred. We conjectured that it might be because of the scanning condition, and we further hypothesized about the "training" regimens that people were using to accommodate themselves to having a self-avatar in the immersive virtual environments. Thus Experiment Five has two conditions: *scanning* and *no scanning* using the indirect walking.

Experiment Five told us that the near-to-far scanning method significantly improved people's performance on distance perception. Experiment Six examined the effect of different training methods and the effect of self-representation. We had four conditions here: avatar with no training, avatar with training similar to Mohler et al. (2010), avatar with training similar to Experiment Four and no avatar with training similar to Experiment Four. We find a significant effect of training, but no effect of self-avatar.

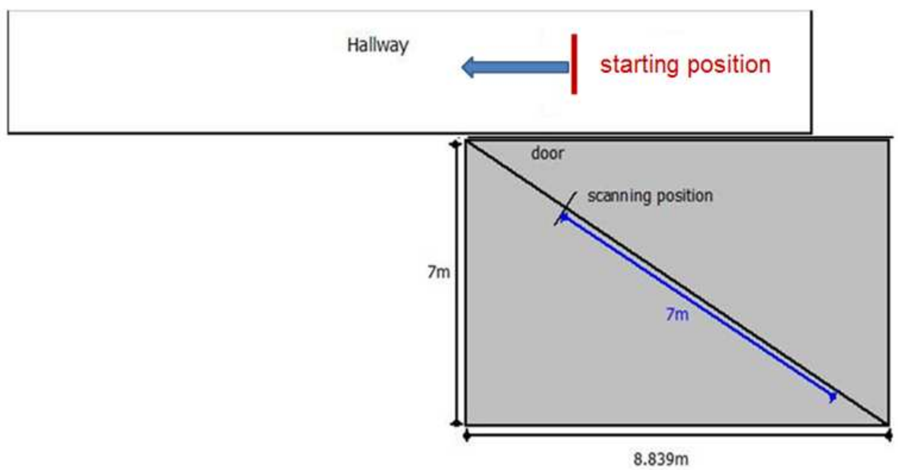

Figure III.1: This figure shows dimensions and geometry of the virtual environments laboratory and hallway outside.

III.1 Experiment One: Replicate Wu et al. (2004) in the real world

In the first experiment, we attempted to replicate the results of one of the experiments in Wu et al. (2004) (their Experiment Four), to establish a baseline for performance in further experiments.

III.1.1 Participants

Eight right-eye dominant subjects, four male and four female, participated in this experiment. These participants all had normal vision or corrected-to-normal vision and were not familiar with the experiment. Subjects were compensated for their time at the rate of $10 per hour.

III.1.2 Method

The method is identical to the Experiment 4 of Wu et al. (2004). There were four conditions: two vertical fields of view (FOVs), 13.6^{o} and 21.1^{o}, and two scanning methods (near-to-far and far-to-near). Four target distances (4m, 5m, 6m and 7m) were used in each condition and each distance was tested twice in a randomized way. In each condition, subjects were asked to wear a pair of goggles, which had a monocular rectangular aperture in the right eye position, that could be open and closed. When open, the goggles would have the specified FOV (13.6^{o} or 21.1^{o}). The subjects were guided to the starting position and instructed to scan the target using one of the scanning methods. In the near-to-far scanning, subjects were asked to look down first, open the goggle and scan out from their feet to the target. Then they would close the goggle, reposition their heads to the look-down position, open the goggle, scan out to the target again, and close the goggle. In the far-to-near scanning, subjects looked ahead first. Then they were instructed to open the goggle, scan down to the target and continue until they saw their feet. Then they closed the goggle, reposition their heads to the look ahead position, scan down again, and close the goggle. After the scanning, the subjects blindly walked to the remembered distance. The location of each trial was jittered after each trial to avoid the acquisition of local features that might provide reference cues.

III.1.3 Results and Discussion

Figure III.2 on page 28 illustrates the results of true distance versus constant errors for each of the conditions. A repeated measures analysis of variance (ANOVA) on the constant

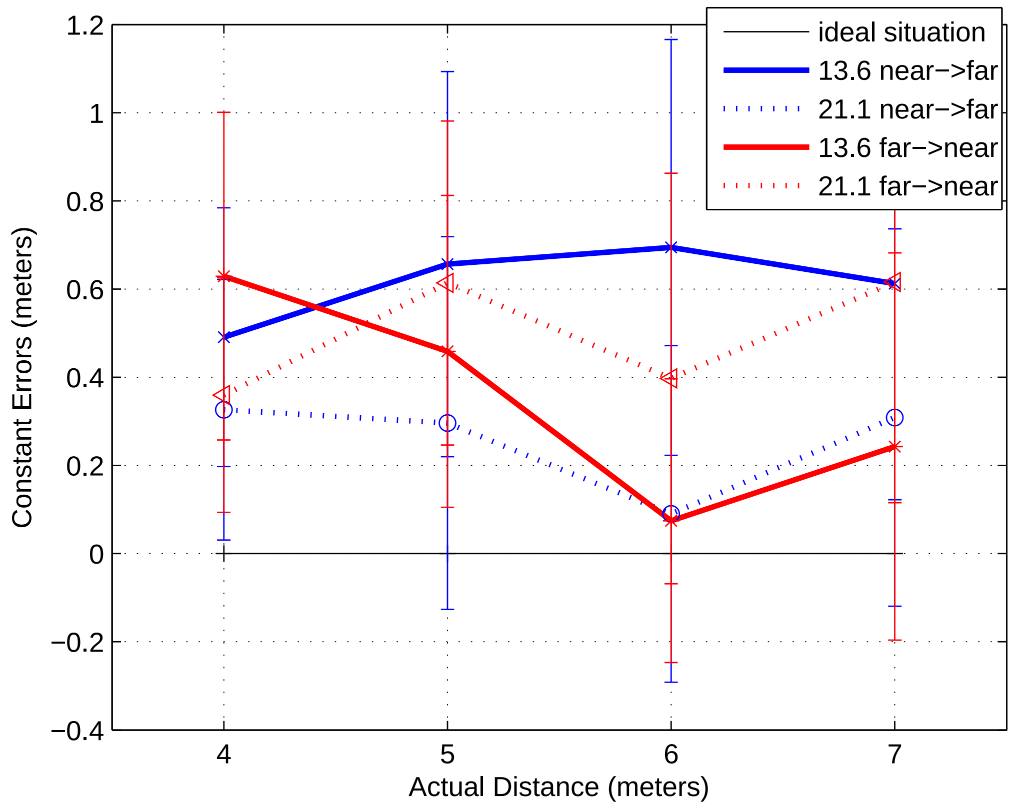

Figure III.2: This figure shows the constant errors as a function of distance in Experiment One. The black line shows the ground truth condition for comparison. The other four lines show the small aperture (13.6°) near-to-far scanning, small aperture far-to-near scanning, large aperture (21.1°) near-to-far scanning, and large aperture far-to-near scanning methods. Error bars indicate standard errors of the mean.

error examining the effect of scanning method (near-to-far, far-to-near) and aperture (small, large) on distance estimation at 4m, 5m, 6m, and 7m, did not reveal any significant effect or interactions.

Thus, we do not replicate the results of Wu et al. (2004). In particular, we do not find distance underestimation in the far-to-near scanning method. Distance estimation is accurate, with a trend towards over-estimation. Correspondence with one of the authors of the Wu et al. (2004) study did not reveal any methodological differences.

III.2 Experiment Two: Distance Estimation with Indirect Walking

Regardless of our failure to replicate Wu et al. (2004), the near-to-far scanning worked as advertised and our intention to try this method in our immersive virtual environment remained, to test calibration of eye height. In this experiment, again done outdoors, we added a different blind-walking condition and compared it to direct blind-walking. We added this blind-walking condition because, as explained in the following, the geometry of our IVE laboratory will not permit direct blind walking to distances beyond about 5m. We performed this experiment outdoors to compare it closely with Experiment One. There is some evidence of distance judgments being affected by the environmental context (Lappin et al., 2006).

III.2.1 Participants

Eight right-eye dominant subjects, four males and four females, participated in this experiment, with ages ranging from 18 to 35. These participants all had normal vision or corrected-to-normal vision and were not familiar with the experiment. Subjects were compensated for their time at the rate of $10 per hour.

III.2.2 Method

This experiment was conducted outside in a campus environment on a large grass field. We repeat the method of Experiment One, with the following changes. We eliminate the far-to-near scanning condition. Subjects did not perform differently in either scanning condition, but in Wu et al. (2004), they were more accurate in near-to-far scanning, so we retain that method. We used the walking method of *indirect blind walking* as mentioned. In this experiment, there were four conditions: two vertical fields of view (FOVs) (13.6° and 21.1°), and two walking methods (*direct blind walking* and *indirect blind walking*). Four target distances (4m, 5m, 6m and 7m) were used in each condition and each distance was tested twice in a randomized way. In each condition, the subjects were instructed to wear the pair of the goggles, stand at the starting position, scan the target using the near-to-far scanning method, and blindly walk to the remembered distance using one of the walking methods.

Subjects performed two methods of blind walking. The first was *direct blind walking* as that in Experiment One. The second was *indirect blind walking*, where they perceived a distance, shuttered the goggles they were wearing, and were guided to a starting position, where they blindly walked to their estimate of the distance. In the *indirect blind walking* condition, they were accompanied by a sighted guide at all times to insure walking in a straight path, but the guide was unaware of the distance that the subject has perceived. There are potential drawbacks with this method. First, there is a delay between the time the distance perception is done and the blind-walking action is taking. However, Rieser et al. (1990) found that such delays caused no errors in direct blind walking. Second, the turning and walking to the start position may prime the sensorimotor system in a way that causes inaccuracies. The purpose of this experiment is to compare indirect blind walking to direct blind walking.

III.2.3 Results and Discussion

Figure III.3 on page 31 shows the results of constant error versus true distance for each of the conditions. A repeated measures ANOVA on the constant errors examining the effect of walking method (direct, indirect) and aperture (13.6° and 21.1°) on distance estimation at 4m, 5m, 6m, and 7m did not find main effects of walking method, or aperture. We examined the constant errors and found that the amount of distance overestimation was significant, $t(7) = 3.31$, $p = .01$. Participants overestimated distances by approximately 10%.

People had similar performance on distance estimation in direct blind walking and indirect blind walking methods in outdoors in the real world. This indirect blind walking method allows people to do distance estimation tasks with much longer distances.

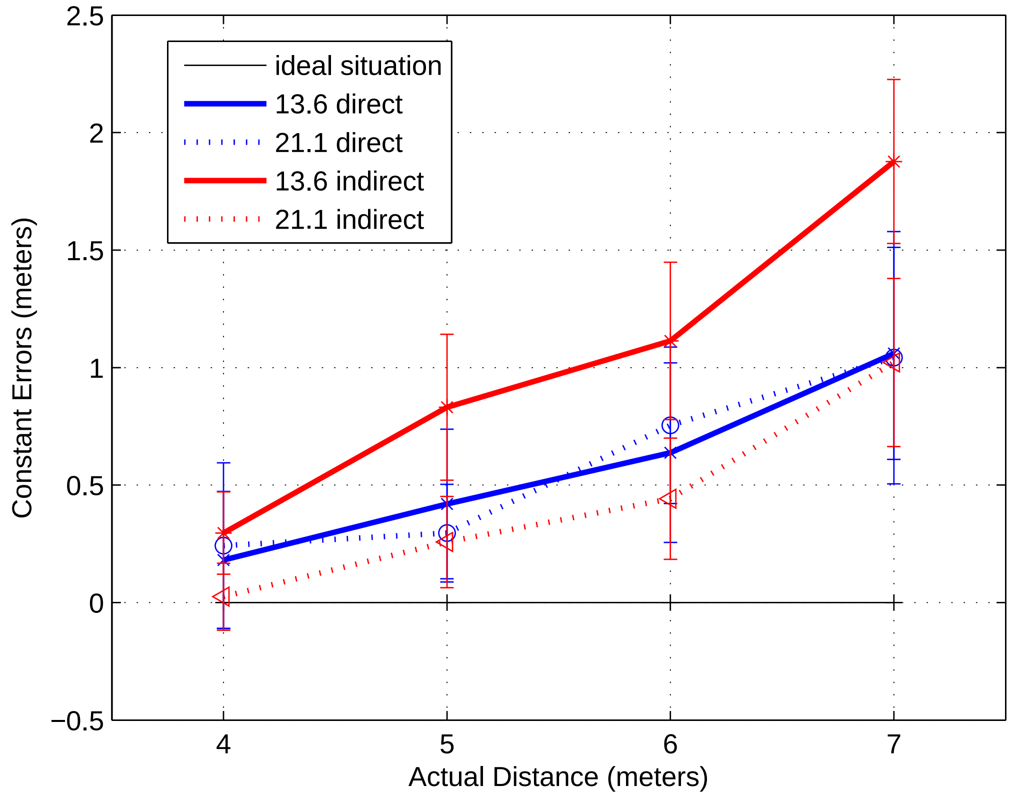

Figure III.3: This figure shows the constant error as a function of distance in Experiment Two. The black line shows the ground truth condition for comparison. The other four lines show the small aperture (13.6°) near-to-far scanning and large aperture (21.1°) near-to-far scanning for each walking method, direct and indirect. Error bars indicate standard errors of the mean.

III.3 Experiment Three: Distance Estimation in Our laboratory

This experiment repeats Experiment Two, but in our immersive virtual environment laboratory (*not* in an immersive virtual environment). The reason for this repetition is to understand the effect of changing the environment, which has been shown to have significant impact on distance estimation (Lappin et al., 2006).

III.3.1 Participants

Eight right-eye dominant subjects, four males and four females, participated in this experiment, with ages ranging from 18 to 35. These participants all had normal vision or corrected-to-normal vision and were not familiar with the experiment. Subjects were compensated for their time at the rate of $10 per hour.

III.3.2 Method

In Experiment Three, we repeated the method of Experiment Two in our immersive virtual environment laboratory approximately 8.9m $\times$ 7m $\times$ 4m, with the following changes. We eliminated the 6m and 7m in the direct blind walking condition, since our laboratory can only permit direct blind-walking up to distances of about 5m as described and shown in Figure III.1 on page 26. Since we have an unbalanced design, we balanced the orders of indirect and direct walking trials by mixing them such that the average incidence of encountering an indirect walking trial or a direct walking trial was the same. That is, because there were fewer of them, the direct walking trials occurred in internal order in the experiment.

III.3.3 Results and Discussion

Figure III.4 on page 33 shows the results of constant error versus true distance for each of the conditions. A linear mixed model was run on our unbalanced data and showed the significant effect of walking method: $F(1,77) = 5.513$, $p = 0.021$. No other effects or interactions were significant. The constant errors for direct blind walking at 4m and 5m and the constant errors for indirect blind walking for 4m, 5m, 6m, and 7m were not significantly different from zero.

The amount of overestimations are 10% in direct blind walking and 12% in indirect blind walking. Although there is significant difference between the two walking methods, the constant errors in both conditions are not significantly different from zero, which suggests that the distance estimations in our experiment are reasonably accurate and we can use both methods to measure people's distance perception. The indirect blind walking method allows us to perform distance estimation with much longer distances.

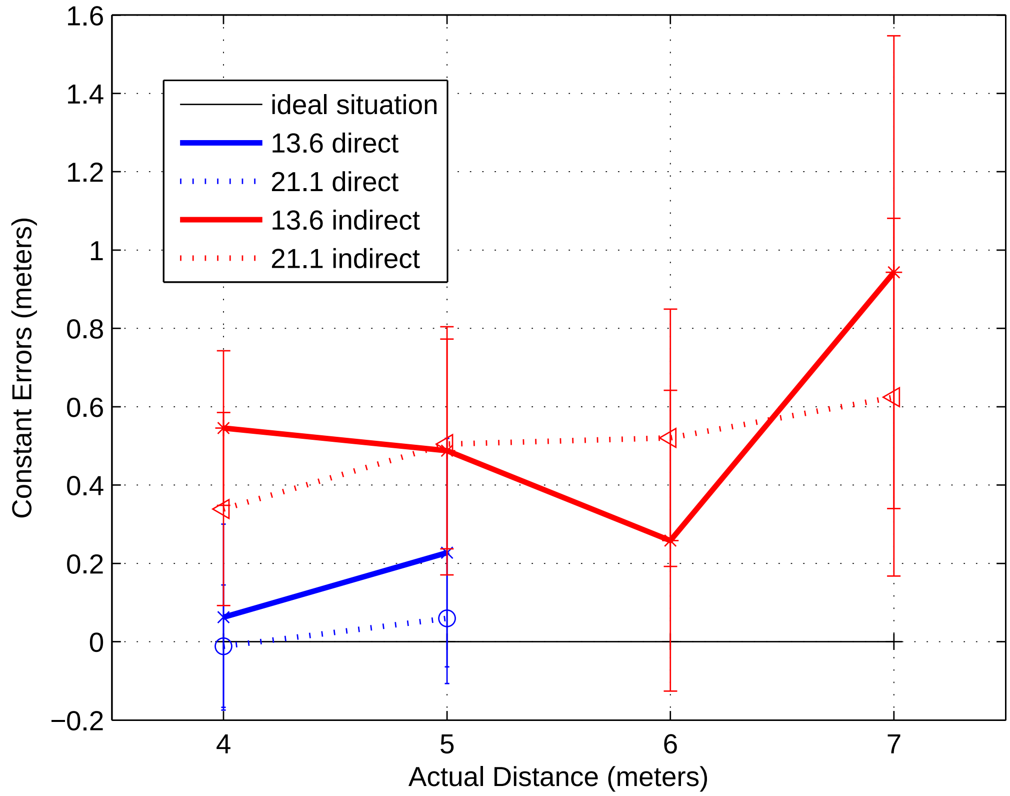

Figure III.4: This figure shows the constant error as a function of distance in Experiment Three. The black line shows the ground truth condition for comparison. The other four lines show the small aperture (13.6°) near-to-far scanning and large aperture (21.1°) near-to-far scanning for each walking method, direct and indirect. Error bars indicate standard errors of the mean.

III.4 Experiment Four: Distance Estimation in Immersive Virtual Environment

The second and third established a baseline for performance in distance estimation using modes of direct and indirect blind walking that can be done in our immersive virtual environment laboratory. We find that subjects do not underestimate distances and are generally accurate within about 10%. We will now perform these experiments in an immersive virtual environment, manipulating self-representation, to see what effect it has on distance perception.

III.4.1 Participants

Twelve right-eye dominant subjects, six male and six female, participated in this experiment. These participants all had normal vision or corrected-to-normal vision and were not familiar with the experiment. Subjects were compensated for their time at the rate of $10 per hour.

