HUMAN PERFORMANCE AND THE PERCEPTION OF ACTIONS IN IMMERSIVE VIRTUAL ENVIRONMENTS

By

Erin McManus

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Computer Science

May, 2012

Nashville, Tennessee

Approved:

Professor Robert E. Bodenheimer

Professor Aniruddha Gokhale

To my parents, Natalie and Richard McManus,

the most trusting, supportive and loving people that I know

ACKNOWLEDGMENTS

I would like to thank the Cornelius Vanderbilt Lanier Honor Scholarship Program, the CRA-W's CREU program, the Max Planck Society and the National Science Foundation through grant 0705863 for the financial support necessary to complete this thesis.

I thank Trevor Dodds, Cengiz Tzeribas, Joachim Tesch, Ivelina Alexandrova, and Markus Leyrer for help with programming the experiments and data collection.

To all of my co-authors, Stephan Streuber, Stephan de la Rosa, and Heinrich Bülthoff, Qiufeng Lin, Stephen Bailey, Aysu Erdemir, and John Rieser, your help with planning, writing, analyzing and reviewing has been greatly appreciated.

I would also like to thank my two advisers, Dr. Betty Mohler and Dr. Bobby Bodenheimer, without whom this work would not have been possible. Betty took me on as an intern for three months in Tübingen, Germany without having ever spoken to me and I am so very grateful for that leap of faith and her continued investment of both time and support for my well-being. And from my first meeting with Dr. Bodenheimer, when he unquestioningly added me to his research team, to our now weekly Monday meetings, he has unfalteringly provided me with the perfect balance of direction, pressure and understanding which has shaped my college career and future for the better.

And finally, to my friends and family who have remained supportive and loving throughout this entire journey–I could never have done it without you.

Table of Contents

List of Tables

Table | Page

List of Figures

Figure | Page

Chapter I

INTRODUCTION

An immersive virtual environment (VE) can be a world where the laws of physics no longer apply and preconceptions about time and space can be broken, creating opportunities to explore aspects of human behavior and perception that are otherwise unreachable. There are numerous applications for virtual reality, including using it as a training tool in areas such as the military and medical professions; in these scenarios, and regardless of the environment that has been created, human performance should closely mirror real world behavior so that learning and training transfer can occur. Due to various limitations of virtual reality, such as discrepancies between feedback from different senses and a limited field of view, human behavior is negatively affected in a number of different areas such as distance perception [Loomis and Knapp 2003; Thompson et al. 2004; Bodenheimer et al. 2007; Jones et al. 2008; Ries et al. 2009; Grechkin et al. 2010; Mohler et al. 2010b] and walking gait [Bodenheimer et al. 2007; Mohler et al. 2007]. The two studies described in this paper are aimed at gaining a better understanding of human perception and action and how using this knowledge can be applied to improving performance within a VE. A common theme among all of the experiments performed throughout both studies is object interaction, as it is important to study not only human interaction within the environment, but to also investigate human interaction *with* the environment.

VEs are high-performance real-time environments. They are important because they can place people into scenarios that are either expensive or impossible to replicate otherwise. They take the place of real-life situations and therefore must mirror the real world as closely as possible. Not only should the environment be aesthetically realistic, but it must also be designed in such a way that performance, behavior and perception are not altered. In order to achieve this form of realism, factors such as latencies and calibration must be taken into consideration. Due to the computational complexity of virtual reality, latency is inevitable; the lag, however, can be minimized using different techniques to decrease the polygon count of models in the scene and using high-speed network connections. Other factors that are inevitably going to place limitations on the degree of fidelity are the restricted field of view and the weight of the head-mounted display as well as the availability of real-world tracking space and sensory feedback (sounds, smells, haptic information). With all of these limitations in mind, overcoming them becomes an important task and a large part of the motivation behind this work.