III.4.2 Materials and Apparatus

We used an eight-camera Vicon (Los Angeles, CA) MX-F40 optical tracking system and Tracker (v. 1.0) for real-time motion capture of subjects. Subjects wore six components placed on the head, waist, right hand, left hand, right and left feet (see Figures IV.4 on page 54 and III.6 on page 35). Motion data was transmitted to a second machine running Motionbuilder (Autodesk, San Rafael, CA), which mapped the motion data to a calibrated character using inverse kinematics and sent the resulting character data to the immersive virtual environment rendering machine. The immersive virtual environment was rendered using Vizard (Worldviz, Santa Barbara, CA) and viewed through a full color stereo NVIS (Reston, VA) nVisor SX head mounted display(HMD) with 1280×1024 resolution per eye, with manufacturer's specification of a field of view of 60° diagonally, and a frame rate of 60Hz. Suiting a subject in the motion capture components and calibrating the virtual avatar to their body size took approximately 15 minutes.

III.4.3 Method

In this experiment, we repeat the method of Experiment Three in our immersive virtual environment with the following changes. We have three avatar conditions (no avatar, still avatar, and animated avatar) for each subject. In order to keep the experiment time of each subject similar to previous experiments, we eliminate the small field of view condition. From Experiment Three, we know that subjects performed better in the large field of view condition.

Figure III.5: This figure a person wearing apparatus for generating a calibrated self-avatar.

Figure III.6: This figure shows the equipment for generating a calibrated self-avatar.

For each avatar condition, subjects participated in a 2-minute training phase that involved walking on stepping stones and reaching for boxes on a virtual grass field. Regardless of avatar condition, subjects were directed to walk on the stepping stones on the floor and touch the red boxes that appeared at different heights and positions. A virtual mirror was placed in front of subjects so that they could observe the immersive virtual environment either by looking down at the ground or forward in the mirror. Figure III.7 on page 37 shows an avatar reflected in the mirror in this environment. Subjects were gender matched to the self-avatar, male to male and female to female.

In the animated avatar condition, subjects could use their virtual hands to touch the virtual boxes; the boxes changed color once they were touched. In the static avatar condition, the avatar was in the pose shown in Figure III.7 on page 37, aligned with the subject's position at the start of the training session, but did not move thereafter; no feedback about when boxes were touched was given. No movement was visible in the no avatar condition, and no feedback about when boxes were touched was given in this condition, either. In the training session, subjects used the HMD with binocular views and with the default FOV of the HMD to observe the IVE. During the experimental task, we switched to an immersive virtual environment the same size and roughly the same appearance as our laboratory (see Figure III.8 on page 38. The image in the left eye was blacked out and the FOV of the right eye was set to 21.1°.

This phase was conducted with the idea of giving subjects body-based information in relation to the virtual environment they are experiencing with the self-avatar. Prior work involving self-avatars has also used such a training phase, with details, possibly significant (Mohler et al., 2010; Phillips et al., 2010), involving length of time and whether the subject was allowed to move differing. Feedback and interaction in immersive virtual environments can change distance estimates (Richardson and Waller, 2005; Mohler et al., 2006; Richardson and Waller, 2007), and this fact is something to be aware of, however this prior work had been more directly focused on task performance.

The procedure in this experiment was conducted identically to Experiment Three except subjects scanned and viewed the distance in an immersive virtual environment wearing an HMD. We blocked on avatar condition and conducted all orders of three conditions across the twelve subjects for a complete within-subjects design. After scanning and viewing the target for the second time, the HMD was blanked and the subjects were instructed to close their eyes. At this point the HMD was removed and subjects put on a blindfold that they were wearing around their neck. In the direct walking condition, the time for the change from HMD to blindfold took about 15s per trial to start with and decreased quickly to about 10s per trial. In the indirect walking condition, the time for this change from HMD to blindfold and then waling to the starting position outside the immersive virtual environment laboratory started at approximately 30s per trial but then decreased to about 15s per trial. All of these times are within delays for distance estimation which Rieser et al. (1990) found did not affect distance judgments.

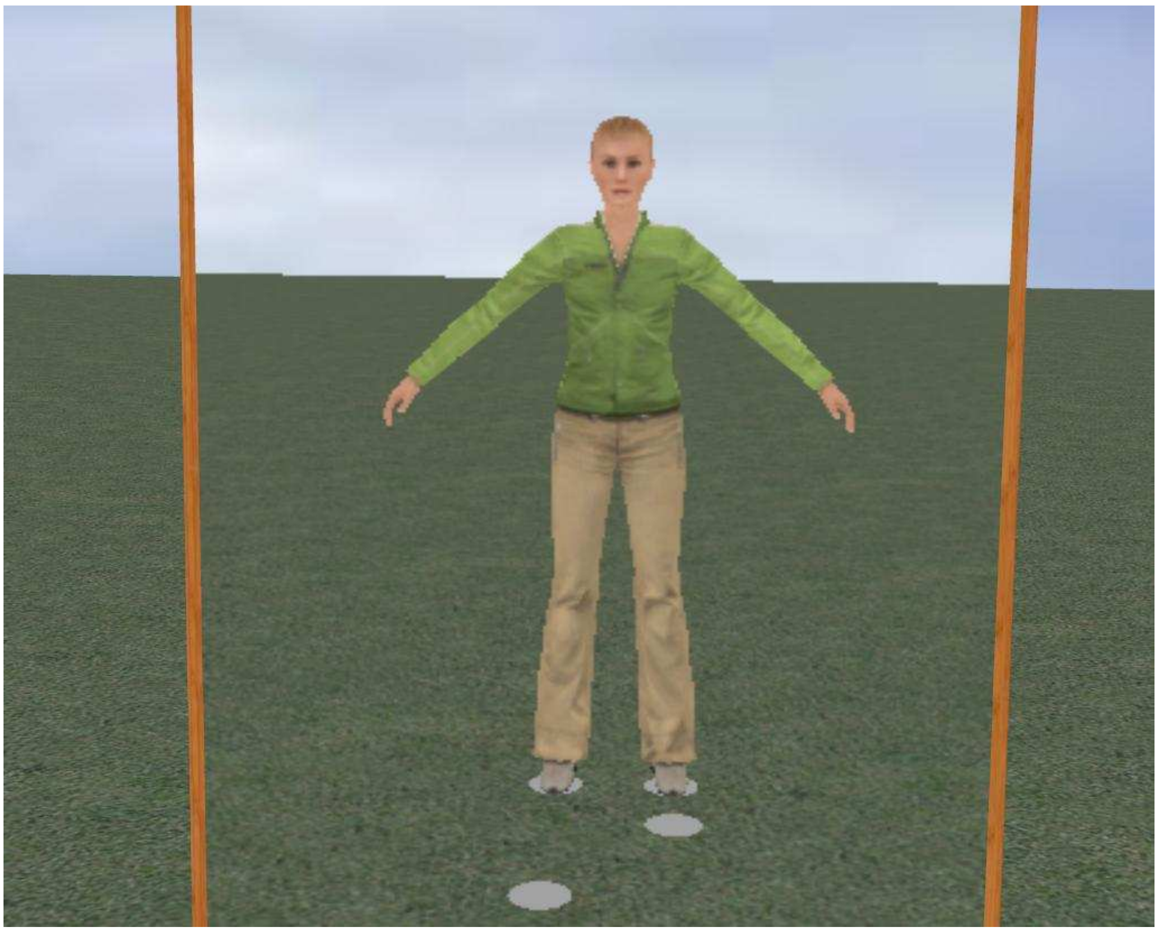

Figure III.7: This figure shows the female avatar in the static avatar pose viewing herself in the mirror. The stepping stones are also shown.

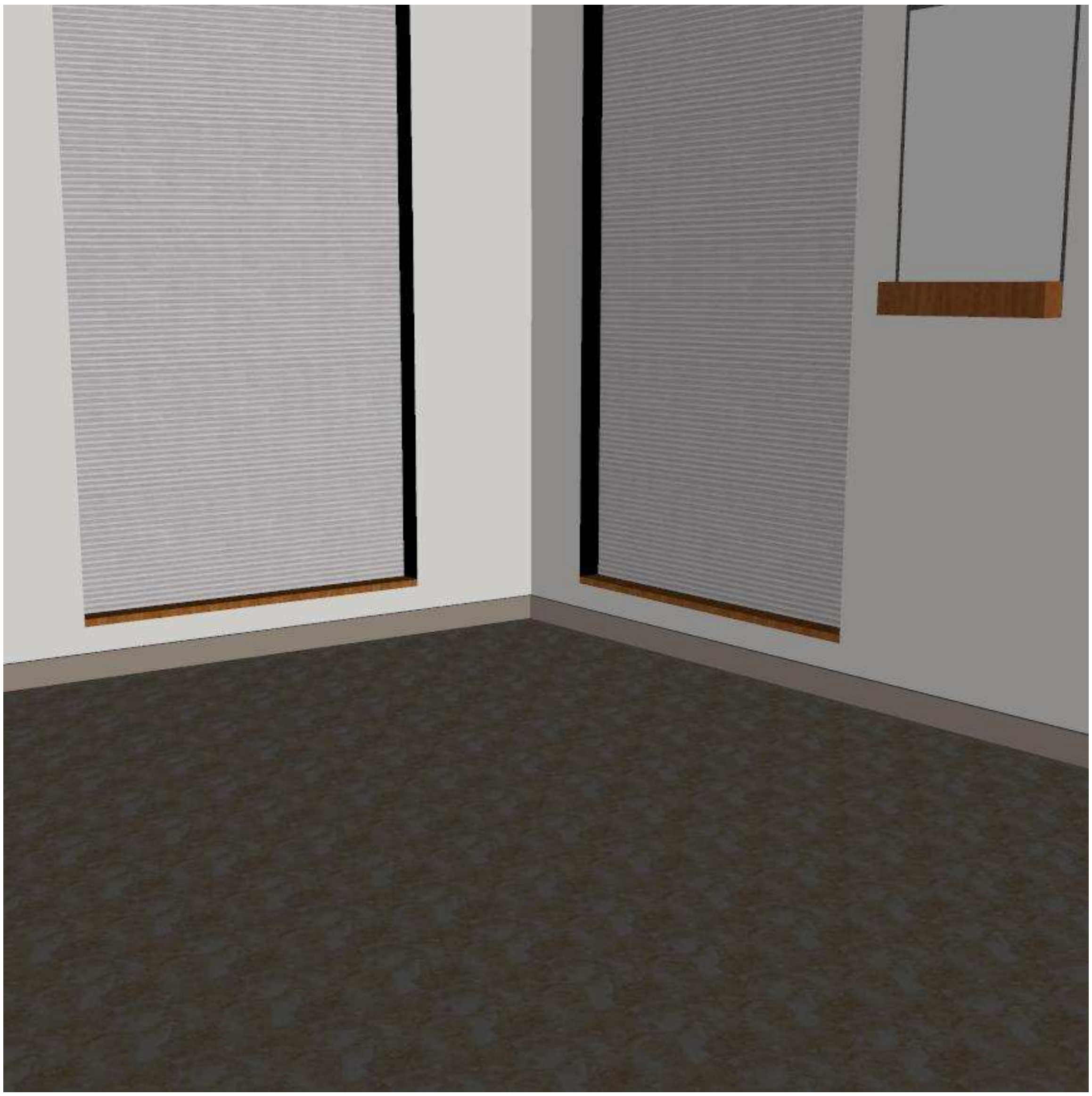

Figure III.8: This figure shows a partial rendering of the laboratory environment where distance estimation was done.

III.4.4 Results and Discussion

Figure III.9 on page 40 shows the results of constant error versus true distance for each of the conditions. A linear mixed model analysis examined the constant errors on the effect of avatar (none, still, animated) and walking method (direct, indirect) on distance estimation at 4m, 5m, 6m, and 7m revealed a main effect walking method, $F(1, 187) = 11.66$, $p < 0.001$ and a main effect of distance, $F(3, 187) = 5.71$, $p = 0.001$. No other effects or interactions were significant. Participants had smaller constant errors at estimating distances in the direct walking condition. The constant errors for direct blind walking at 4m and 5m were not significantly different from zero. However, the constant errors for indirect blind walking were significant from zero, $t(11) = 2.55$, $p = 0.03$. In the indirect blind walking case, subjects overestimated the distance by 20.6%.

As our results show subjects were accurate in direct walking condition. It is not surprising that there is no effect of the self-avatar because of this accuracy. Likewise, subjects performed similarly to real world condition of Experiment Three in the indirect blind walking condition. Again, it is not surprising that there is no effect of the self-avatar.

III.5 Experiment Five: Near-to-far Scanning versus No Scanning in Immersive Virtual Environment

The first four experiment established a baseline of people's performance on distance perception in the real world and the immersive virtual environment using near-to-far scanning, restricted field of view, and indirect blind walking. We tried to find whether the self-avatar adds value to people's distance perception as showed in some previous research. But we did not find any significant effect of the self-avatar. People perceived more than 90% of the target distance, which had much less underestimation compared to previous research. We hypothesized that the near-to-far scanning used in our experiment might be responsible for improving distance estimation in immersive virtual environment. Thus we designed this experiment with two conditions: *no scan* and *scan* to compare people's performance on

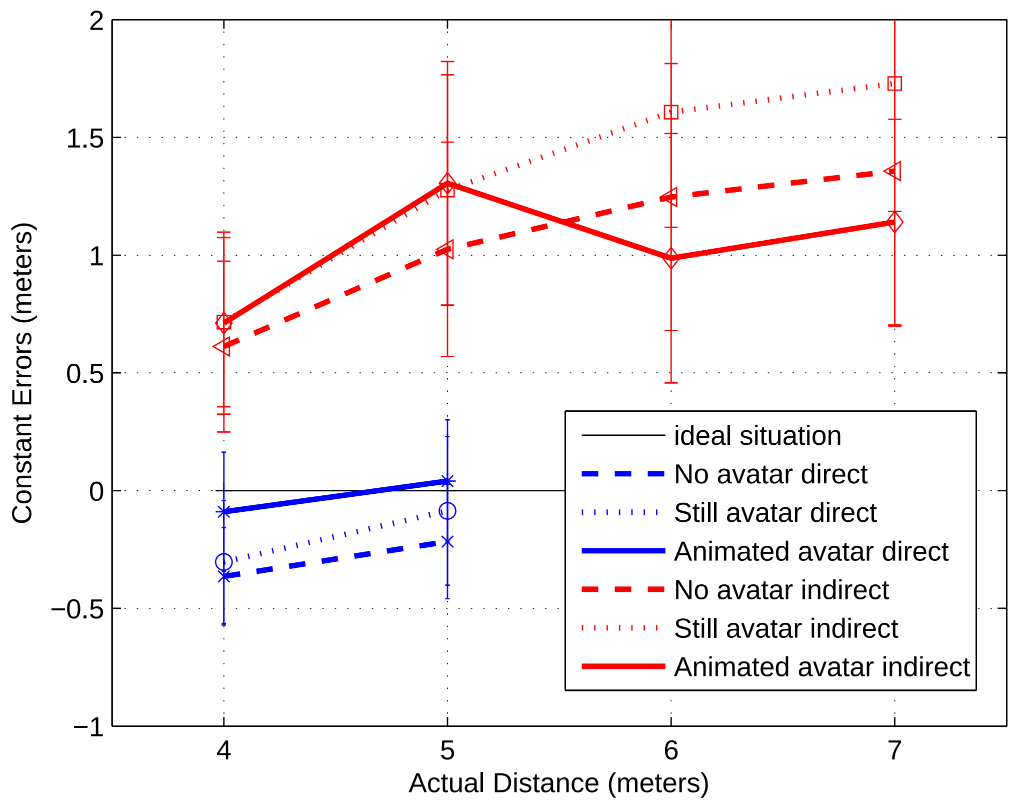

Figure III.9: This figure shows the constant error as a function of distance in Experiment Four. The black line shows the ground truth condition for comparison. The other six lines show the no avatar, still avatar and animated avatar for each walking method, direct and indirect. Error bars indicate standard errors of the mean.

distance perception. Here, we tested on three target distance of 5m, 7.5m and 10m.

III.5.1 Participants

Twenty four right-eye dominant subjects (12 males and 12 females) participated in this virtual environment experiment with ages ranging from 18 to 35. One half of the subjects (6 males and 6 females) participated in the *no scanning* condition and the other half in the *scanning* condition. All participants had normal or corrected to normal vision and they were recruited from Vanderbilt University. Participants were paid at the rate of $10 per hour for their participation.

III.5.2 Materials and Apparatus

The immersive virtual environment was rendered using Vizard (Worldviz, Santa Barbara, CA) and viewed through a full color stereo NVIS (Reston, VA) nVisor SX head mounted display (HMD) with 1280×1024 resolution per eye, with manufacturer's specification of a field of view of 60° diagonally, and a frame rate of 60Hz. Subjects wore a head-mounted display (HMD) to view the immersive virtual environment.

III.5.3 Method

From Experiment Two, we know that people were better at estimating distances with the large aperture. We eliminated the small FOV condition in immersive virtual environment. In the real world, people behaved similarly in direct blind walking and indirect blind walking. Multiple studies have reported that distances beyond a few meters appear to be compressed in immersive virtual environments presented through head-mounted displays (HMDs) using direct blind walking (Thompson et al., 2004; Knapp and Loomis, 2004; Ries et al., 2009; Grechkin et al., 2010; Mohler et al., 2010). To keep the experimental time short, we used only the indirect blind walking in our immersive virtual environment experiment. We chose 5m, 7.5m and 10m as our target distances because we also wanted to know how people behave at long target distances.

The goal of this experiment was to explore whether the near-to-far scanning method would improve people's performance on distance perception. We designed two scanning conditions in our immersive virtual environment experiment: *no scanning* and *scanning*. Subjects perceived distances in an immersive virtual environment wearing an HMD. The view in the left eye position of the HMD was blocked to guarantee the monocular viewing and the FOV of the right eye is 21.1°. In the *no scanning* condition, subjects wore a collar (see Figure III.10 on page 43) around their neck to avoid unintentional scanning, wore the HMD to view the immersive virtual environment and kept looking at the target until they were confident that they remembered the distance in their minds. In the *scanning* condition, subjects used similar scanning procedure as described in Wu et al. (2004). The only difference was that subjects did not need to operate the goggle. The open and close of the HMD were controlled by the experimenter through the rendering computer. There were two white stepping stones on the floor to indicate people's feet positions. Three target distances (5m, 7.5m and 10m) were used in each condition and each distance was tested three times in a randomized way. In each condition, the subjects were guided to the starting position and instructed to perceive the distance through the HMD using one of the scanning methods. After that, the HMD was blanked and the subjects were instructed to close their eyes. At this point the HMD was removed and subjects put on a blindfold that they were wearing around their neck. Then the subjects blindly walked to the remembered distance using the indirect blind walking after they took off the HMD.

III.5.4 Results and Discussion

Figure III.11 on page 44 illustrates the results of constant errors versus target distances for the two conditions. The black line shows the ground truth condition for comparison. The other two lines show the two scanning methods. A repeated measures analysis of variance (ANOVA) on the constant errors examining the effect of scanning method (scanning, no scanning) on distance estimations find a significant effect of scanning, $F(1,22) = 89.24$,

Figure III.10: This figure shows the collar worn around subjects' neck.

$p < 0.001$. Asking people to scan significantly improve distance perception in our immersive virtual environment. People can perceive more than 85% of the target distances. We examined the constant errors and found that the amount of distance underestimation in the *scanning* condition was significant, t(35) = -37.16, $p < 0.001$. Based on this result, it seems reasonable to conclude that it was the scanning method that produced near veridical results in the prior experiments. This finding has implications for the methods used in distance estimation in the immersive virtual environment literature, a topic we discuss further in Chapter V.