We first describe a self-avatar study composed of four experiments, one in the VE and three in the real world, exploring human performance and behavior on three different tasks: distance estimation, object interaction and stepping stone locomotion. All tasks were performed in both the real and virtual worlds in order to allow for comparisons to be made. It is known that humans behave differently in the virtual world than in the real one in a number of ways, and if it is ever going to be widely used in applications such as military or medical training, or even entertainment, then this gap needs to be bridged. Experiment 1a asked the question, does adding an animated self-avatar or other character to the VE improve performance on any of three tasks? Due to improvements in virtual reality technology, providing a self-avatar or other human characters to an environment is now both feasible, and possibly necessary. Until recently, the technology to effectively add a self-avatar to a virtual scene did not exist; it requires sophisticated hardware and is computationally expensive for a real-time system. Using Xsens MVN Motion Capture Suits, however, we were able to provide a self-avatar with a latency of only ∼120ms. We used a 2x2 experimental design, varying both the learning phase animation (cone or character) and task phase animation (self-avatar or no self-avatar). Subjects were asked to watch a task being performed by either an animated human character or a cone animation for about 1 minute in the learning phase, and then asked to perform the same task directly afterwards and were either provided with a self-avatar or not. Results show that adding either an animated character or a self-avatar to the scene will improve task performance for tasks involving visual interaction with the environment (the distance estimation task was direct blind walking and therefore showed no improvement across conditions), but the effects are not additive. We also find learning, regardless of condition, occurs over time on all three tasks. Experiments 1b, 1c and 1d were the distance estimation, object interaction and stepping stone locomotion tasks, each performed in the real world. We find that, as expected, subjects perform significantly better in the real world than in the virtual one, but we were pleased to see how close the virtual world results were to the real-world performance.

We next describe a study composed of three experiments, all performed in the virtual world, that are designed to investigate perception and action. Experiment 2a was motivated by two experiments performed by Reitsma et al. [2003; 2008] who introduced errors into the trajectory of a ball being shot out of a cannon as well as that of a human character leaping and asked participants to detect the trajectories with errors. Errors were introduced to the initial vertical or horizontal velocities of the trajectory or to the magnitude of the gravitational constant acting on the object as it traveled along its flight path. Reitsma et al. found that users could discriminate horizontal velocity errors more easily than vertical velocity errors and that preparatory motion associated with these trajectories helped to increase users' sensitivities to the errors. Our study differs from Reitsma et al.'s in two important ways. We placed the subject into an active role by replacing the third-person point-of-view used by Reitsma et al. with a first-person perspective and the preparatory motion with the motion of the user throwing a ball underhandedly. We find no difference between subjects' ability to discriminate between vertical and horizontal velocity perturbations and suggest that this may be due to the added involvement of the user in the action. The second experiment in this set builds upon the first and explores the importance of the ending point of the ball when making the judgments from Experiment 2a. In order to do this, the trajectory of the ball was hidden and users had to discern between two endpoint options as to which ball they had thrown. We find that users' detection thresholds do not vary significantly from those of Experiment 2a and therefore conclude that while the endpoint does seem to be a contributing factor to making a judgment, we do not believe that it is the only factor. Finally, after finding that motor information, along with endpoint location of the ball, allows users to correctly identify their own ball, we wanted to investigate the importance of visual and motor information in Experiment 2c. Experiment 2c placed users into the roles of both actor and observer, as they were able to see an avatar representing themselves or a character representing someone else throw the ball , while performing the same task as in Experiment 2b. We find no effect of role on the users' discrimination thresholds and therefore conclude that visual information alone is good enough to perform the task, and that motor information adds no significant improvement.

This thesis is organized as follows: chapter 2 presents the self-avatar study, chapter 3 presents the throwing study and chapter 4 contains concluding remarks.

Chapter II

THE INFLUENCE OF HUMAN ANIMATIONS ON ACTION PERFORMANCE IN IMMERSIVE VIRTUAL ENVIRONMENTS

2.1 Abstract

2.1.1 Background

Humans have been shown to perform actions differently in immersive virtual environments as compared to the real world. Immersive virtual environments often lack the presence of virtual characters; users are rarely presented with a representation of their own body (self-avatar) and have little or no experience with other human avatars/characters. However, virtual characters and avatars are becoming more common in immersive virtual environments.

2.1.2 Principal Findings

In four experiments, we explored real and virtual world task performance for common tasks that are commonly performed in the real world. In particular, we examined performance on three different behavioral tasks: participants performed distance estimation, object interaction, and stepping stone locomotion tasks within the virtual environment or in the real world. In a two-phase virtual reality experiment (Experiment 1a), we investigated the impact of seeing an animated character or a self-avatar in a head-mounted display virtual environment on task performance in

the virtual environment. In a learning phase, participants either saw a character animation or an animation of an abstract geometric shape, a cone. In the task performance phase, we varied whether participants saw a co-located animated self-avatar. In Experiments 1b, 1c, and 1d, we measured real world performance in the same three behavioral tasks as a baseline for comparison.

2.1.3 Conclusions

In Experiment 1a, we find that both the character animation and the self-avatar influenced task performance for the object interaction and the stepping stone tasks. Overall the participants performed the tasks faster and more accurately when they either had a self-avatar or saw a character animation. We find no impact of the character animation or the self-avatar on distance estimates. Finally, we see that for all tasks in the virtual environment performance improves with repetition. Our real world baseline results demonstrate that participants are still not able to perform these tasks in the virtual world with the same speed and/or accuracy as they do in the real world. However, we are encouraged by how close their virtual and real world behavior is and suggest that this type of evaluation is a good method for determining the similarity of behavior in real and virtual worlds. The results of our virtual reality experiment indicate that including character animations or self-avatars before or during task execution is beneficial to performance of some common interaction tasks that involve the on-line visual control of action within the virtual environment. To enable the users of virtual environments to act more closely to how they would in the real world, we suggest that developers of such environments add animations of virtual humans (self-avatars or characters) and allow users to gain experience with the environments before performing complex tasks.

2.2 Introduction

Virtual environments offer humans the ability to experience, train, and learn in situations that are different from their local surroundings, and particularly offer the opportunity to encounter settings that are difficult or expensive to simulate in the real world [Jayaram et al. 1995; Rizzo et al. 2006; van Wyk and de Villiers 2009]. For virtual environment technology to be successful at this task, enough fidelity in the environment that humans behave similarly in the virtual and real environments is desirable. However, it is has been shown that humans behave differently in virtual environments when compared to the real world in a number of ways. For example, many studies have shown that humans underestimate egocentric distances in head-mounted display (HMD) VEs [Loomis and Knapp 2003; Thompson et al. 2004; Bodenheimer et al. 2007; Jones et al. 2008; Ries et al. 2009; Grechkin et al. 2010; Mohler et al. 2010b], with the studies reporting varying amounts of distance compression from 6% to over 20%. Further, gait parameters (stride length, walking velocity and head-trunk angle) are different when humans walk in a VE compared to the real world [Bodenheimer et al. 2007; Mohler et al. 2007]. There are also discrepancies in completion time and behavior when performing simple tasks in the real world and the VE [Streuber et al. 2009; Williams et al. 2006].

Integrating dynamic virtual humans into complex real-time graphical environments is a challenging research problem that continues to have open questions [O'Sullivan and Ennis 2011; McDonnell et al. 2009]. Representing realistic real-time avatars presents difficulties in part because of technical challenges in rendering their behavior and appearance, and in part because it is unknown the extent to which users of such environments can profit from seeing a visual representation of their own body (self-avatar) or of other human bodies (characters), and therefore what effort should be expended in providing these features over other features. However, technical advances in the ability to represent realistic self-avatars in real-time make investigating their potential benefits and limits a high priority. For instance, self-avatars have been used to systematically investigate physical and emotional presence and interaction and communication in VEs [Dodds et al. 2010; Durlach and Slater 2000; Lok 2004; Slater and Usoh 1994].

For which situations or tasks might having an avatar prove helpful? Candidate tasks are visually guided action coordination tasks where one's own limb movements are changing relative to the physical environment and require temporal and spatial adjustments to successfully conduct a task, e.g. when grasping a cup. Support for this idea comes from several studies that show that the visual feedback about the object and/or limb movements influences task performance. For example, in the real world the visibility of one's own arm leads to improved reaching accuracy [Bard et al. 1985; Spijkers and Spellerberg 1995; Proteau et al. 2000] and faster adjustments of incorrect arm movements [Reichenbach et al. 2009]. Additionally, it has been shown that for adaptation of visual-vestibular mapping (reaching and whole body movements) a first person or a mirror view of a self-avatar is preferred for reducing variability of motor performance as compared to the top-down view [Schomaker et al. 2011]. Further, seeing one's own arm movements leads to improved social interaction performance in a real world table tennis task [Streuber et al. 2011]. This leads to the hypothesis that seeing a virtual representation of one's own body movements (such as a self-avatar) also leads to performance improvements within a VE.