III.6 Experiment Six: Training

Based on Experiment Five, we can reproduce distance underestimation in an immersive virtual environment using the no scanning method. With this method, we can again ask

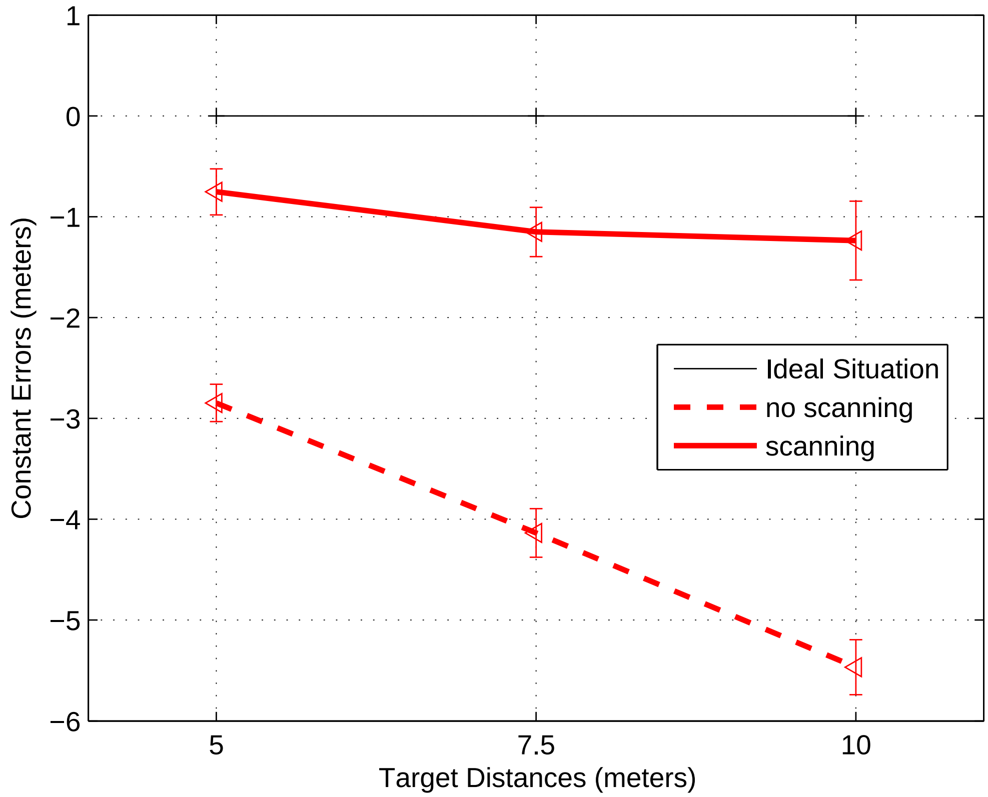

Figure III.11: This figure shows the constant error as a function of distance in the immersive virtual environment experiment. Error bars indicate standard errors of the mean.

the question of whether a self-avatar makes a difference in distance estimation. In this experiment, we will examine two factors, the presence or absence of a self-avatar, and the effect of the training regime, that is, the accommodation period in which people accustom themselves to their self-avatar.

Our experiment was between subjects. For each subject, there were two phases: a training phase and a task phase. The subjects did the training before the distance perception task phase. We used different virtual environments in the two phases. The training phase was done in a large grass virtual field. The task phase was done in a campus like virtual environment, with some virtual buildings, trees, and people in the surroundings. We will talk more about the training conditions in the design section. During the task phase, subjects were not allowed to scan target distances. They used the *no scanning* method described in the previous experiment and wore the collar around their neck to avoid unintentional scanning.

III.6.1 Participants

Forty-eight subjects (24 males and 24 females) participated in this virtual environment experiment with their ages range from 18 to 35. All participants had normal or corrected to normal vision and were recruited from Vanderbilt University. Participants were paid at the rate of $10 per hour for their participation.

III.6.2 Materials and Apparatus

We used an eight-camera Vicon (Los Angeles, CA) MX-F40 optical tracking system and Tracker (v. 1.0) for real-time motion capture of subjects. Subjects wore six components placed on the head, waist, right hand, left hand, right and left feet (see Figure IV.4). Motion data was transmitted to a second machine running Motionbuilder (Autodesk, San Rafael, CA), which mapped the motion data to a calibrated character using inverse kinematics and sent the resulting character data to the immersive virtual environment rendering machine. We used the same rendering equipment as that in Experiment Three. Suiting a subject in the motion capture components and calibrating the virtual avatar to their body size took approximately 10 minutes. The self-avatars were gender matched to the participants and they were co-located with the participants. The leg and arm lengths of the avatar were calibrated to the actual leg and arm lengths of the subject.

III.6.3 Design and Procedure

We designed three training conditions to test the effect of training: no training at all, training process similar to that in Mohler et al. (2010) and training process as that in Experiment Four. We used *avatar - no training*, *avatar - training 1* and *avatar - training 2* to represent the three training conditions, respectively. All subjects were instructed to wear the six motion capture components and the experimenter calibrated a self-avatar for them. Then they were instructed to do the training. After the training, they did the task phase as in the *no scanning* condition of Experiment Three. In the *avatar - no training* condition, subjects were not allowed to look down to see the self-avatar during the training phase. They could only look around to see the virtual environment and see their hands when they put their hands in front of the HMD. In the *avatar - training 1* condition, subjects kept standing at a position and they were not allowed to move out of that position. They were asked to obeserve the virtual environment and the self-avatar carefully, which lasted around 4 minutes. In the *avatar - training 2* condition, we did the same training as that in Experiment Four, during which subjects were instructed to get familiar with their self-avatars by looking into the mirror or by looking down. Then, they were instructed to walk on stepping stones and touch the three red virtual boxes. The training in this condition also lasted around 4 minutes. After the training, subject did the distance perception task. The same distances (5m, 7.5m and 10m) were used in each condition and each distance was tested three times in a randomized way.

To determine the effect of the self-avatar, we had another condition, which we called *no avatar - training 2* condition, in which subjects did not possess a self-avatar and needed to do the training as that in *avatar - training 2* condition before the task phase. Thus we had a direct comparison between the *no avatar - training 2* condition and *avatar - training 2* condition and know whether the presence of the self-avatar played a role in improving people's distance perception. Although there was no self-avatar in this condition, we still asked subjects to wear the motion capture components to insure they had the same equipment as the the other three conditions.

Then we had four conditions (*avatar - no training*, *avatar - training 1*, *avatar - training 2* and *no avatar - training 2*) in this experiment. We chose a between subjects design to avoid possible carry-over effects with twelve subjects (6 males, 6 females) in each of these conditions.

Subjects read the experimental instruction and we explained the task procedures again orally to insure that they understood them. Subjects wore six motion capture components and we calibrated matched-size self-avatars for them in the immersive virtual environment. Then they did the training phase followed by the task phase.

III.6.4 Results and Discussion

Figure III.12 on page 47 illustrates the results of constant errors versus target distances for the four conditions. The black line shows the ground truth condition for comparison. The other four lines show the four conditions.

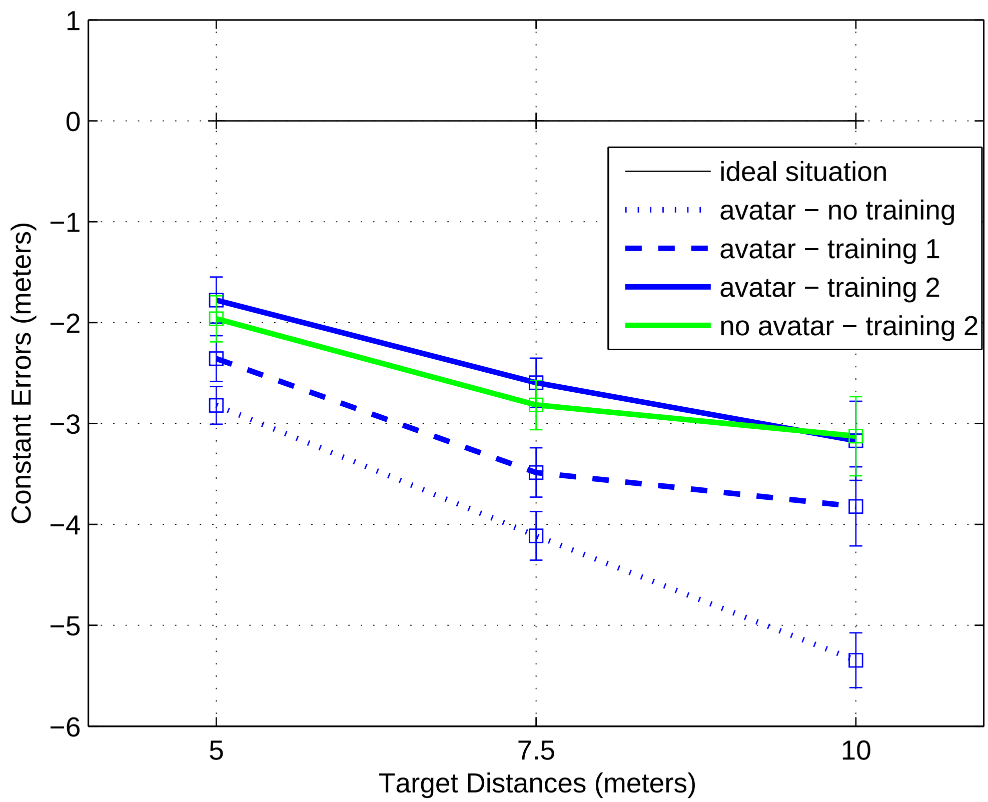

Figure III.12: This figure shows the constant error as a function of distance in the avatar experiment. Error bars indicate standard errors of the mean.

A repeated measures analysis of variance (ANOVA) on the constant errors examining the four conditions (*avatar - no training*, *avatar - training 1*, *avatar - training 2*, and *no avatar - training 2*) on constant errors found a significant effect of training, $F(3,44) = 3.79$, $p = 0.017$ and a significant effect of target distances, $F(2,66) = 69.76$, $p < 0.001$. The post-hoc analysis found a significant difference on constant errors between *avatar - no training* condition and *avatar - training 2* condition, $F(1,22) = 17.87$, $p = 0.021$ and a signficant difference between *avatar - no training* and *no avatar - training 2*, $F(1,22) = 16.58$, $p = 0.038$. We did not find significant effect of the training in the other condition combinations.

A repeated measures analysis of variance (ANOVA) on the constant errors examining the effect of self-avatar (*no avatar - training 2* and *avatar - training 2*) on distance estimations did not find a significant effect of the presence of the self-avatar.

Compared to the result in the previous experiment, a repeated measures analysis of variance (ANOVA) on the constant errors examining the effect of training without the self-avatar (*no scanning* and *no avatar - training 2*) on distance estimations find a significant effect of training when the self-avatar is not present, $F(1,22) = 7.936$, $p = 0.01$.

The significant effect of training has implications for people to spend more time interacting with immersive virtual environments before real tasks. The presence of the self-avatar does not seem to have any effect, which is a very different result from Mohler et al. (2010), where a strong effect of the avatar was found. More work is needed to examine the effect of the self-avatar to quantify what aspects of self-avatar are important for people to perform better on distance estimation.

CHAPTER IV

AFFORDANCE EXPERIMENTS

People judge what they can and cannot do all the time when acting in the physical world. Can I step over that fence or do I need to duck under it? Can I step off of that ledge or do I need to climb off it? These qualities of the environment that people perceive that allow them to act are called affordances. This chapter compares people's judgments of affordances on two tasks in both the real world and in virtual environments presented with head-mounted displays. The two tasks were stepping over or ducking under a pole, and stepping straight off of a ledge. Comparisons between the real world and virtual environments such as done in this article are important because they allow us to evaluate the fidelity of virtual environments. Another reason is that virtual environment technologies enable precise control of the myriad perceptual cues at work in the physical world and deepen our understanding of how people use vision to decide how to act. In the experiments presented here, the presence or absence of a self-avatar — an animated graphical representation of a person embedded in the virtual environment — was a central factor. Another important factor was the presence or absence of action, that is, whether people performed the task or reported that they could or could not perform the task.

Besides these factors, whether the size of the self-avatar will change people's behavior inside immersive virtual environment is another interesting topic we explore in this chapter. There are two affordance tasks here: stepping over or ducking under a pole and walking through or ducking under doorways.

IV.1 Experiment One: Pilot Study: Stepping Over or Ducking Under a Pole in the Real World

To illustrate the nature of the affordance, we conducted a small pilot study in the real world with a real pole. For each of six members of our laboratory (none of them the authors),

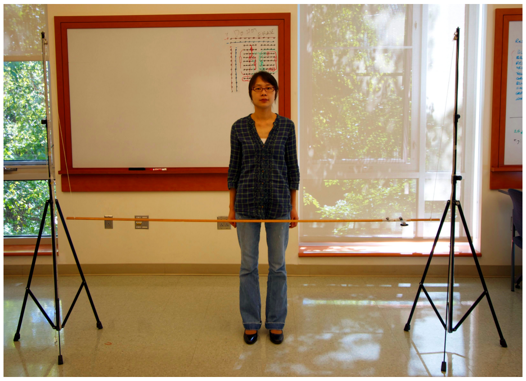

Figure IV.1: The physical pole apparatus used for the pilot and following experiments.

Figure IV.2: First person view used in judging whether to step over or duck under the pole.

we measured their eyeheight, and starting with a horizontal pole at a height of 2m, lowered it repeatedly, at each step asking each member of the lab whether they could duck under the pole. We used the apparatus shown in Figure IV.1 with a first-person view shown in Figure IV.2 on page 50. We stopped when individuals reported that they did not think that they could, and recorded the height of the pole. We then started the pole on the floor and successively raised it, asked the participant at each height if he or she could step over the pole, again stopping and recording the height when a negative response was obtained. The results of this pilot study are shown in Figure IV.3 on page 51. In this figure, we normalize the thresholds by the eyeheight of the participants. The average ducking under threshold was 0.52 (SE=0.02), and the average stepping over threshold was 0.53 (SE=0.02). Note that passability by ducking under can occur for every height above the minimum ratio, and passability by stepping over can occur for every height below the minimum ratio. These results give a loose indication that there is not a significant gap in the responses — as seems reasonable, that is, at least one or the other is always possible — and that the range where *both* are reported to be possible is narrow. We will now quantify these results, more elaborately, first in a virtual environment, and then in the real world.

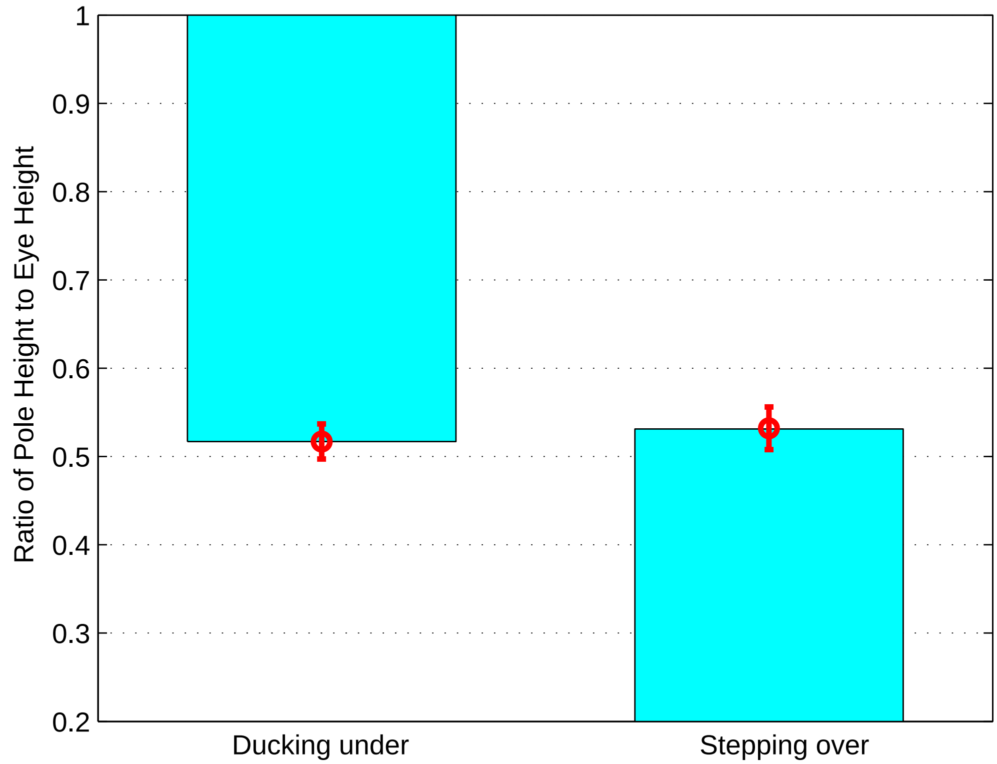

Figure IV.3: Results from pilot experiment (N=6). Error bars are standard errors of the mean. The minimum ratio for ducking under was 0.52 (0.02) and the maximum ratio for stepping over was 0.53 (0.02).

IV.2 Experiment Two: Stepping Over or Ducking Under a Pole in an IVE

In this experiment, we want to learn whether requiring people to do the action in the task of stepping over or under a pole will change people's performance. Thus we designed a within subjects condition: *no action* or *action*. In the *no action* condition, subjects made a judgment as to whether they would step over or duck under a pole, but did not perform the action; in the *action* condition, subjects both made the judgment and performed the action. We blocked the experiment so that subjects performed the *no action* trials first to avoid possible carry-over effects from the *action* trials. In addition to the *action/no action* condition, we also explored whether the presence of an animated match-sized self-avatar would change people's behavior on the task, leading to two avatar conditions: *no avatar* and

Table IV.1: Experimental Design

	Avatar Condition	
Pole Condition	No avatar	Avatar
With Pole	Action/No Action	Action/No Action
Without Pole	Action/No Action	Action/No Action

The Pole and Avatar conditions are between groups. The Action/No Action condition is within subjects. For this design N=48, with 12 subjects in each cell.

avatar. In the *no avatar* condition, subjects are presented with a view of the world in which they receive no view of their own body; that is, they view the world from the viewpoint of a disembodied virtual camera. In the *avatar* condition, subjects had an articulated graphical model that moved as they moved, and was matched in gender and size to them.

An immersive virtual environment showing a graphical pole that subjects elect to step over or duck under is fundamentally different from the real world in yet another way: subjects in an action condition may receive haptic and tactile sensations from the pole if they brush against it. These sensations provide additional feedback that might be used on a trial-to-trial basis to refine subjects' estimates of their affordance threshold, and may make subjects more conservative because it introduces a risk of falling. To simulate this condition in our immersive virtual environment, we added an additional factor: a physical pole that was tracked to the location of the virtual pole in all experiments. This factor manifests itself in a *pole* and *no pole* condition. In practice, this condition can only have an effect in action conditions, since no tactile feedback can be gained in judgment-only (no action) conditions.