On the other hand, seeing another avatar doing a task allows for mental preparation of the task, for example, by imitating a successful solution or by recognizing and starting to solve tasks constraints prior to the execution of the task. Hence, observing an animated character (even an abstract one) performing a specific task before the user performs the identical task should help improve performance. There is evidence that this type of learning by imitating body movements is an important building block in social interactions because it allows a person to transfer knowledge about the function and meaning of objects within the context of a specific task [Prinz ; Saunders and Knill ; Sebanz et al. ; Streuber and de la Rosa 2010].

Finally, seeing an avatar (regardless whether this is one's own or another person's avatar) may provide familiar size cues that can be used to improve spatial or distance judgments. Recent research has suggested that when spatial perception is involved, people may obtain familiar size cues from other human beings, even themselves, to scale virtual space [Henry and Furness 1993; Mohler et al. 2010b]. Participants significantly improved distance judgments when a static avatar was presented within the VE prior to the distance estimation task [Mohler et al. 2010b; Mohler et al. 2008; Ries et al. 2008]. Mohler et al. [2010b] provided a fully animated self-avatar to participants that yielded similar results, but with an even greater increase in task performance. They suggest that this increased improvement could be due to visual-motor feedback about body location over time. Interestingly, an avatar with reduced visual or motion fidelity (point lights or a stiff self-avatar) was shown not to influence egocentric distance judgments [Ries et al. 2009]. Similarly, the effect of showing a pre-recorded animation of an avatar on distance estimations has also been shown not to have a significant impact on egocentric distance judgments [Mohler et al. 2010a]. The full explanation for why and how an animated character or avatar (before task execution) might influence distance estimates in HMD-VEs is still not fully understood.

In sum, animations of virtual humans might enhance user performance in various ways. A self-avatar might lead to improved visual monitoring of own actions and/or increased presence in the virtual space that in turn might lead to better task performance in VEs. User performance might also profit from observing an instructional avatar prior to task performance because it might inform the user about how a task can be solved by means of a combination of goal directed body movements rather than a combination of verbal instructions. In addition, virtual self or character animations might provide familiar size cues that are used for task performance, specifically spatial estimates. The previous research suggests that presenting self and character animations in VEs should both enhance user performance, unless the visual mapping between self-movement and visual motion is essential to performance. This study seeks to test the influence of self and character animations on user's performance in three tasks: estimating distances, stepping stones and object interaction.

2.3 Materials and Methods

2.3.1 Experiment 1a: HMD VE Experiment

2.3.1.1 Participants

Thirty-one participants, 20 male and 11 female, participated in Experiment 1a. Their age ranged from 22 to 56 (mean = 31.5, SD = 7.5). All participants were volunteers from the local university community, had normal or corrected-to-normal stereo vision and received standard monetary compensation. Participants were pre-screened for a height that would fit into our medium or large Xsens MVN motion capture suit (see Figure 2.1). Participants' height ranged from 1.62 to 1.87 meters (mean= 1.73, SD =0.07). Participants' height, arm span, and foot length were also measured for the Xsens MVN motion capture suit software to more accurately estimate the position and orientation of the body and in the case of height and arm span for appropriate scaling of the self-avatar in the virtual world. Participants had an arm span which ranged from 1.57 to 1.98 meters (mean = 1.73, SD = 0.10) and a foot length which ranged from 0.21 to 0.30 meters (mean = 0.25, SD = 0.02).

2.3.1.2 Experimental Setup

Each participant viewed the scene through an nVis nVisor SX60 HMD that displays a stereoscopic image of the virtual world with a resolution of 1280 x 1024 pixels per eye, a frame rate of 60 Hz for each eye, and manufacturer-specified FOV of 60 degrees diagonally. Absolute tracking for the head and pelvis of the participant was provided by a motion capture system with 16 Vicon MX13 cameras and Vicon IQ 2.5 software. The VE was run in Virtools 4.1 on a Dell Inspiron M6400 laptop with an Intel Core2 Duo 2.53 GHz processor. The graphics were rendered on an nVidia Quadro FX3700M with 1024 MB of RAM. The end-to-end latency for head and pelvis vicon tracking, network and rendering was approximately ~50 ms for this experiment (using the method described by Di Luca et al. [2010]). The study was carried out in a fully tracked free-walking space, 11.13m width x 12.60m length x 8.4m height.

Participants' relative body orientation and position were tracked with an Xsens MVN motion capture suit (see Figure 2.1). This motion capture suit consists of 17 inertial and magnetic sensor modules. The captured data is transmitted via a wireless connection to a laptop computer (a second computer with the same specifications as the visualization computer) on which the processing is performed and transmitted to the visualization computer. The suit is used for quick and convenient placement of sensors and cables (setup and calibration took approximately 10-15 minutes for each participant). The system runs in real-time with a maximum update rate for all kinematics of 120 Hz and an approximate latency of 120 ms (using the method described by Di Luca et al. [2010]). Since Moven does not provide reliable absolute position and orientation information, the pelvis was tracked using Vicon technology. For the pelvis, only heading values for orientation were used, as the other orientation values for the pelvis from the motion capture suit were sufficient and necessary for rendering the self-avatar as accurately as possible.

Figure 2.1: (left) Xsens MVN motion capture suit used to capture the motion of each participant. (right) The unisex character used for the self-avatar during the task phase.

Pre-recorded animations of a virtual character performing the three tasks were made using Xsens MVN motion capture suit data captured from the first author. This data was then applied to an accurately scaled RocketBox avatar in 3DStudio Max 2009. Object data in the case of the object interaction task was captured using Vicon data and this data was used to animate the tool object used in the task. For the distance estimation task, the chosen animation was an avatar walking 7 meters at 1.4 m/s to a target on the ground plane. For the object interaction and stepping stone tasks, the character animation was of the first author completing the task as quickly as she could (15.3 and 5.5 seconds, respectively). In addition to the character animations, animations were made which were intended to retain the instructional and timing value of the character animations but not the biological motion and familiar size of an animated virtual character. These animations were achieved by using a green 3D cone model and the invisible character animations in Virtools 4.1. We fixed the height of the green cone at 1.7 meters in height and made the cone follow the position of the character's head in the X (lateral) and Z (depth) planes. The character was not visible as we played the cone animations. See Figures 2.2 to 2.4 for representative screen shots of all the character (left) and cone (right) animations.

There were two different environments in which the participants completed the three tasks. The environment used for the distance estimation task was an office room (10.27m length x 7.25m width x 2.77m height) with a starting line and a green target placed on the floor (see Figure 2.5). Both the stepping stone and the object interaction tasks were completed in an environment which consisted of a 12 x 12 meter

Figure 2.2: Learning phase animations for the distance estimation task: The animated character avatar walking towards the green target (left). The green cone moving in the xz-plane of the invisible character (right).

wooden floor along with the objects needed to complete each task (see Figure 2.6). Both environments contained a full-length mirror with a black frame located along the west wall, which was seen only during the task performance phase in all conditions and used to help the participants more easily see the self-avatar.

During the entire learning phase, the trajectories of the person's head and pelvis as well as their full body motion (from the Xsens MVN motion capture suits) were recorded for additional analysis of the participants' viewing behavior.