In each combination of the conditions, we tried to find the threshold values of the affordance at which people changed their decision or behavior from stepping over to ducking under in our immersive virtual environment. The experiment was conducted as a mixed design, with the *action/no action* condition manipulated within subjects, as mentioned above, and the other conditions (*avatar/no avatar*, *pole/no pole*) manipulated between groups in counterbalanced order.

IV.2.1 Participants

Forty-eight subjects (24 males and 24 females) participated in this virtual environment experiment with ages ranging from 18 to 30. All participants had normal or corrected to normal vision and they were recruited from Vanderbilt University. Participants were paid at the rate of $10 per hour for their participation.

IV.2.2 Materials and Apparatus

We used an eight-camera Vicon (Los Angeles, CA) MX-F40 optical tracking system and Tracker (v. 1.0) for real-time motion capture of subjects. Subjects wore six components placed on the head, waist, right hand, left hand, right foot and left foot (see Figure IV.4 on page 54). Motion data were transmitted to a second machine running Motionbuilder (Autodesk, San Rafael, CA), which mapped the motion data to a calibrated character using inverse kinematics and sent the resulting character data to the immersive virtual environment rendering machine. The immersive virtual environment was rendered using Vizard (Worldviz, Santa Barbara, CA) and viewed through a full color stereo NVIS (Reston, VA) nVisor SX head mounted display (HMD) with 1280×1024 resolution per eye, with manufacturer's specification of a field of view of 60° diagonally, and a frame rate of 60Hz. Suiting a subject in the motion capture components and calibrating the virtual avatar to their body size took approximately 10 minutes. For the *avatar* condition, prototypical self-avatars were gender matched to the participants and they were co-located with the participants. The leg and arm lengths of the avatar were calibrated to the actual leg and arm lengths of the subject, and the eye height of the avatar was calibrated to the eye height of the person.

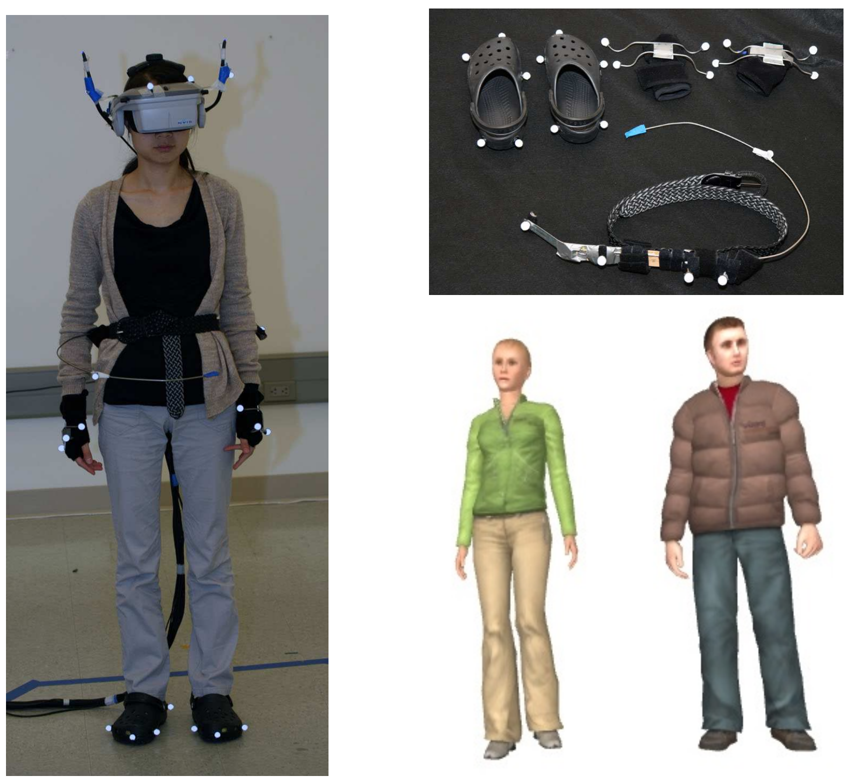

Figure IV.4: This figure shows a person wearing the HMD and motion capture components (left), motion capture components along (top right) and the avatar models used for the self-avatar (bottom right).

IV.2.3 Design and Procedure

We manipulated the *action* condition within subjects, and the *avatar* and *pole* conditions as between-subjects as described previously. The $2 \times 2 \times 2$ design is illustrated in Table IV.1 on page 52, with each group having 12 people. Thus, 24 subjects (12 male, 12 female) participated in each between subject condition, with 12 subjects (six male, six female) in each of the *avatar*, *no avatar*, *pole* and *no pole* conditions. Each group of 12 subjects performed trials first in the *no action* or judgment-only mode, and then performed trials in the *action* condition. Thus, all *no action* trials were performed before any *action* trials.

Subjects read the experimental instructions and listened to the task procedures again orally to insure that they understood them. All subjects were required to wear six motion capture components in all conditions and we calibrated an avatar in the immersive virtual environment for subjects in the avatar conditions. All subjects were asked to look around, look up, and look down to get familiar with the immersive virtual environment after wearing the HMD. Subjects in the *avatar* conditions were asked to notice their self-avatar and engage with it when they looked down, but no formal training was done in this experiment as in Chapter III. In the *no action* conditions, the subjects started from the start position and walked across the room. At the start position, they were asked to observe the height of the pole carefully by looking forward, looking to the left and the right. Then they started to walk. Upon reaching the pole they paused and were asked what their action (step over or duck under) would be. After reporting, the pole in the immersive virtual environment disappeared and subjects walked to the other side of the virtual room. Then participants turned back and started a new trial. In the *action* conditions, we had a similar procedure. However, the subjects performed the action when they reached the pole and the pole in the immersive virtual environment was kept present. People could potentially see the pole when they tried to step over or duck under. Subjects were able to touch the pole in the action conditions with the physical pole. They were not given special instructions to avoid contact with the pole.

In each condition, there were 25 trials of stepping over or under the virtual pole in the immersive virtual environment. So there were 50 trials in total for each subject, which took approximately 40 minutes per subject. In the avatar conditions, including the time for the real-time motion capture, each subject needed about 50 minutes to finish the experiment.

For each trial, the height of the pole was determined by an adaptive maximum-likelihood stimulus procedure as described in Grassi and Soranzo (2009). The values of the midpoint, slope and false alarm rate used in this algorithm were all chosen through a pilot experiment using the method of constant stimuli in a manner similar to that described in Wu et al. (2009). The procedure iterates using prior responses and the magnitude of the next height is calculated by optimizing over candidate psychometric functions and eventually converges to a threshold value. In each condition, the maximum and minimum of the pole height is 1.0 $\times$ participants' eye height and 0.2 $\times$ participant's eye height. An example of the maximum likelihood process for determining the pole height threshold for one subject in the experiment is shown in Figure IV.5.

IV.2.4 Results and Discussion

The maximum likelihood procedure used in this task converged quickly. To compute the threshold value for each subject, we averaged values in the last four trials for each subject in each avatar condition. The average threshold values for all subjects with standard errors of the mean for the ratio of pole height to eye height in the eight conditions are shown in Figure IV.6 on page 59. The values for the means and standard errors are shown in Table IV.2 on page 58

An omnibus ANOVA on the pole threshold values examining the effect of action condition (no action, action), avatar condition (no avatar, avatar) and pole condition (no physical pole, physical pole) revealed a main effect of requiring participants to do the actual action: $F(1,46) = 76.737$, $p < 0.001$ and a main effect of the presence of the self-avatar, $F(1,46) = 13.89$, $p < 0.001$. The analysis also showed a significant interaction between

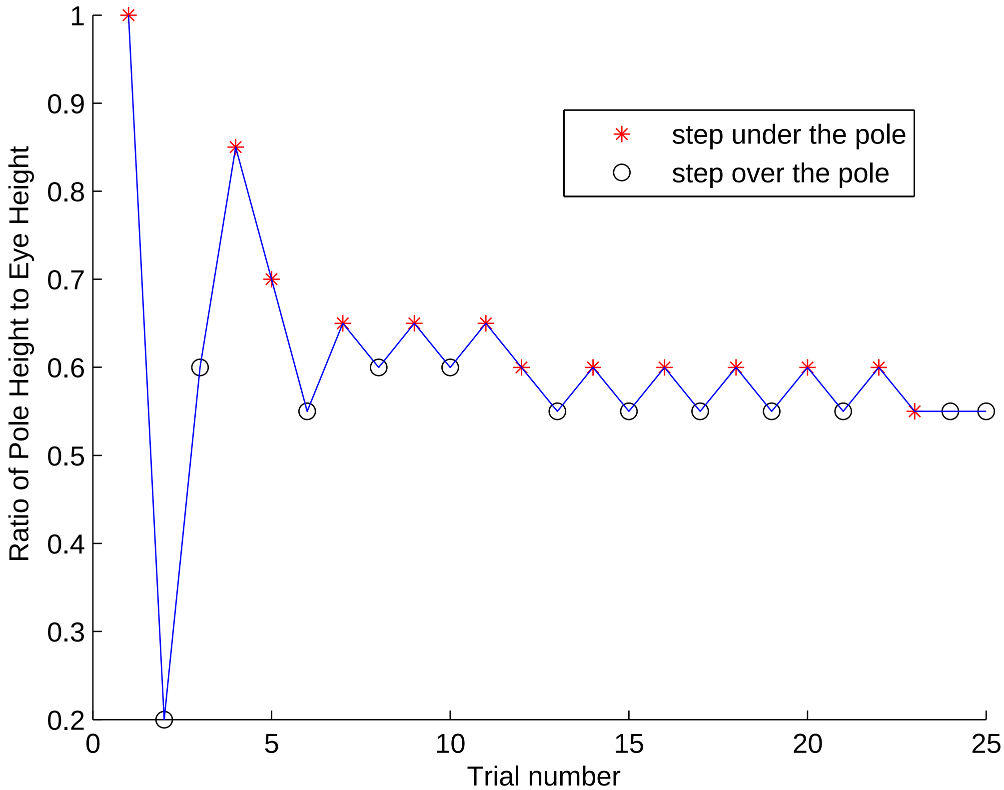

Figure IV.5: This figure shows the example of the maximum-likelihood procedure in the stepping over or under a pole in no avatar condition. The Y axis is the percentage of the pole height to eye height. A red cross means the participant selected to step under the pole. A black circle means the participant indicated to step over the pole. The final threshold is 0.55.

the action condition and the pole condition, $F(1,44) = 4.49$, $p = 0.0398 < 0.05$ and a significant interaction between the avatar condition and the action condition, $F(1,44) = 4.19$, $p = 0.0467 < 0.05$. These interactions are shown in Figures IV.7 on page 60 and IV.8 on page 61, respectively. For the pole $\times$ action interaction, Figure IV.7 on page 60 illustrates that occasional, inadvertent physical feedback can lower the threshold. For the avatar $\times$ action interaction, Figure IV.8 on page 61 illustrates the large and significant lowering of threshold with the avatar when no action occurs, and a slightly smaller lowering of threshold with the avatar when the action occurs.

In summary, both performing the action and having a self-avatar significantly reduced

Table IV.2: Quantitative Results for Experiment One (Ducking Under/Stepping Over in the Immersive Virtual Environment)

	No Avatar No Action No Pole	No Avatar Action No Pole	Avatar No Action No Pole	Avatar Action No Pole	No Avatar No Action Pole	No Avatar Action Pole	Avatar No Action Pole	Avatar Action Pole
Mean	0.60	0.54	0.54	0.50	0.59	0.49	0.54	0.48
Std Error	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01

Means and standard errors of the mean are expressed as a ratio normalized to eye height.

the thresholds at which subjects judged they could step over or duck under the virtual pole. The presence of a physical pole further reduced the threshold in the action condition. Since the presence of an avatar is optional in the virtual world, but people are necessarily embodied in the real world, and since requiring subjects to perform the action changed their performance in all the conditions, a question that clearly arises is how these results compare to a similar experiment in the real world, an experiment we conducted next.

These findings — that action significantly changes judgment — are consistent with perception studies in the real world that show less accurate spatial judgments when subjects are asked to make verbal descriptions of what they perceive as opposed to when subjects perceive and the produce a physical action (Creem-Regehr and Kunz, 2010). The action condition allows for the possibility of feedback, as subjects may be able to see where they are in relation to the pole as they perform the task (although the limited FOV of the HMD makes this challenging). Another possibility is that when stepping over a horizontal obstruction, people may think that there is a chance to lose their balance, and thus ducking under may be a more conservative behavior that people give preference to.

IV.3 Experiment Three: Stepping Over or Ducking Under a Pole in the Real World

Requiring subjects to perform the action changed people's behavior in all the immersive virtual environment conditions in the task of stepping over or ducking under a pole. A comparison to the real world data is therefore useful, as it will inform us that whether the action is necessary in immersive virtual environments for people to behave similarly to that

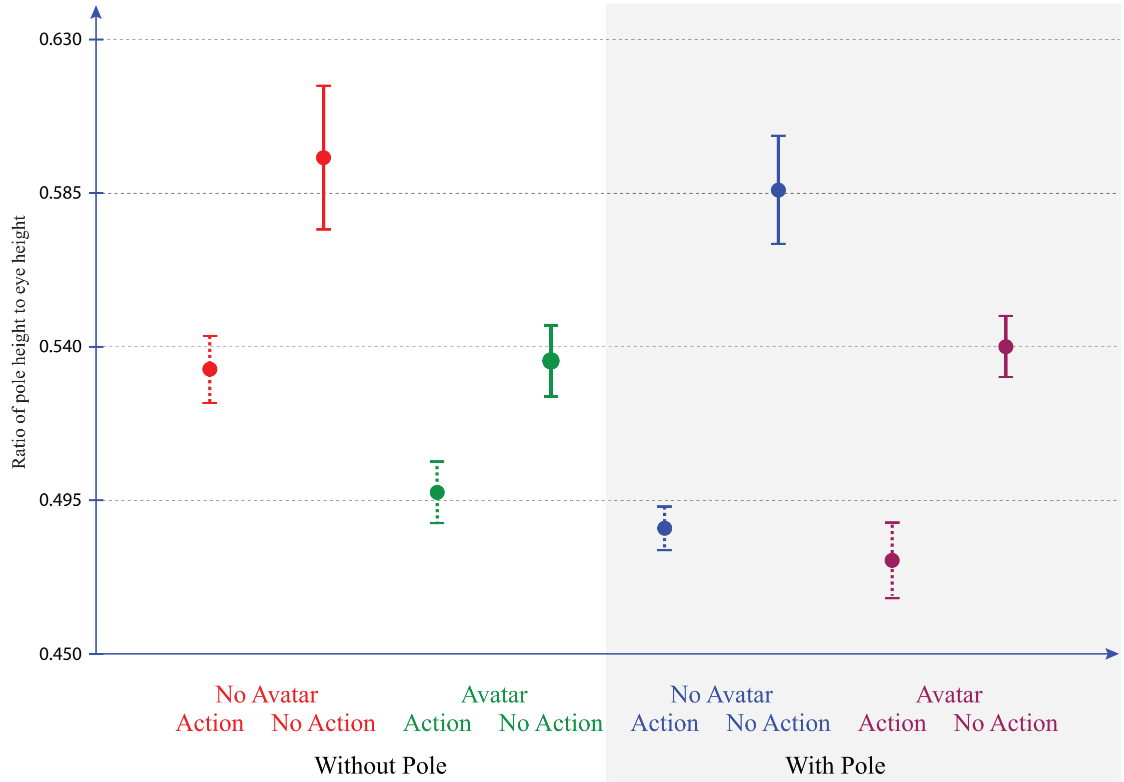

Figure IV.6: Threshold magnitudes in the task of stepping over or under a pole. The y-axis is the ratio of the pole height to subject's eye height. The error bars show standard errors of the mean. In the figure, "No Avatar" indicates no self-avatar was present, whereas "Avatar" indicates the presence of a self-avatar; "No Action" indicates subjects only perform a judgment, whereas "Action" indicates subjects step over or duck under the pole; "Without Pole" indicates no physical pole was present, whereas the shaded "With Pole" indicates conditions where there a physical pole was collocated with the virtual pole. The colors differentiate each group of participants, as explained in the text. There were 48 total participants, with 12 participants indicated by each color.

in the real world. Thus, we did the experiment in the real world to provide a baseline for people's performance in the virtual environment.

This experiment mimics the experiment performed previously, but done in the real world. We used a within subject design with two conditions: *no action* and *action*. We tried to find the threshold values for people to change their behavior from stepping over to ducking under. Note that in this experiment, we could not reasonably have an *avatar/no avatar* condition as subjects could always see their bodies, and we could not reasonably

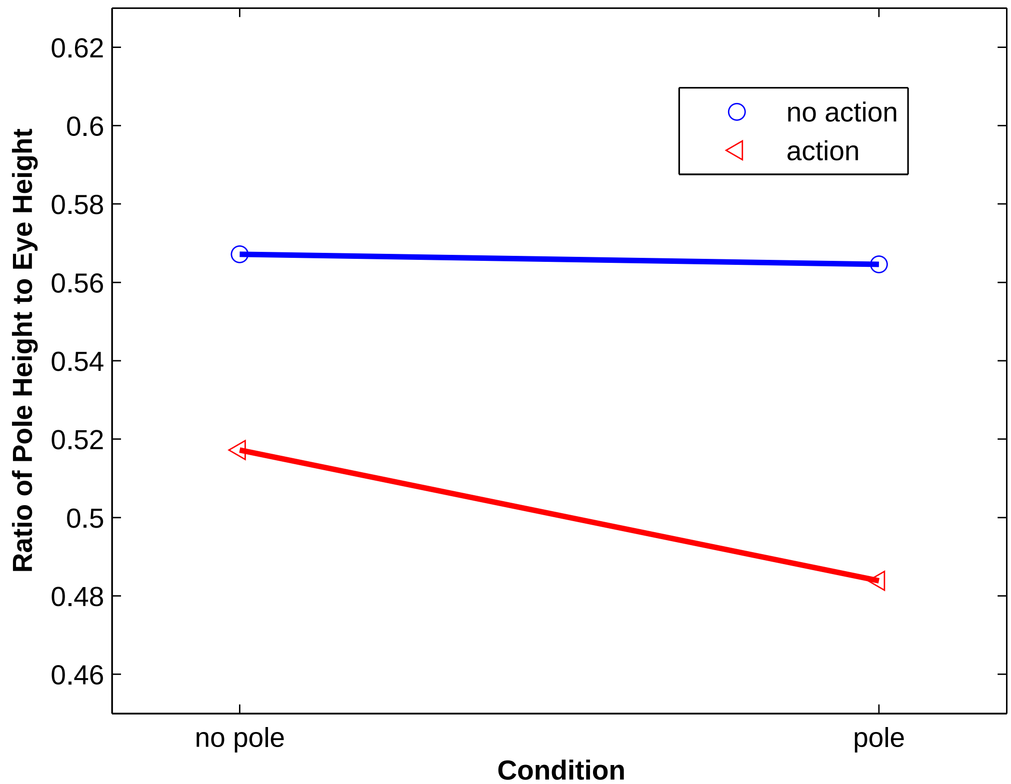

Figure IV.7: The pole by action interaction. The y-axis shows the ratio of the pole height to the subject's eye height. The blue line indicates the no action condition, and the red line indicates the action condition.

have a *pole/no pole* condition as there could be no analog to a virtual pole.

IV.3.1 Participants

Twelve subjects (6 males and 6 females) with different heights participated in this study. Their ages ranged from 18 to 30. All subjects had normal or corrected to normal vision and were recruited from Vanderbilt University. They were paid at the rate of $10 per hour for their participation.

IV.3.2 Materials and Apparatus

A horizontal rod approximately 1.5m long and 2cm in diameter was suspended on a rack such that it could be placed easily at a variable height.

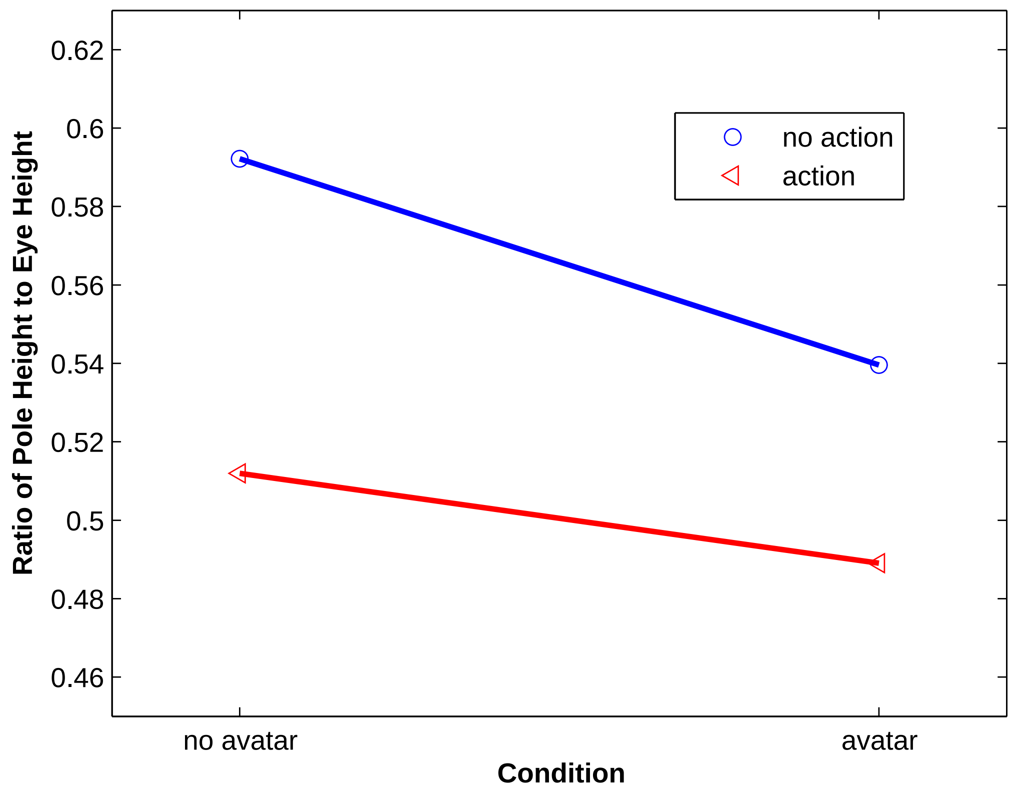

Figure IV.8: The avatar by action interaction. The y-axis shows the ratio of the pole height to the subject's eye height. The blue line indicates the no action condition, and the red line indicates the action condition.

IV.3.3 Design and Procedure

In this experiment, we also chose to do the *no action* condition first followed by the *action* condition, because we tried to avoid people using the feedback information provided in the *action* in the *no action* condition. Subjects read the experimental instructions and we explained the task procedures again orally to insure that they understood them. Then subjects began from the start position and walked across the room. At the start position, they were asked to observe the height of the pole carefully. Then they started to walk. Upon reaching the pole they paused and were asked what their action (step over or duck under) would be. In the *no action* condition, after reporting, the experimenter took out the pole and the subject walked to the other side of room without the stepping over or under action. The pole was repositioned while subjects faced the other direction, then they turned back and started a new trial. In the *action* conditions, subjects were required to step over or duck under the pole upon reaching it. Subjects sometimes brushed against the pole as they performed this action.

IV.3.4 Results and Discussion

The results for the real world are shown in Figure IV.9. For this experiment, the mean and standard error of the mean for the no action condition are 0.51 and 0.01, expressed as a ratio of pole height to eye height; the mean and standard error of the mean for the action condition are 0.47 and 0.01. In this experiment, we found a significant effect of asking people to do the actual action in the real world. When people were required to do the action, people had a significantly lower threshold value to step over a pole, $t(11) = 5.83$, $p < .001$.

We also compared the results in the IVE to those in the real world. Levene's test indicated that the samples had equal variances. We compared the effect of the environment condition between the real world and each of the IVE conditions using a Bonferroni-corrected Mann-Whitney U test. We found a significant difference between the IVE condition of *no*

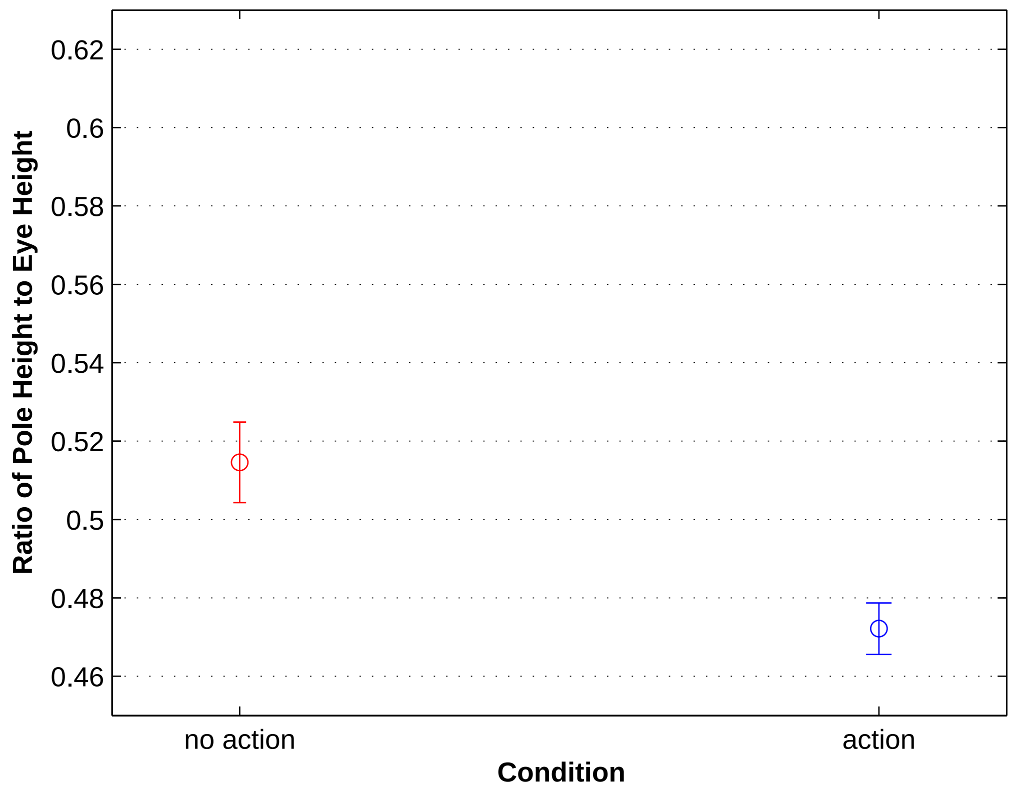

Figure IV.9: This figure shows the threshold values in stepping over or under a pole in the real world with and without action. Error bars indicate standard errors of the mean.

avatar, *no action*, *no pole* and the real world *no action* condition (medians of IVE condition and the real world condition were 0.54 and 0.53, respectively; the mean ranks were 7.5 and 17.5, respectively; $U = 132$, $Z = 3.52$, $p < 0.001$, $r = 1.02$). There was a significant difference between the IVE condition of *no avatar*, *action*, *no pole* and the real world *action* condition (medians of IVE condition and the real world condition were 0.51 and 0.48, respectively; the mean ranks were 7.35 and 17.8, respectively; $U = 135$, $Z = 3.74$, $p < 0.001$, $r = 1.08$). There was a significant difference between the IVE condition of *no avatar*, *no action*, *pole* and the real world condition *no action* (medians of the IVE condition and real world condition were 0.58 and 0.53, respectively; the mean ranks were 8 and 17, respectively; $U = 126$, $Z = 3.18$, $p < 0.001$, $r = 0.917$). Nothing else was significant.

We interpret this plethora of results as follows. In our ecological approach, we assume that behavior in a real world condition is behavior we want the IVE to approach. In our real-world experiment, as in many situations, people had multiple sources of information with which to judge the height of the pole — they could use visual information (our *no action* condition), and they could use visual information, kinesthetic information gained from motor action, and possibly tactile information from brushing against the pole (our *action* condition). In our experiment, we show that these two results are close but lead to discernibly different behaviors.

In the IVE, when the subject has no avatar, performs no action, and gets no physical feedback, their answer is significantly different from the closest real world condition. If we allow only action (no avatar and no pole), then the result is also significantly different from the closest real world condition. All other answers are within 7% of the closest real world conditions (and statistically indistinguishable in our experiments).

IV.4 Experiment Four: Stepping Off a Ledge in an Immersive Virtual Environment

As a second affordance task in our immersive virtual environment, we examine the issue of perception of a visual cliff from an action or no action viewpoint. In particular, we determine threshold values at which people in an immersive virtual environment are willing to step off from a height so as to investigate the effect of the presence of an animated matched-size self-avatar on their decision. More specifically, we determine the height at which people change their decision between stepping off and not stepping off. We performed this experiment in a between subjects manner with one half of the subjects not having a self-avatar (the *no avatar* condition) and the other half possessing a gender-matched, size-matched self-avatar (the *avatar* condition). In this experiment, subjects did not perform an action, as we deemed stepping off a ledge in the real world while wearing an HMD would be too dangerous; thus, we only recorded a report of what they would do.

IV.4.1 Participants

Twenty-four subjects (12 males and 12 females) participated in this virtual environment experiment with ages ranging from 18 to 30. All participants had normal or corrected-to-normal vision and they were recruited from Vanderbilt University. Participants were paid at the rate of $10 per hour for their participation.

IV.4.2 Materials and Apparatus

We used the same materials and apparatus as those in Experiment Two.

IV.4.3 Design and Procedure

We again chose a between subjects design. One-half (6 males, 6 females) of the subjects participated in the *no avatar* condition and the other half (6 males, 6 females) participated in the *avatar* condition. Subjects read the experimental instructions and we explained the task procedures again orally to insure that they understood them. Self-avatars were calibrated as described previously. All subjects were asked to look around, look up, and look down to get familiar with the immersive virtual environment after wearing the HMD. Again, subjects in the *avatar* condition were asked to notice their self-avatar and engage with it when they looked down, but no formal training was done in this experiment.

Subjects were then asked to walk until they stood beside the edge of a ledge; we then asked them to look down to observe the height of the ledge carefully. An example of what subjects saw in both the avatar and no avatar conditions is shown in Figure IV.10 on page 66. After this they were asked the question: are you able to step off the ledge gracefully and comfortably without losing your balance? Subjects responded and told the experimenter their decision. A new trial began after the experimenter recorded subjects' decisions.

The height of the ledge in each trial was determined using the same maximum-likelihood stimulus procedure as described in Section IV.2.3. There were 25 trials in each condition, which took around 40 minutes per subject, including the motion capture procedure.

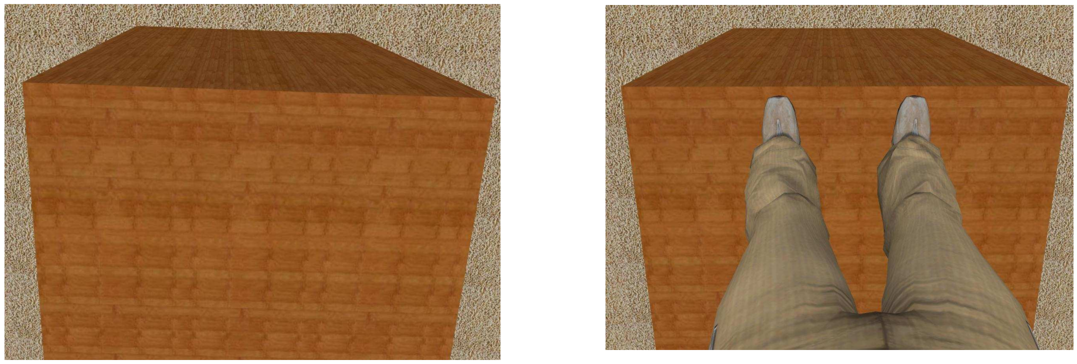

Figure IV.10: A view that a subject might see in the immersive virtual environment for the stepping off the ledge experiment in the *no avatar* condition (left) and *avatar* condition (right).

IV.4.4 Results and Discussion

To compute the threshold value for each subject, we again averaged values in the last four trials for each subject in each avatar condition. The average threshold values for all subjects with standard errors of the mean expressed as a proportion of ledge height to eye height in the two conditions are shown in Figure IV.11 on page 67. The mean and the standard error of the mean in the *no avatar* condition are 0.54 and 0.06, respectively. The mean and the standard error of the mean in the *avatar* condition are 0.27 and 0.03, respectively.

In the virtual stepping task, an independent sample t-test between the *no avatar* and *avatar* condition showed a significant effect of the presence of the self-avatar, $t(22) = 3.942$, $p < .001$. As can be seen from Figure IV.11, the presence of a self-avatar significantly reduced the threshold magnitude. This result is thus similar to Experiment Two, but like Experiment Two, there is a clear question as to how these results correlate to results in the real world.

Before addressing the comparison to the real world, however, we note that visual cliffs and ledges in virtual environments are anecdotally considered one of the more compelling types of virtual environments in terms of the sense of immersion that one feels in them ((Slater et al., 1995; Usoh et al., 1999; Meehan et al., 2002; Zimmons and Panter, 2003;

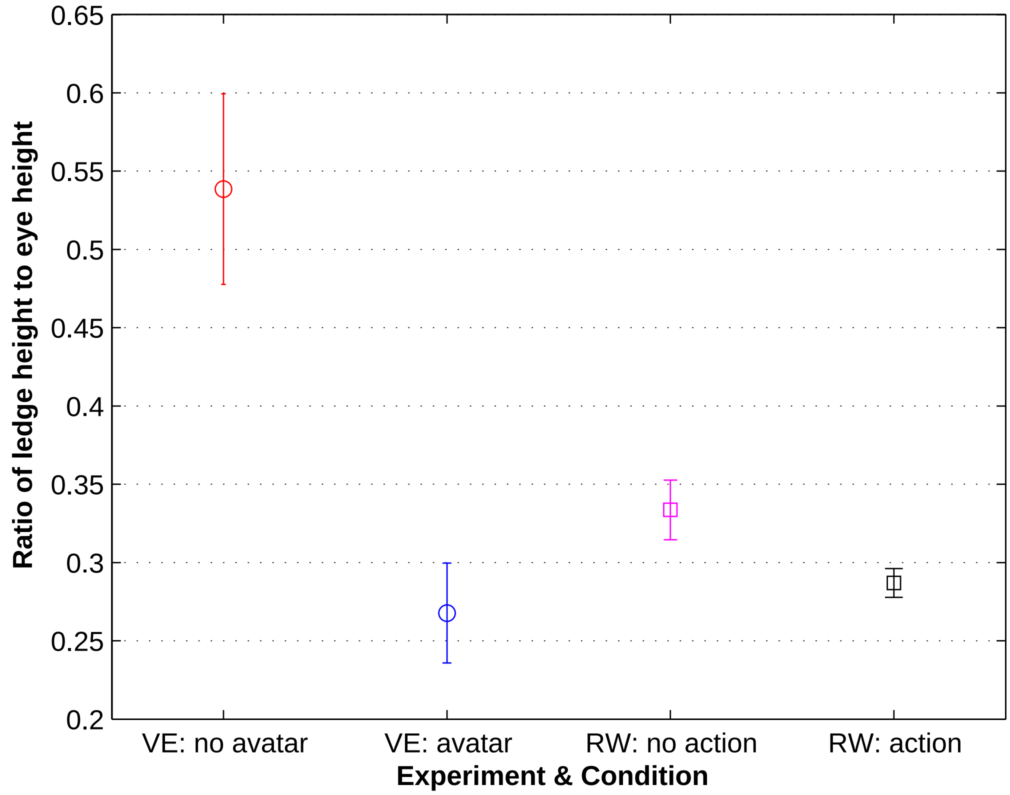

Figure IV.11: Threshold magnitudes in the two experiments involving stepping off the ledge. The Y-axis is the proportion of the ledge height to subject's eye height. The error bars show standard errors of the mean. The left two error bars (red and blue) show Experiment Four, done in a virtual environment (VE), where no avatar and avatar indicate the state of the *avatar* condition, with a self-avatar calibrated to the size of the subject. The right two error bars (magenta and black) show the results of Experiment Five, done in the real world (RW), where no action and action indicate the state of the *action* condition.

Slater et al., 2009a), although these visual cliffs are typically much larger than ours). Also, our experiment did not address whether the avatar needed to be an animated self-avatar or if a non-animated self-avatar would suffice. This question did not seem of particular interest to us since in our pilot studies and in observations of people during the study, most people fidget and adjust their feet. We felt that it would be odd to have feet that did not move. Another decision we made was not to have a prop in the real environment like a board that would give haptic feedback of an edge (see, for example, Figure 3 of Meehan et al. (2002)). It would be interesting to pursue such a course to see if the haptic feedback induced a greater sense of height or fear of heights, since it may induce emotional responses, some of which are known to influence the perception of heights (Stefanucci and Proffitt, 2009).

IV.5 Experiment Five: Stepping Off a Ledge in the Real World

The presence of the self-avatar resulted in significantly lower threshold values in our visual cliff virtual environment. A comparison to real-world data would be useful, as it would inform us whether people in the *avatar* condition are, in fact, behaving similarly to their behavior in the real world. We performed a similar experiment using the ledge outside the handrail of an accessibility ramp found on the Vanderbilt campus and depicted in Figure IV.12 on page 69.

IV.5.1 Participants

Twelve (6 males and 6 females) with different heights participated in this study. Their ages range from 22 to 32. All subjects had normal or corrected to normal vision and were recruited from Vanderbilt University. They were paid at the rate of $10 per hour for their participation.

IV.5.2 Materials and Apparatus

The ramp shown in Figure IV.12 on page 69 was used in this experiment. At its lowest point, it is 0.2m high, and at its highest it is 0.73m high. Measuring the length of the ramp

Figure IV.12: The ledge used for the stepping study in the real world.

provides the slope, and we determined the height to place subjects on a trial to trial basis by calculating the length along the ramp that they needed to be (as measured by the measuring tape that can be seen in Figure IV.12). We initially ran a pilot study among our lab members to determine if the ramp was high enough, and in practice no one reported or tried to step off the upper end.