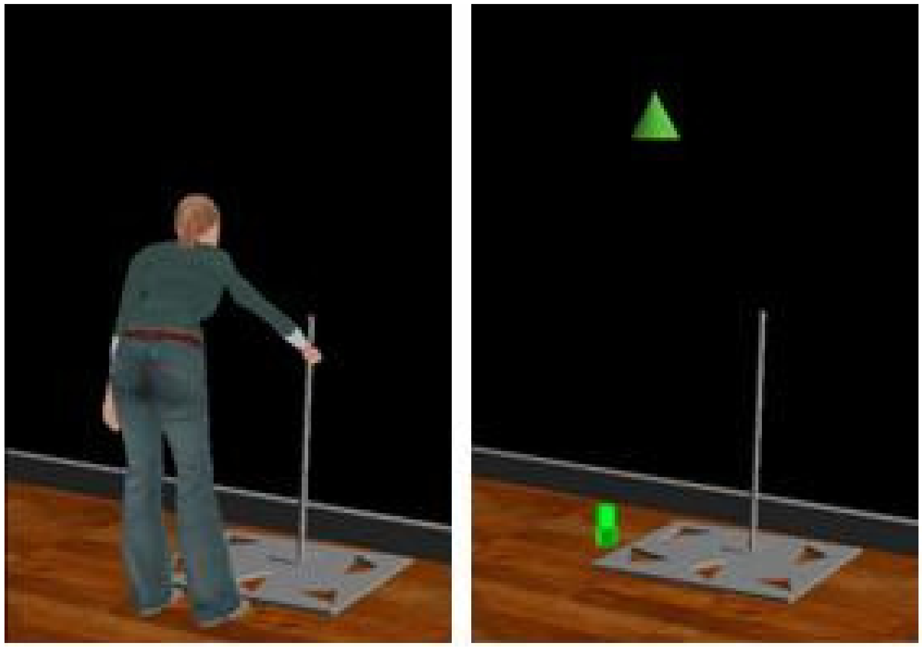

Figure 2.3: Learning phase animations for the object interaction task: The animated character avatar placing the stamp tool into the second of six triangle orientations cut out of the board (left). The green cone moving in the xz-plane of the head of the invisible character (right).

2.3.1.3 Task Performance Phase

In the task performance phase of the experiment, participants in half of the conditions were able to see first person, real-time animations of a virtual self-animated avatar, and could use the mirrors in each environment to see their entire virtual self-avatar. In these conditions, a unisex avatar (see Figure 2.1) was used to represent each participant. Those participants who saw their self-avatar could interact with the self-avatar in the mirror for as long as they liked before beginning the task and in all but the distance estimation task (which was performed without vision) they could view their own body during task execution. The other participants were not

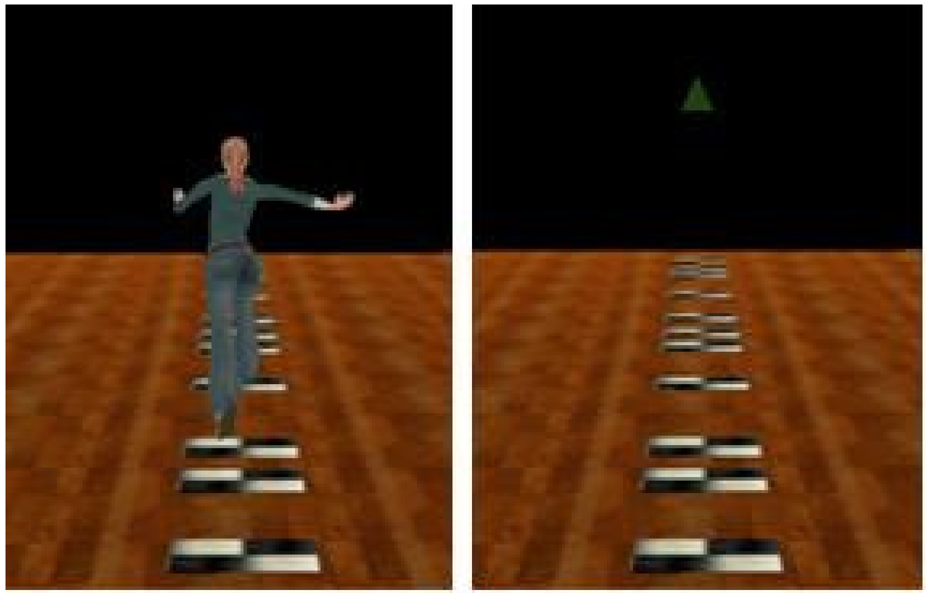

Figure 2.4: Learning phase animations for the stepping stone task: The animated character avatar jumping from the third to the fourth block (left). The green cone moving in the xz-plane of the invisible character (right).

given the capability to see their first person avatar and had no representation of their own body within the environments, but were still encouraged to visually explore the environment for as long as they liked before and during task performance (for all but the distance estimation task).

Each participant was required to complete all three tasks. Each participant performed nine trials of the distance estimation task and four trials of both the object interaction task and locomotion task. The experiment was designed to last between 45 minutes and an hour, which is considered to be a reasonable amount of time to expect someone to stay immersed in an HMD VE. During the entire task performance phase, the trajectories of the person's head and pelvis as well as their full body motion (from the Xsens MVN motion capture suits) were recorded for additional analysis

Figure 2.5: Virtual room used for the distance estimation task: The environment for the distance estimation task was a 10.27m length x 7.25m width x 2.77m height office room modeled off of an office located in the building in which the experiment was conducted. Experiment 1b was conducted in the physical room of which this model was a virtual replica.