IV.5.3 Design and Procedure

In this experiment, we again chose to do the *no action* condition followed by the *action* condition as within-subjects trials. Trials were blocked on these conditions. There was no *avatar* condition. In the *no action* condition subjects were asked if they could step off the ledge in a graceful and balanced manner; in the *action* condition, they were asked to perform the action if they indicated that they could (otherwise they did not). Subjects were instructed in this way so that would not attempt stepping off of a height that was dangerous to them. The procedure was otherwise identical to the that of Experiment Three. As mentioned previously, no one stepped off the ledge at its maximum height.

IV.5.4 Results and Discussion

Data were analyzed as in the previous experiment (Experiment 3). The mean threshold values for all subjects with standard errors of the mean expressed as a proportion of eye height to ledge height in both the *action* and *no action* conditions is shown on the right side of Figure IV.11 in black and magenta, respectively. The mean and standard error of the mean for the *no action* condition are 0.33 and 0.02, respectively. The mean and standard error of the mean for the *action* condition are 0.29 and 0.01, respectively.

For this experiment, a paired sample t-test showed a significant effect of *action*, $t(11) = 2.97$, $p < 0.01$. Subjects performing the action had a significantly lower threshold for stepping off the ledge than subjects making a judgment only. We also compared Experiment 3 to this experiment using a Bonferroni-corrected Mann-Whitney U test. Levene's test indicated equality of variances. In particular, we compared and found a significant difference between the IVE *no avatar* condition, which has no action, and the real world *no action* condition (the medians of the IVE and real world conditions were 0.5 and 0.33, respectively; the mean ranks of the IVE and real world groups were 17 and 8, respectively; $U = 126$, $Z = 3.12$, $p = 0.001$, $r = 0.90$). Thus, the absence of a self-avatar in the virtual environment results in significantly different performance from that of the corresponding real world condition.

IV.6 Experiment Six: Two Affordance Tasks to Explore the Effect of Changing the Size of the Self-avatar

Our approach to investigating the benefits of avatars was to assess the threshold values at which people decided to step over or duck under poles and walk through or duck under doorways. In this experiment, we determine the pole height at which people change their behavior from stepping over a pole to ducking under it and door height at which people change their behavior from walking straight to stooping under it. To test the effect of the avatar on these tasks in the immersive virtual environment, we had *no avatar* and *avatar* conditions. In the avatar condition, we hypothesized that participants with longer leg length avatars would able to step over higher poles. Then we designed two sizes: first, an avatar with the same leg length as that of the subject (referred to as avatar 1.00) and next, an avatar with a 15% longer leg length (referred to as avatar 1.15). The 15% stretch was chosen because it was noticeable yet did not stretch the avatars so much that they looked imhuman.

IV.6.1 Participants

Twenty four subjects, 12 males and 12 females participated in this study with their ages ranging from 18 to 30 and normal or corrected to normal vision. All participants were recruited from Vanderbilt University and they were compensated $10 per hour for their participation.

IV.6.2 Materials and Apparatus

We used an eight-camera Vicon optical tracking system and Vicon Tracker for real-time motion capture. Subjects wore six components placed on the head, waist, right hand, left hand, right and left feet. Motion data was transmitted to Motionbuilder, which mapped the motion data to a calibrated character and sent the resulting character data to Vizard, which rendered the immersive virtual environment. The immersive virtual environment was viewed binocularly through a full color stereo NVIS nVisor SX head mounted display(HMD) with 1280×1024 resolution per eye and a manufacturer specified field of view (FOV) of 60° diagonally.

IV.6.3 Design and Procedure

In pilot studies, we became concerned about carry-over effects between the conditions of the test. Thus, we opted for a between subjects design. One-half (6 males, 6 females) of the subjects participated in the no avatar condition and the other half (6 males, 6 females) participated in the avatar condition. To examine the size effect of the avatar on people's height perception, we used a within-subject design on the avatar sizes in the avatar condition. Avatars were co-located with the participants and they were gender matched to the participants. Avatars were calibrated to the arm length and head height of the person. In the avatar 1.00 condition, the leg length of the avatar was calibrated to the actual leg length of the subject. And in avatar 1.15, the leg length was 15% longer. The head height (eye height) was not changed; thus the torso:leg length ratio was altered by 15%. We only scaled the leg length of the avatar to be longer because we wanted to save time in order to keep the time of the subject in the immersive virtual environment around one hour. Body proportions such as foot size and width of limbs were not calibrated.

Each subject came in and read the instructions. Then we explained the experimental task and procedures again to make sure the subject understood them. Subjects in the avatar condition wore the motion capture components and we calibrated an avatar in the immersive virtual environment. At this point, a similar procedure for all subjects ensued. Subjects with a self-avatar were instructed to familiarize themselves with it. The subjects started from the start position and walked across the room. Upon reaching the pole or doorway they paused and were asked what their action (step over, duck under, etc.) would be. After reporting, the pole or door in the immersive virtual environment disappeared and the subject walked to the other side of the virtual room. They turned back and started a new trial. Subjects were allowed to view the floor and their self-avatar as much or as little as they wanted. Although we did not ask subjects to do the actual action, having people walk in the virtual environment might immerse people in the virtual environment and increase their presence. We made the pole or door disappear because we did not want the subject to have any feedback during the experiment.

At each trial, the height of the pole or the height of the door was determined by a maximum-likelihood stimulus procedure described in Grassi and Soranzo (2009). The values of the midpoint, slope and false alarm rate used in this procedure were all chosen in a pilot experiment in a manner similar to that of Wu et al. (2009). The procedure iterates and the magnitude of the next height is calculated by optimizing over candidate psychometric functions and eventually converges to a threshold value. In the pole scenario, the maximum and minimum of the pole height is 1.0× participants' eye height and 0.2× participant's eye height. In the door scenario, the maximum and minimum of the door is 1.3× participants' eye height and 0.8× participants' eye height. An example of pole height process for a subject in the experiment is shown in Figure IV.5.

In the no-avatar condition, there were 25 trials in stepping over or under a virtual pole in the immersive virtual environment. So there were 50 trials in total, which took around 30 minutes. In the avatar condition, there were 25 trials in each of the avatar size conditions in each scenario. So there were 100 trials in total, which took around 1 hour. Including the time for the real-time motion capture, each subject needed around 1 hour and 20 minutes to finish the experiment in the avatar condition. For the two sizes in avatar condition, one-half of the subjects started with avatar1 and the other half started with avatar 1.15.

The maximum likelihood procedure converged quickly. To compute the threshold values, we average the last four trials of the maximum likelihood method for each subject in each condition. The mean threshold values for all subjects with standard errors of the mean for the percentage of pole height to eye height are shown in Figure IV.13 on page 74 and that of door heights in Figure IV.14 on page 75.

IV.6.4 Result and Discussion

This experiment showed that having a matched (size, gender) self-avatar resulted in people behaving differently in the immersive virtual environment when performing the step over or duck under pole task. Subjects without an avatar judged they were able to step over a higher pole than subjects possessing a self-avatar. When we manipulated the size of the self-avatar, the judgment scaled approximately with the size of the scaled self-avatar: for a 15% increase in avatar's leg length, subjects judged a 12% increase in the critical pole height. On the other hand, the presence of a self-avatar, scaled or matched, had no effect

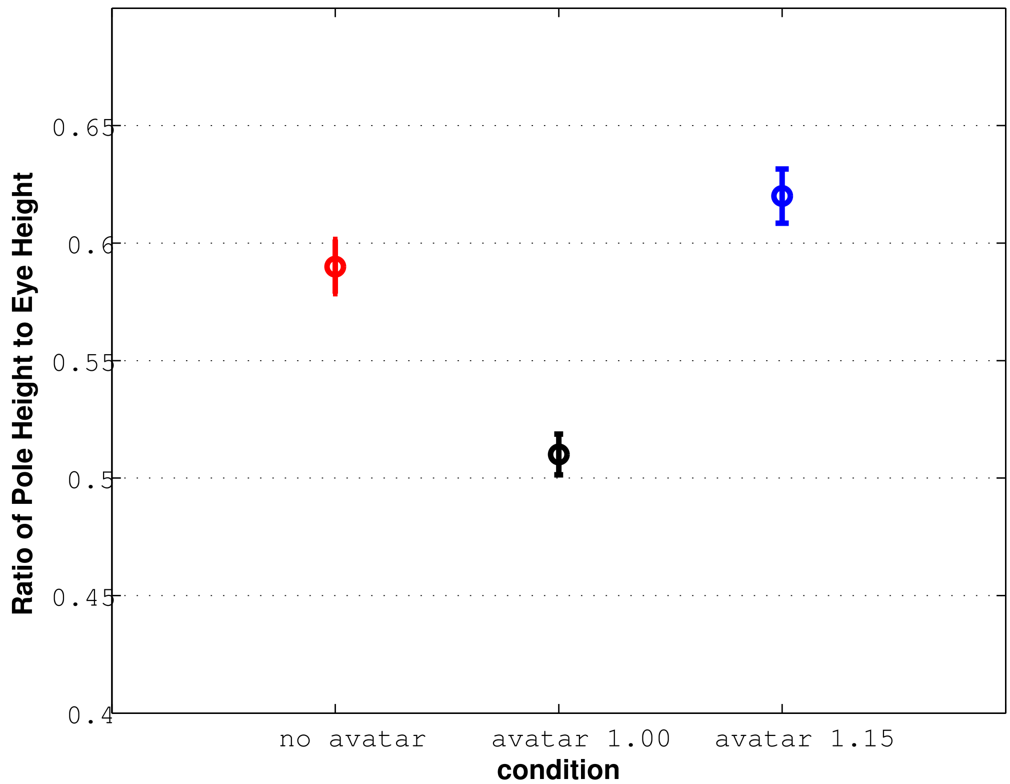

Figure IV.13: Threshold magnitudes in stepping over or under a pole in the three conditions. The Y axis is the percentage of the pole height to subject's eye height. The error bars show standard errors of the mean. In the figure, no avatar means no avatar condition, avatar 1 means the avatar's size is the same as the subject's size, avatar 1.15 means an avatar with 15% longer leg length.

on judgments in the ducking under door height task.

The results of the pole task are consistent with prior work (Slater and Usoh, 1994; Biocca and Rolland, 1998), and there are several potential explanations for this effect. First, the self-avatar enabled participants to ground themselves on the floor of the immersive virtual environment, which might help them identify their eye height and use this reference to judge the pole height with respect to their ability to act. Second, by observing the leg and foot sizes of the self-avatar when they looked down, the participants might get a good sense of the size metric in the immersive virtual environment. Immersive virtual environments look different from the real world environment and lack scale and context

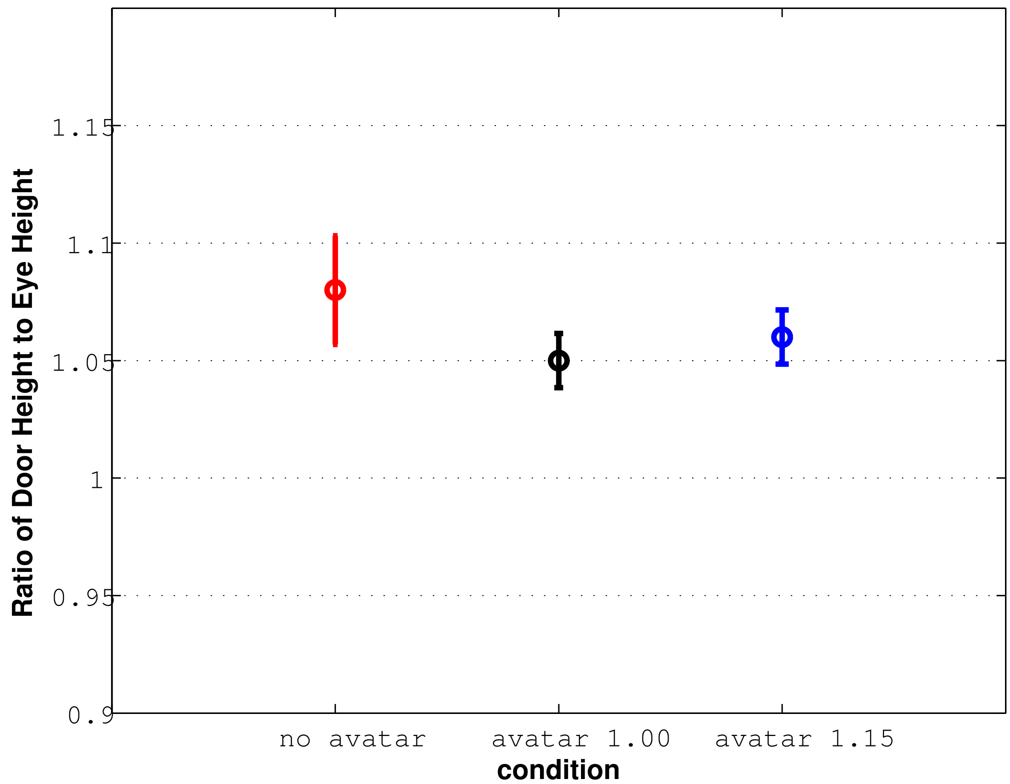

Figure IV.14: Threshold magnitudes in walking straight or stooping under a door in the three conditions. The Y axis is the percentage of the door height to subject's eye height. The error bars show standard errors of the mean. The no avatar, avatar 1 and avatar 1.15 have the same meaning as those in Figure IV.13.

information. The self-avatar might serve as a ruler to provide size information. In the real world, leg length performs an important role when people try to step over a pole. In immersive virtual environments, the presence of matched length legs might provide a direct comparison between the pole height and leg lengths and helped their judgments in the stepping task. Third, the self-avatar might increase people's presence in the immersive virtual environment. People currently experience most immersive virtual environments as a disembodied virtual camera. But humans are accustomed to being embodied in real world task. A final reason might be that that an animated self-avatar may provide participants motor information while performing the tasks that helps them calibrate body scale.

People seem to ignore the presence of the self-avatar in judging whether to duck. This result may be because there is dynamic geometric information that supercedes use of the avatar. When one walks toward a doorway that is higher than one's eyes, the visual angle to the top of the door rises relative to the horizon, whereas when one walks toward a doorway that is lower than one's eyes, the visual angle to the top of the door falls relative to the horizon. Additionally, it is difficult to calibrate the size of the self-avatar from the perspective of ducking.

CHAPTER V

CONCLUSION AND FUTURE DIRECTION

V.1 Conclusion and Discussion

This thesis develops novel methods for humans to experience and interact with immersive virtual environments. More specifically, this thesis looks at perception and action in immersive virtual environments. We would like to create virtual environment simulations in which people are able to act in the same way as they act in the real world. While this goal is currently beyond our reach, this thesis has improved our understanding of cues that appear to enable veridical action. More particularly, in this thesis, we examined distance perception and several affordance tasks and the effect of self-avatars and other factors on them. The following summarizes our findings.

V.1.1 Distance Perception

We have four principal findings on distance perception in immersive virtual environments and the real world. The first finding is the effect of the near-to-far scanning. When people use near-to-far scanning, there is less underestimation in our immersive virtual environment. The second finding involves the training in the immersive virtual environment. Training significantly improves people's performance on distance estimation and it might be a factor in much of the variation in results reported in the past literature. But third we do not find a significant effect of the presence of the self-avatars on distance estimation in our experimental setup. Self-avatars do not seem to help at all under our experimental conditions. The last finding is that we do not replicate the results of Wu et al. (2004) with regard to their far-to-near scanning. In Experiment Four of Wu et al. (2004), people underestimated distances with far-to-near scanning while they were accurate with near-to-far scanning. People were accurate in both scanning methods in our experimental setup. We have four principal findings on distance perception in immersive virtual environments and the real world. The first finding is the effect of the near-to-far scanning. When people use near-to-far scanning, there is less underestimation in our immersive virtual environment. The second finding involves the training in the immersive virtual environment. Training significantly improves people's performance on distance estimation and it might be a factor in much of the variation in results reported in the past literature. But third we do not find a significant effect of the presence of the self-avatars on distance estimation in our experimental setup. Self-avatars do not seem to help at all under our experimental conditions. The last finding is that we do not replicate the results of Wu et al. (2004) with regard to their far-to-near scanning. In Experiment Four of Wu et al. (2004), people underestimated distances with far-to-near scanning while they were accurate with near-to-far scanning. People were accurate in both scanning methods in our experimental setup.

There are a variety of methods people use in egocentric distance estimation to perceive distances. The most common one is asking people to look at the target and judge or remember the distance in their mind (Elliott, 1987; Rieser et al., 1990; Loomis et al., 1992, 1996; Fukusima et al., 1997; Philbeck and Loomis, 1997; Loomis and Beall, 1998). The other is using scanning method (Wu et al., 2004). How these different methods affect people on distance perception is still an open question to some extent. Wu et al. (2004) showed the importance of integrating the near ground information to facilitate the distance estimation. The authors first showed that a relatively wide expanse of the ground surface is required for accurate distance judgments. Then as proof of surface integration, they showed that even with restricted field of view, people could accurately judge absolute distance by scanning local patches of the ground surface, bit by bit, from near to far, but not in the reverse direction. One potential explanation is that when people use the near-to-far scanning, they get more information from the near ground, which help them to identify the distance better. This explanation is actually consistent with some previous research. Some studies (Gibson, 1950; Sinai et al., 1998) showed that a continuous surface of the ground is important for estimating absolute distances. Disrupting the homogeneity of the ground surface by manipulating the texture gradient or introducing a gap (Sinai et al., 1998) reduced people's accuracy on distance perception in the real world. As mentioned, people underestimated egocentric distance when this factor was not explicitly controlled (Witmer and Sadowski, 1998; Thompson et al., 2004; Knapp and Loomis, 2004; Jones et al., 2008; Ries et al., 2009; Grechkin et al., 2010; Mohler et al., 2010). It is interesting that distance estimation in immersive virtual environments may therefore be more susceptible to variations in experimental methodology than findings in the real world. How much of this is due to the technology of the HMD is not known, but it is worth pointing out that recent results with light weight, wide FOV commodity level hardware found no distance compression (Young et al., 2014; Li et al., 2014).

Prior research (Mohler et al., 2010) has shown the salient effect of the self-avatar on distance perception. In their experiment, they had a training phase before the experimental phase, during which subjects engaged with the avatars around five minutes. Experiment Six was designed to examine the effect of the training in immersive virtual environment. When people had a training, whether described in Mohler et al. (2010) or in our Experiment Four on distance perception, the training significantly improved people's performance on distance perception when they did not use the near-to-far scanning. But people still underestimated distance even they had the training. The training might provide people more information for distance perception. Unfortunately, the present experiments were not designed to determine the relevant cues that make up the critical properties of a self-avatar involved in the training. Body scaling might be a possible reason that a self-avatar helped people. During the traing phase, subjects interacted with a matched-size avatar, which might give them more information about their body scaling, compared to a disembodied virtual environment. People might use this information to identify the angle of declination, which has shown to falicitate distance perception in the real world (Sedgwick, 1983).

The use of self-avatars as a potential aid in distance estimation in immersive virtual environments is a recent development, and the theoretical reasons why it may affect spatial perception is not well understood. Our method differs from prior research (Mohler et al., 2010; Ries et al., 2009) in potentially significant ways. The training phase was different, and involved more interaction, and the use of a mirror. The mirror meant that a full-body view of the avatar was available. We also gender-matched the avatar. We were thinking of the presence of the self-avatar might give concrete size information, which might help people on distance perception. But the presence of the self-avatar did not help at all in our experimental setup. One explanation might be that the scanning is already sufficient for people to make accurate distance perception inside immersive virtual environments. People just did not need extra information from the self-avatar.

We failed to replicate the results of far-to-near scanning in Wu et al. (2004). People were accurate using this scanning method. One explanation is that the continuous surface during the distance perception is important for people to perceive distance accurately. Gibson (1950) and Sinai et al. (1998) showed that a continuous surface of the ground is important for estimating absolute distances. Disrupting the homogeneity of the ground surface by manipulating the texture gradient or introducing a gap (Sinai et al., 1998) reduced people's accuracy on distance perception in the real world. And it might be that people get enough information to judge their eye height and angle of declination from the process of scanning either using near-to-far or far-to-near.

There are some other minor findings on distance estimation. The effect of the walking method in the real world on distance estimation in Experiment Three raises interesting issues on the robustness of spatial representation in short term memory. Such representations appear less robust to retention when disturbing by priming sensorimotor action, particularly when the the acquisition of the representation was obtained through a small field of view. It may be that with small field of view, people are unable to construct a representation of the environment that is then robust to spatial disturbance, c.f., spatial updating (Rieser and Pick, 2007). The trend toward overestimation that these experiments show could then be attributed to uncertainty: people keep walking until they are certain they have walked far enough. These findings raise several interesting issues for spatial navigation and processing involving environmental and featural cues (Kelly et al., 2008, 2009) that could be a topic of future work.

V.1.2 Affordance Judgments

Chapter IV presented three studies of affordances in the real world and in immersive virtual environments, that of stepping over or ducking under a pole, that of stepping off of a ledge, and that of ducking under a doorway. A perception task and perception-action task were used. For the perception task, participants were asked to look at the situation and then decide whether they would step over or duck under or could step off the ledge or could walk through a doorway. For the perception-action task, participants were asked to physically duck under or step over the pole or to step off the ledge. The experiments were designed to see whether the presence of a matched-size self-avatar affected performance in the virtual environment, and also whether reporting a judgment was different than acting on a judgment. In the experiment for stepping over or ducking under in the virtual environment, we also examined whether potential, occasional, and inadvertent haptic feedback from a real pole collocated with the virtual pole would affect performance. When we manipulated the size of the self-avatar, the judgment scaled approximately with the size of the scaled self-avatar: for a 15% increase in avatar's leg length, subjects judged a 12% increase in the critical pole height. On the other hand, the presence of a self-avatar, scaled or matched, had no effect on judgments in the ducking under doorway task. Comparing performance in the first two tasks tasks between the real and virtual environments offers a quantifiable measure of the fidelity of the virtual environment.

Our primary finding is that the presence of the self-avatar provides important information in stepping over or ducking under a pole and in stepping off a ledge. The lack of a self-avatar in the virtual environment resulted in significantly different thresholds between the virtual environment and the real world, whereas thresholds with a self-avatar present were not significantly different from thresholds in the real world. This finding is important because creating a matched-size self-avatar is cumbersome, requiring extra software and hardware for the virtual environment, and additional time to calibrate the self-avatar to a user of the virtual environment.

What information does the self-avatar provide that makes such a significant difference in thresholds in these two tasks? Unfortunately, the present experiments were not designed to determine the relevant cues that make up the critical properties of a self-avatar. Body scaling is an obvious possibility, since people could see the size of the self avatar relative to the height of the pole and of the ledge. It is possible that the self-avatar gives subjects concrete information about their eye-height. With this information, subjects may be able to compare the extrinsic information about pole or ledge height to their eye-height and determine a critical threshold based on this information. Some support for this view can be found in Chapter IV, which performed an eye-height manipulation in a virtual environment and found that affordance judgments closely tracked the eye-height manipulation. This work is roughly parallel to some of the findings of Mark (1987) who manipulated eye-height in his affordance studies of sitability and stair climbability and found that subjects adapted their thresholds according to their new eye-heights. Of course, it may be that several cues from the self-avatar and its presence in the environment are combined to improve affordance judgments (Ernst and Bülthoff, 2004; Hillis et al., 2002).

People seem to ignore the presence of the self-avatar in judging whether to duck in the task of ducking under a doorway. This result may be because there is dynamic geometric information that supercedes use of the avatar. When one walks toward a doorway that is higher than one's eyes, the visual angle to the top of the door rises relative to the horizon, whereas when one walks toward a doorway that is lower than ones eyes, the visual angle to the top of the door falls relative to the horizon. Additionally, it is difficult to calibrate the size of the self-avatar from the perspective of ducking.

Another finding was that people had lower thresholds for switching behavior when they performed the behavior as contrasted with when they simply reported what they intended to do. People had the possibility of receiving at least visual feedback when they performed the action, and the result may be a reflection of that. However, the feedback in the no avatar case is somewhat difficult to explain as they could not compare it to their body, they simply got a potential view of it as they crossed over or under. The result is also consistent with the results of studies of perception in the real world, showing lower thresholds and less accurate spatial judgments when people are asked to look and describe what they see verbally than when they look and then produce a physical action (Creem-Regehr and Kunz, 2010). For the stepping over or ducking under task, we regard ducking under as the more conservative behavior than stepping over due to the smaller likelihood of losing one's balance. Then in all experiments people chose the more conservative behavior when they must perform the action rather than when they simply report having to perform it. This finding is consistent with the results of Warren and Whang (1987) in their study of passable or not-passable aperture widths in the physical world. This result has potential implications in using virtual environments for learning and training, so that virtual environments that permit action may be preferable for such activities than other types of virtual environments.

Note that the experiments were not designed to exactly replicate conditions between a virtual environment and the real world. For example, we did not restrict the field-of-view in the real-world experiments to match that of the HMD displaying the virtual environment. There are a number of factors involved in experiencing a virtual environment, and these may include (for certain types of virtual environments) wearing a limited field-of-view HMD of significant mass and inertia to affect the dynamics of head motion (Willemsen et al., 2004), wearing a set of potentially cumbersome gear to allow the rendering of a self-avatar, or moving in awkward ways to navigate large virtual environments (Xie et al., 2010; Peck et al., 2012). Within these constraints, we would like to have performance in the virtual environment match performance in the real world, where the real world is experienced in a natural, unencumbered state. From the perspective of learning, training, and simulation, understanding how functionally similar these two scenarios are to one another, and being able to quantify the similarity, is the goal of this work.

Other perspectives are possible. From the perspective of cognitive science, virtual environment technologies enable precise control of the myriad perceptual cues at work in the physical world and deepen our understanding of how people use vision to decide how to act. Additionally, our self-representation controls our actions, interactions, and judgments, and changing it can affect these actions, interactions, and judgments (Witt et al., 2005; Stefanucci and Geuss, 2009). The present work demonstrates the importance of the self-representation, and how comparable performance to the real world can be achieved with a matched-size self-representation in a virtual environment. Collectively, these issues have been discussed in the virtual environments and psychology literature, e.g., (Kilteni et al., 2012; de Vignemont, 2011), and neural correlates to a sense of body ownership have been examined (Evans and Blanke, 2013; Blanke, 2012), but no detailed examination of them from the point of view of how they apply to virtual environments has been conducted.

In conclusion, our results support the utility of using affordances to judge the perceptual fidelity of virtual environments. Actions and action possibilities are important components of virtual environments when they are used for learning and training, and having them matched to the real world increases the similarity of the virtual environment to its real counterpart. The result that providing a matched self-avatar is a key factor enabling this increase in similarity between actions and action possibilities, at least for the two tasks tested here, is important information for designers of virtual environments. Future work should extend the investigation to other affordances and attempt to determine a theoretical and empirical foundation for why self-avatars are such a key factor. Examining which aspects of the self-avatar model are salient would seem to be a particularly relevant line.

V.2 Listed Contributions

The following are contributions of my work on distance perception:

1. In the real world, people perform better with larger field of view and people tend to walk further with the indirect walking method described in this thesis.

2. In immersive virtual environments, training before the distance task is important. Appropriate training does not eliminate distance underestimation, but it does improve people's performance on distance estimation significantly.

3. In immersive virtual environments, asking people to scan using the near-to-far scanning method significantly improve people's distance estimation. The underestimation error is reduced to less than 15% by requiring people to scan before the indirect blind walking. It is a new contribution on distance perception in immersive virtual environments. No previous papers have controlled the scanning before.

4. We did not find any effect of the self-avatar on distance perception as Mohler et al. (2010) did. We tried to compare the difference between no avatar and with the self-avatar under the condition of the direct walking with short distances, the indirect walking with short and long distances, with and without scanning. We did not find any effect of the self-avatar on distance perception in our immersive virtual environments.

The following are contributions of my work on affordances:

1. The presence of a self-avatar significantly improves people's performances in our affordance tasks of stepping over and ducking under a pole and stepping off a ledge. When the self-avatar is present, people's performances are closer to their behavior in the real world.

2. Requiring subjects to perform the action changed people's behavior in all the immersive virtual environment conditions in the task of stepping over or ducking under a pole and stepping off a ledge. Comparison to the results in the real world shows that

performing the action makes people's action closer to what they behave in the real world.

3. The presence of a physical pole reduced the threshold in the action condition in the task of stepping over or ducking under a pole. Subjects in an action condition may receive haptic and tactile sensations from the pole if they brush against it. These sensations provide additional feedback that might be used on a trial-to-trial basis to refine subjects' estimates of their affordance threshold, and may make subjects more conservative because it introduces a risk of falling.

V.3 Future Directions

I have opened up these areas for new research questions that I have not answered in this thesis.

V.3.1 Avatars in Immersive Virtual Environments

There is still discrepancy in findings regarding the effects of avatars on distance perception. This thesis looked very carefully at the issue and could not find an effect of the avatars. As mentioned, we did not find any effect of the self-avatar on distance estimation as Mohler et al. (2010) did. There are methodology differences between our experiment and their experiment. One research direction to determine which aspects of avatars have the most effect on distance perception or other kinds of user experience in virtual worlds is valuable for better use of avatars and design of immersive virtual environments. There are several questions we could ask about avatars when they are present in immersive virtual environments. Do avatars simply provide a familiar size cue when people interact with immersive virtual environments? Or do avatars provide a notion of embodiment when they are collocated with people in immersive virtual environments? Does the avatar need to be animated? Answering these questions might lead to answers to questions on the calibrations of avatars. How close does the calibration of self-avatars needs to be? We are using high cost motion capture equipment (Vicon motion capture equipment), which has low latency and high accuracy. Commodity level equipment, such as Kinect, can also provide self-avatars, but with more latency and less accuracy (Bailey and Bodenheimer, 2012). Then what kind of calibration is needed for people to effectively interact with immersive virtual environments is an interesting research topic. In terms of size, if avatars do provide a familiar size cue, the accurate calibration in body size needs to be guaranteed for people to use this familiarity. Then it becomes interesting to explore the effect of changing the size of avatars. In my work, we were very careful to calibrate the size of the self-avatar to the size of a person. But we don't know how close the calibration really needs to be for good judgments. Some work in this thesis examined changing the calibration by 15% and the affordance judgment changed about 15%. So the real question is how good the avatar needs to be in terms of model quality, animation quality to convey the properties of the environment. Besides the size information, the calibration may also affect the notion of embodiment. Can we give people a self-avatar of an NBA player and have them experience a different virtual reality of basketball or does it just seem fake and weird?

V.3.2 Training in Immersive Virtual Environments

Experiment Six in Chapter III shows that when people had training, their performance on distance perception is significantly improved. The training seems provide people more information during the task. Training does provide people more time to engage and interact with immersive virtual environments, which might enable recalibration of immersive virtual environments. But the complete reason of why training can improve distance estimation is still not clear. Then what kinds of training and how training really improve people's performance on distance perception is interesting questions to explore in immersive virtual environment.

V.3.3 The Effect of Action

Requiring people to perform the action in immersive virtual environments change people's behavior in stepping over or ducking under a pole and stepping off a ledge tasks. And the action makes people's performance closer to that in real world situation. The research to identify the importance of performing the action in different tasks is another research direction. And it might be valuable to examine why and how requiring people to do the action change people's behavior.

V.3.4 The User Experience in Immersive Virtual Environments

Chapter IV shows that the presence of self-avatar adds value to the two affordance tasks in this thesis but not on walking under the doorway nor on distance perception task. Then another interesting research topic could be to determine which aspects of the user experience in virtual environments can be improved or affected by the proper use of avatars.

V.3.5 Physical Objects in Virtual World Experiments

The presence of the physical pole make the stepping threshold lower in stepping over or ducking under a pole arise the question of what kinds of real world feedback are important when people interact with immersive virtual environments and make their experience closer to real world experience.

BIBLIOGRAPHY

ANDRE, J. and ROGERS, S. (2006). Using verbal and blind-walking distance estimates to investigate the two visual systems hypothesis. *Perception & Psychophysics*, 68(3):353–361.

BAILENSON, J., PATEL, K., NIELSEN, A., BAJCSY, R., JUNG, S.-H. and KURILLO, G. (2008). The effect of interactivity on learning physical actions in virtual reality. *Media Psychology*, 11:354–376.

BAILENSON, J. N. and BLASCOVICH, J. (2004). Avatars. *In* BAINBRIDGE, W. S., éditeur : *Encyclopedia of Human-Computer Interaction*, pages 64–68. Berkshire Publishing Group.

BAILENSON, J. N., BLASCOVICH, J., BEALL, A. C. and LOOMIS, J. M. (2001). Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators and Virtual Environments*, 10(6):583–598.

BAILENSON, J. N. and SEGOVIA, K. Y. (2010). Virtual doppelgangers: Psychological effects of avatars who ignore their owners. *In* BAINBRIDGE, W. S., éditeur : *Online worlds: Convergence of the real and the virtual*, pages 175–186. Springer, New York, NY.

BAILEY, S. W. and BODENHEIMER, B. (2012). A comparison of motion capture data recorded from a vicon system and a microsoft kinect sensor. *In Proceedings of the ACM Symposium on Applied Perception*, SAP '12, pages 121–121, New York, NY, USA. ACM.

BEALL, A. C., LOOMIS, J. M., PHILBECK, J. W. and FIKES, T. G. (1995). Absolute motion parallax weakly determines visual scale in real and virtual environments. *In Proc. SPIE*, volume 2411, pages 288–297.

BIOCCA, F. A. and ROLLAND, J. (1998). Virtual eyes can rearrange your body: Adaptation to visual displacement in see-through, head-mounted displays. *Presence: Teleoperators and Virtual Environments*, 7(3):262–277.

BLANKE, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, 13(8):556–571.

BLASCOVICH, J., LOOMIS, J., BEALL, A. C., SWINTH, K. R., HOYT, C. L. and BAILENSON, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2):103–124.

BODENHEIMER, B., MING, J., WU, H., NARASIMHAM, G., RUMP, B., MCNAMARA, T. P., CARR, T. H. and RIESER, J. J. (2007). Distance estimation in virtual and real environments using bisection. *In Symposium on Applied Perception in Graphics and Visualization*, pages 35–40, Tübingen, Germany.

CREEM-REGEHR, S. H. and KUNZ, B. R. (2010). Perception and action. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6):800–810.

CREEM-REGEHR, S. H., WILLEMSEN, P., GOOCH, A. A. and THOMPSON, W. B. (2005). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. *Perception*, 34(2):191–204.

CRUZ-NERA, c., SANDIN, D. J., DEFANTI, T. A., KENYON, R. V. and HART, J. C. (1992). The cave: audio visual experience automatic virtual environment. *Communication ACM*, 35(6):64–72.

CUTTING, J. E. and VISHTON, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth.

DALGARNO, B. and LEE, M. J. (2010). What are the learning affordances of 3-d virtual environments? *British Journal of Educational Technology*, 41(1):10–32.

de VIGNEMONT, F. (2011). Embodiment, ownership and disownership. *Consciousness and Cognition*, 20(1):82 – 93. Brain and Self: Bridging the Gap.

DODDS, T., MOHLER, B. and BULTHOFF, H. (2010). A communication task in hmd virtual environments: Speaker and listener movement improves communication. *In Proceedings of the 23rd Annual Conference on Computer Animation and Social Agents*, CASA 2010.

DRAPER, M. (1995). Exploring the influence of a virtual body on spatial awareness. Mémoire de D.E.A., Department of Engineering, University of Washington, Seattle, WA.

ELLIOTT, D. (1987). The influence of walking speed and prior practice on locomotor distance estimation. *Journal of Motor Behavior*, 19(4):476–485.

ENNIS, C., MCDONNELL, R. and O'SULLIVAN, C. (2010). Seeing is believing: body motion dominates in multisensory conversations. *ACM Trans. Graph.*, 29:91:1–91:9.

ERNST, M. O. and BÜLTHOFF, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8(4):162 – 169.

EVANS, N. and BLANKE, O. (2013). Shared electrophysiology mechanisms of body ownership and motor imagery. *NeuroImage*, 64(0):216 – 228.

FERWERDA, J. A. (2003). Three varieties of realism in computer graphics. *In Electronic Imaging 2003*, pages 290–297. International Society for Optics and Photonics.

FLACH, J. M. and HOLDEN, J. G. (1998). The reality of experience: Gibson's way. *Presence: Teleoperators and virtual environments*, 7(1):90–95.

FUKUSIMA, S. S., LOOMIS, J. M. and SILVA, J. A. D. (1997). Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1):86–100.

GARAU, M., SLATER, M., VINAYAGAMOORTHY, V., BROGNI, A., STEED, A. and SASSE, M. A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. *In Proc. SIGCHI Conference on Human Factors in Computing Systems*, pages 529–536.

GEUSS, M., STEFANUCCI, J., CREEM-REGEHR, S. and THOMPSON, W. B. (2010). Can i pass?: Using affordances to measure perceived size in virtual environments. *In Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, APGV '10, pages 61–64, New York, NY, USA. ACM.

GIBSON, E. J. and WALK, R. D. (1960). Visual cliff. *Scientific American*.

GIBSON, J. J. (1950). *The perception of the visual world*. Greenwood Press, Westport, Conn.

GIBSON, J. J. (1966). The senses considered as perceptual systems. *Leonardo*, 1(1):89.

GIBSON, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin, Boston, MA.

GILINSKY, A. S. (1951). Perceived size and distance in visual space. *Psychological Review*, 58(6):460.

GOODALE, M. A. and MILNER, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1):20–25.

GRASSI, M. and SORANZO, A. (2009). MLP: a matlab toolbox for rapid and reliable auditory threshold estimation. *Behavior Research Methods*, 41(1):21–28.