(i.e., accuracy analysis).

The three tasks were chosen so as to require the user to perform a range of actions that one could reasonably expect user of a HMD VE applications to perform. Distance estimation is a fundamental task in VEs that involves space perception and in this case involved an action-based response. The object interaction task tests a user's dexterous manipulation of objects in the environment and the stepping stone task requires goal-oriented locomotion, as opposed to free exploration of a space.

The first task that each participant completed was an egocentric distance estimation task (direct blind walking) taken from Mohler et al. [2010b]. This task was performed first to ensure that no additional feedback was given for the distance estimates since it has been shown that interaction with an immersive VE corrects the typical underestimation of distance estimates observed in HMDs [Mohler et al. 2006;

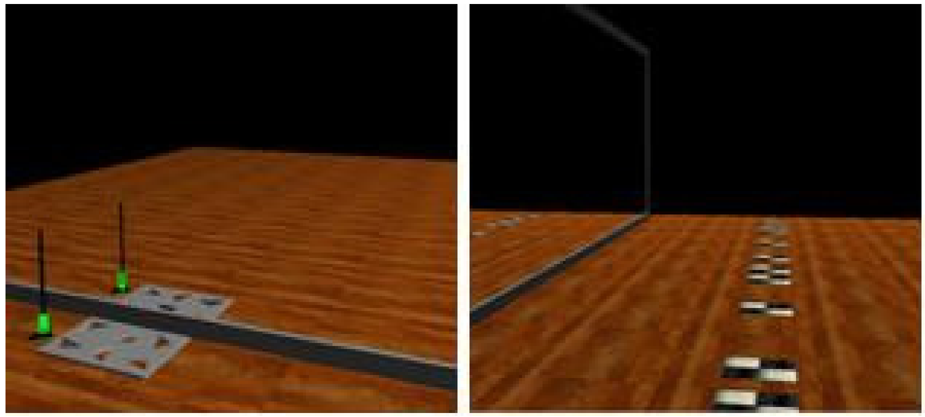

Figure 2.6: The room used for the object interaction task (left) and the locomotion task (right) was a 12 x 12 meter wooden floor with a full length mirror. The objects available to the participant in the environment for each task varied; a stamp tool and triangle board were put in front of the full-length mirror in the object interaction task, and stepping stones were placed in a line parallel and to the right of the mirror.

Richardson and Waller 2007]. The second and third task order was counterbalanced across all participants in each condition.

For the distance estimation task, the participants were asked to look at a target located on the floor placed at 3, 5, or 7 meters from the participant, with the distance randomly assigned in three blocks over the nine trials. The target was viewed from a fixed starting point as indicated by a yellow and black striped line which can be seen, along with the green target, in Figure 2.5. When the participant indicated that they were ready, the experimenter blanked the screen and the participant was asked to walk to where they believed the target was located without vision. For this task, the self-avatar was only visible before the task was performed (since the task was to walk without vision to the previously seen target). Participants were allowed to look down and at the mirror as long as they liked before each trial. The primary measure was the percent distance walked.

The object interaction task required the participants to take a triangle that was fixed to a 70cm long handle and insert it into six triangular cutouts, all with differing orientations, in a board that was lying on the ground (dimensions: 100 x 70 x 20cm) (see Figure 2.7). The completion of the task was insured by the experimenter who observed the motion of the stamping tool being fully inserted into the cutouts and recorded the completion time of the overall task. For the object interaction task, the participants could view the self-avatar by looking in the mirror or by looking down for as long as they liked.