GRECHKIN, T. Y., NGUYEN, T. D., PLUMERT, J. M., CREMER, J. F. and KEARNEY, J. K. (2010). How does presentation method and measurement protocol affect distance estimation in real and virtual environments? *ACM Trans. Appl. Percept.*, 7:26:1–26:18.

GREEN, D. M. and SWETS, J. A. (1966). *Signal Detection Theory and Psychophysics*. John Wiley and Sons, New York.

GROSS, D. C., STANNEY, K. M. and COHN, L. J. (2005). Evoking affordances in virtual environments via sensory-stimuli substitution. *Presence: Teleoperators and Virtual Environments*, 14(4):482–491.

HALL, E. T. (1966). *The Hidden Dimension*. Anchor Books.

HARWAY, N. I. (1963). Judgment of distance in children and adults. *Journal of Experimental Psychology*, 65(4):385.

HAYDUK, L. A. (1983). Personal space: Where we now stand. *Psychological Bulletin*, 94(2):293–335.

HEIM, M. (1998). Virtual realism.

HILLIS, J. M., ERNST, M. O., BANKS, M. S. and LANDY, M. S. (2002). Combining sensory information: Mandatory fusion within, but not between, senses. *Science*, 298(5598):1627–1630.

HODGINS, J., JÖRG, S., O'SULLIVAN, C., PARK, S. I. and MAHLER, M. (2010). The saliency of anomalies in animated human characters. *ACM Trans. Appl. Percept.*, 7:22:1–22:14.

INTERRANTE, V., RIES, B., LINDQUIST, J. and ANDERSON, L. (2007). Elucidating the factors that can facilitate veridical spatial perception in immersive virtual environments. *In Proc. IEEE Symposium on 3D User Interfaces*, pages 11–18.

JONES, J. A., SWAN, II, J. E., SINGH, G., KOLSTAD, E. and ELLIS, S. R. (2008). The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. *In Proceedings of the 5th symposium on Applied perception in graphics and visualization*, APGV '08, pages 9–14, New York, NY, USA. ACM.

KELLY, J., DONALDSON, L., SJOLUND, L. and FREIBERG, J. (2013). More than just perception-action recalibration: Walking through a virtual environment causes rescaling of perceived space. *Attention, Perception and Psychophysics*, 75:1473–1485.

KELLY, J., HAMMEL, W., SIEGEL, Z. D. and SJOLUND, L. (2014). Recalibration of perceived distance in virtual environments occurs rapidly and transfers asymmetrically across scale. *IEEE Transactions on Visualization and Computer Graphics*, 20(4):588–595.

KELLY, J. W., MCNAMARA, T. P., BODENHEIMER, B., CARR, T. H. and RIESER, J. J. (2008). The shape of human navigation: How environmental geometry is used in maintenance of spatial orientation. *Cognition*, 109(2):281 – 286.

KELLY, J. W., MCNAMARA, T. P., BODENHEIMER, B., CARR, T. H. and RIESER, J. J. (2009). Individual differences in using geometric and featural cues to maintain spatial orientation: Cue quantity and cue ambiguity are more important than cue type. *Psychonomic Bulletin & Review*, 16(1):176–181.

KILTENI, K., GROTEN, R. and SLATER, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387.

KNAPP, J. M. and LOOMIS, J. M. (2004). Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(5):572–577.

KRETCH, K. S. and ADOLPH, K. E. (2013). Cliff or step? posture-specific learning at the edge of a drop-off. *Child Dev*, 84(1):226–240.

KUNZ, B. R., WOUTERS, L., SMITH, D., THOMPSON, W. B. and CREEM-REGEHR, S. H. (2009). Revisiting the effect of quality of graphics on distance judgments in virtual environments: A comparison of verbal reports and blind walking. *Attention, Perception, & Psychophysics*, 71(6):1284–1293.

LAPPIN, J. S., SHELTON, A. L. and RIESER, J. J. (2006). Environmental context influences visually perceived distance. *Perception & psychophysics*, 68(4):571–581.

LENGGENHAGER, B., MOUTHON, M. and BLANKE, O. (2009). Spatial aspects of bodily self-consciousness. *Consciousness & Cognition*, 18:110–117.

LENGGENHAGER, B., TADI, T., METZINGER, T. and BLANKE, O. (2007). Video ergo sum: Manipulating bodily self-consciousness. *Science*, 317(5841):1096.

LEPECQ, J.-C., BRINGOUX, L., PERGANDI, J.-M., COYLE, T. and MESTRE, D. (2009). Afforded actions as a behavioral assessment of physical presence in virtual environments. *Virtual reality*, 13(3):141–151.

LI, B., ZHANG, R. and KUHL, S. (2014). Minication affects action-based distance judgments in oculus rift hmds. *In Proceedings of the ACM Symposium on Applied Perception*, SAP '14, pages 91–94, New York, NY, USA. ACM.

LIN, Q., RIESER, J. and BODENHEIMER, B. (2012). Stepping over and ducking under: the influence of an avatar on locomotion in an hmd-based immersive virtual environment. *In Proceedings of the ACM Symposium on Applied Perception*, SAP '12, pages 7–10, New York, NY, USA. ACM.

LIN, Q., RIESER, J. J. and BODENHEIMER, B. (2013). Stepping off a ledge in an hmd-based immersive virtual environment. *In Proceedings of the ACM Symposium on Applied Perception*, SAP '13, pages 107–110, New York, NY, USA. ACM.

LOK, B., NAIK, S., WHITTON, M. and BROOKS, J. F. P. (2003). Effects of handling real objects and self-avatar fidelity on cognitive task performance and sense of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 12(6):615–628.

LOKKI, T. and GROHN, M. (2005). Navigation with auditory curs in a virtual environment. *IEEE MultiMedia*, 12(2):80–86.

LOOMIS, J. M. and BEALL, A. C. (1998). Visually controlled locomotion: Its dependence on optic flow, three-dimensional space perception, and cognition. *Ecological Psychology*, 10(3-4):271–285.

LOOMIS, J. M., DA SILVA, J. A., PHILBECK, J. W. and FUKUSIMA, S. S. (1996). Visual perception of location and distance. *Current Directions in Psychological Science*, 5(3): 72–77.

LOOMIS, J. M., KLATZKY, R. L., PHILBECK, J. W. and GOLLEDGE, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Perception & Psychophysics*, 60(6):966–980.

LOOMIS, J. M. and KNAPP, J. (2003). Visual perception of egocentric distance in real and virtual environments. *In* HETTINGER, L. J. and HAAS, M. W., éditeurs : *Virtual and Adaptive Environments*, chapitre 2, pages 21–46. Erlbaum, Mahwah, NJ.

LOOMIS, J. M., LIPPA, Y., KLATZKY, R. L. and GOLLEDGE, R. G. (2002). Spatial updating of locations specified by 3-d sound and spatial language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(2):335.

LOOMIS, J. M., SILVA, J. A. D., FUJITA, N. and FUKUSIMA, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4):906–921.

MARK, L. S. (1987). Eyeheight-scaled information about affordances: a study of sitting and stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3):361.

MARK, L. S., BALLIETT, J. A., CRAVER, K. D., DOUGLAS, S. D. and FOX, T. (1990). What an actor must do in order to perceive the affordance for sitting. *Ecological Psychology*, 2(4):325–366.

MCMANUS, E. A., BODENHEIMER, B., STREUBER, S., de la ROSA, S., BÜLTHOFF, H. H. and MOHLER, B. J. (2011). The influence of avatar (self and character) animations on distance estimation, object interaction and locomotion in immersive virtual environments. *In Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, APGV '11, pages 37–44, New York, NY, USA. ACM.

MEEHAN, M., INSKO, B., WHITTON, M. and BROOKS, Jr., F. P. (2002). Physiological measures of presence in stressful virtual environments. *ACM Trans. Graph.*, 21:645–652.

MESSING, R. and DURGIN, F. H. (2005). Distance perception and the visual horizon in head-mounted displays. *ACM Trans. Appl. Percept.*, 2:234–250.

MICHAELS, C. F. (2003). Affordances: Four points of debate. *Ecological Psychology*, 15(2):135–148.

MILNER, A. and GOODALE, M. (2008). Two visual systems re-viewed. *Neuropsychologia*, 46(3):774 – 785.

MOHLER, B. J., CREEM-REGEHR, S. H. and THOMPSON, W. B. (2006). The influence of feedback on egocentric distance judgments in real and virtual environments. *In Proc. Third Symposium on Applied Perception in Graphics and Visualization*, pages 9–14.

MOHLER, B. J., CREEM-REGEHR, S. H., THOMPSON, W. B. and BÜLTHOFF, H. H. (2010). The effect of viewing a self-avatar on distance judgments in an HMD-based virtual environment. *Presence: Teleoperators and Virtual Environments*, 19(3):230–242.

MOHLER, B. J., THOMPSON, W. B., CREEM-REGEHR, S., HERBERT L. PICK, JR., WILLIAM WARREN, JR., RIESER, J. J. and WILLEMSEN, P. (2004). Visual motion influences locomotion in a treadmill virtual environment. *In Proc. First Symposium on Applied Perception in Graphics and Visualization*, pages 19–22.

NORMAN, D. A. (2002). *The design of everyday things*. Basic books.

OOI, T. L., WU, B. and HE, Z. J. (2001). Distance determined by the angular declination below the horizon. *Nature*, 414(6860):197–200.

PALMER, S. E. (1999). *Vision science : photons to phenomenology*. MIT Press.

PECK, T., FUCHS, H. and WHITTON, M. (2012). The design and evaluation of a large-scale real-walking locomotion interface. *Visualization and Computer Graphics, IEEE Transactions on*, 18(7):1053–1067.

PHILBECK, J. W. and LOOMIS, J. M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1):72.

PHILLIPS, L., RIES, B., KAEDING, M. and INTERRANTE, V. (2010). Avatar self-embodiment enhances distance perception accuracy in non-photorealistic immersive virtual environments. *In Proc. IEEE Virtual Reality Conference*, pages 115–118.

PLUMERT, J. M., KEARNEY, J. K. and CREMER, J. F. (2004). Distance perception in real and virtual environments. *ACM Transactions on Applied Percpetion*, 2:216–233.

PUFALL, P. B. and DUNBAR, C. (1992). Perceiving whether or not the world affords stepping onto and over: A developmental study. 4:17–38.

REGIA-CORTE, T., MARCHAL, M., CIRIO, G. and LÉCUYER, A. (2013). Perceiving affordances in virtual reality: influence of person and environmental properties in perception of standing on virtual grounds. *Virtual Reality*, 17(1):17–28.

RICHARDSON, A. R. and WALLER, D. (2005). The effect of feedback training on distance estimation in virtual environments. *Applied Cognitive Psychology*, 19:1089–1108.

RICHARDSON, A. R. and WALLER, D. (2007). Interaction with an immersive virtual environment corrects users' distance estimates. *Human Factors*, 49(3):507–517.

RIES, B., INTERRANTE, V., KAEDING, M. and ANDERSON, L. (2008). The effect of self-embodiment on distance perception in immersive virtual environments. *In Proc. ACM Symposium on Virtual Reality Software and Technology*, pages 167–170.

RIES, B., INTERRANTE, V., KAEDING, M. and PHILLIPS, L. (2009). Analyzing the effect of a virtual avatar's geometric and motion fidelity on ego-centric spatial perception in immersive virtual environments. *In Proc. ACM Symposium on Virtual Reality Software and Technology (VRST)*, pages 59–66.

RIESER, J. J., ASHMEAD, D. H., TAYOR, C. R. and YOUNGQUIST, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception*, 19:675–689.

RIESER, J. J. and PICK, H. L. (2007). Using locomotion to update spatial orientation: What changes with learning and development? *In* PLUMERT, J. and SPENCER, J., éditeurs : *The Emerging Spatial Mind*, pages 77–103. Oxford University Press.

RIZZO, A. A., KLIMCHUK, D., MITURA, R., BOWERLY, T., BUCKWALTER, J. G., R, K., ADAMS, R., FINN, P., TARNANAS, I., OLLENDICK, T. H. and ching YEH, S. (2006). A virtual reality scenario for all seasons: the virtual classroom. *CNS Spectrum*, 11:35–44.

ROTHBAUM, B. O. and HODGES, L. F. (1999). The use of virtual reality exposure in the treatment of anxiety disorders. *Behavior Modification*, 23(4):507–525.

ROTHBAUM, B. O. and HODGES, L. F. (2000). A controlled study of virtual reality exposure therapy for the fear of flying. *Journal of Consulting and Clinical Psychology*, 68(6):1020–1026.

SCHWAIGER, M., THUMMEL, T. and ULBRICH, H. (2007). Cyberwalk: An advanced prototype of a belt array platform. *In Haptic, Audio and Visual Environments and Games, 2007. HAVE 2007. IEEE International Workshop on*, pages 50–55. IEEE.

SEDGWICK, H. A. (1983). Environment-centered representation of spatial layout: Available information from texture and perspective. *Human and Machine Vision*, pages 425 – 428.

SHERMAN, W. R. and CRAIG, A. B. (2003). Understanding virtual reality: Interface, application, and design.

SINAI, M. J., OOI, T. L. and HE, Z. J. (1998). Terrain influences the accurate judgment of distance. *Nature*, 395:497–500.

SLATER, M., KHANNA, P., MORTENSEN, J. and YU, I. (2009a). Visual realism enhances realistic response in an immersive virtual environment. *Computer Graphics and Applications, IEEE*, 29(3):76–84.

SLATER, M., PEREZ-MARCOS, D., EHRSSON, H. H. and SANCHES-VIVES, M. V. (2009b). Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience*, 3:214–220.

SLATER, M., PEREZ-MARCOS, D., EHRSSON, H. H. and SANCHEZ-VIVES, M. V. (2008). Towards a digital body: the virtual arm illusion. *Frontiers in Human Neuroscience*, 2:1–8.

SLATER, M. and USOH, M. (1994). Body centered interaction in immersive virtual eneironments. *In* THALMANN, N. M. and THALMANN, D., éditeurs : *Artificial Life and Virtual Reality*, pages 125–148. John Wiley and Sons.

SLATER, M., USOH, M. and STEED, A. (1995). Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction*, 2(3):201–219.

SLATER, M. and WILBUR, S. (1997). A framework for immersive virtual environments (five): Speculation on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 6:603–616.

SMETS, G., STAPPERS, P., OVERBEEKE, K. and van der MAST, C. (1995). Designing in virtual reality: Perception-action coupling and affordances. *Simulated and Virtual Realities: Elements of Perception*, pages 189–208.

STEFANUCCI, J. K. and GEUSS, M. N. (2009). Big people, little world: The body influences size perception. *Perception*, 38(12):1782–1795.

STEFANUCCI, J. K. and GEUSS, M. N. (2010). Duck! scaling the height of a horizontal barrier to body height. *Attention, Perception, & Psychophysics*, 72(5):1338–1349.

STEFANUCCI, J. K. and PROFFITT, D. R. (2009). The roles of altitude and fear in the perception of height. *Journal of Experimental Psychology: Human Perception and Performance*, 35(2):424.

STREUBER, S., de la ROSA, S., TRUTOIU, L., BÜLTHOFF, H. H. and MOHLER, B. J. (2009). Does brief exposure to a self-avatar affect common human behaviors in immersive virtual environments? *Eurographics 2009*, pages 1–4.

THELEN, E. (1995). Motor development: A new synthesis. *American psychologist*, 50(2): 79.

THOMPSON, W. B., WILLEMSEN, P., GOOCH, A. A., CREEM-REGEHR, S. H., LOOMIS, J. M. and BEALL, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments*, 13(5):560–571.

USOH, M., ARTHUR, K., WHITTON, M. C., BASTOS, R., STEED, A., SLATER, M. and BROOKS, JR., F. P. (1999). Walking > walking-in-place > flying in virtual environments. *In Proc. ACM SIGGRAPH*, pages 359–364.

van WYK, E. and de VILLIERS, R. (2009). Virtual reality training applications for the mining industry. *In Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, AFRIGRAPH '09, pages 53–63, New York, NY, USA. ACM.

WALLER, D. and RICHARDSON, A. (2008). Correcting distance estimate by interacting with immersive virtual environments: Effects of task and available sensory information. *Journal of Experimental Psychology: Applied*, 14:61–72.

WARREN, W. H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, 10:683–703.

WARREN, W. H. and WHANG, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 13:371–383.

WILLEMSEN, P., COLTON, M. B., CREEM-REGEHR, S. H. and THOMPSON, W. B. (2004). The effects of head-mounted display mechanics on distance judgments in virtual environments. *In Proc. First Symposium on Applied Perception in Graphics and Visualization*, pages 35–38.

WILLEMSEN, P., COLTON, M. B., CREEM-REGEHR, S. H. and THOMPSON, W. B. (2009). The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments. *ACM Transactions on Applied Perception (TAP)*, 6(2):8.

WILLEMSEN, P., GOOCH, A. A., THOMPSON, W. B. and CREEM-REGEHR, S. H. (2008). Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(1):91–101.

WILLIAMS, B., NARASIMHAM, G., WESTERMAN, C., RIESER, J. and BODENHEIMER, B. (2007). Functional similarities in spatial representations between real and virtual environments. *Transactions on Applied Perception*, 4(2).

WITMER, B. G. and SADOWSKI, W. J. J. (1998). Nonvisually guided locomotion to a previously viewed target in real and virutal environments. *Human Factors*, 40:478–488.

WITMER, B. G. and SINGER, M. J. (1998). Measuing presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240.

WITT, J. K., PROFFITT, D. R. and EPSTEIN, W. (2004). Perceiving distance: A role of effort and intent. *PERCEPTION-LONDON-*, 33:577–590.

WITT, J. K., PROFFITT, D. R. and EPSTEIN, W. (2005). Tool use affects perceived distance but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance*, 31:880–888.

WU, B., OOI, T. L. and HE, Z. J. (2004). Perceiving distance accurately by a directional process of integrating ground information. *Nature*, 428:73–77.

WU, H., ASHMEAD, D. H. and BODENHEIMER, B. (2009). Using immersive virtual reality to evaluate pedestrian street crossing decisions at a roundabout. *In Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, APGV '09, pages 35–40, New York, NY, USA. ACM.

XIE, X., LIN, Q., WU, H., NARASIMHAM, G., MCNAMARA, T. P., RIESER, J. and BODENHEIMER, B. (2010). A system for exploring large virtual environments that combines scaled translational gain and interventions. *In Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, APGV '10, pages 65–72, New York, NY, USA. ACM.

YOUNG, M. K., GAYLOR, G. B., ANDRUS, S. M. and BODENHEIMER, B. (2014). A comparison of two cost-differentiated virtual reality systems for perception and action tasks. *In Proceedings of the ACM Symposium on Applied Perception*, SAP '14, pages 83–90, New York, NY, USA. ACM.

ZAHORIK, P. and JENISON, R. L. (1998). Presence as being-in-the-world. *Presence: Teleoperators and virtual environments*, 7(1):78–89.

ZIMMONS, P. and PANTER, A. (2003). The influence of rendering quality on presence and task performance in a virtual environment. *In Virtual Reality, 2003. Proceedings. IEEE*, pages 293–294.