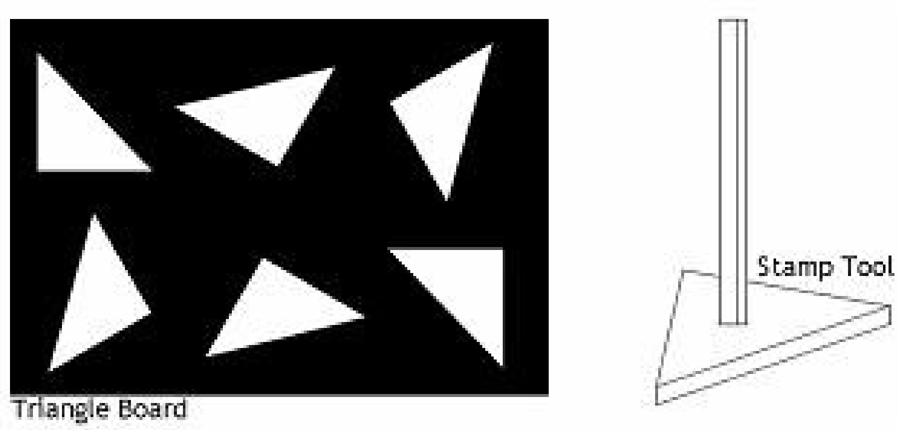

Figure 2.7: Schematic of the triangle board (left) and stamp tool (right) used during the object interaction task.

Additionally, it is important to note that the hands could not be perfectly calibrated for object interaction. The Xsens MVN motion capture suit provides only relative joint angle information and does not capture the motion of the fingers, thus with the MVN motion capture suits alone there is no exact measurement of the position of the hand in absolute world coordinates. Additionally, we do not capture motion of the fingers. To alleviate this situation, the hand was snapped to the end of the stamp tool when it was calculated to be within a certain vicinity of the tool. This process introduced small errors in limb geometry but increased the naturalness of the visual experience of the stamp tool and the visualization of the hand used for interaction.

In the stepping stone locomotion task, participants were asked to stand on one black and white, checkered square, measuring 42 x 59 cm (about twice the size of a standard piece of paper), and step/jump from one to another until they reached the last square. There were 10 squares that were placed at varying distances (0.47 to 1.9 meters, mean = 1.19, SD= 0.48) from each other in a straight line across the entire wooden floor environment (see Figure 2.6). The distance between the center of the first square and the last square was 8.8 meters. The participants were asked to complete the task as quickly and as accurately as possible, making sure that they believed they were stepping on the black and white squares rather than arbitrarily running from one side of the space to another. The completion time of the task was recorded and the footsteps were extracted from data taken from the Vicon pelvis corrected Xsens MVN motion capture suit data information about the foot position.

2.3.2 Experiment 1b: Real World Distance Estimation Task

2.3.2.1 Participants

Sixteen participants, 8 male and 8 female, participated in Experiment 1b. Their age ranged from 25 to 34 (mean= 28.57, SD = 3.20). All participants were volunteers from the local university community, had normal or corrected-to-normal stereo vision and received standard monetary compensation (8 €/hour).

2.3.2.2 Experimental Setup

Participants viewed targets at each of nine distances ranging from 2.0-6.0 meters (0.5 meter intervals) three times in a blocked order. This resulted in a total of 27 trials. They viewed each target in the real world and then were asked to lower a blindfold and walk without vision to the previously seen target. As they were walking without vision the previously seen target was pulled to the side by a string (invisible fish line) by the experimenter. The measurement of distance walked for each trial was recorded and then the participant was subsequently guided back indirectly to the starting location. No feedback about accuracy was provided to the participants.

2.3.3 Experiment 1c: Real World Object Interaction Task

2.3.3.1 Participants

Ten participants, 5 males and 5 females, participated in Experiment 1c. Their age ranged from 19 to 30 (mean = 24.07, SD = 3.06). All participants were volunteers from the local university community, had normal or corrected-to-normal stereo vision and received standard monetary compensation (8 €/hour).

2.3.3.2 Experimental Setup

The object interaction task used in this real world control experiment was identical to the task used in Experiment 1a. Participants performed four trials of the object interaction task in the real-world, using the same stamp tool and board which were used in the virtual object interaction task. They were given the same instructions to perform the task as quickly as possible.

2.3.4 Experiment 1d: Real World Stepping Stone Task

2.3.4.1 Participants

Eight participants, 3 male and 5 female, participated in Experiment 1d. Their age ranged from 23 to 47 (mean = 30.3, SD = 7.9). All participants were volunteers from the local university community, had normal or corrected-to-normal stereo vision and received standard monetary compensation (8 €/hour).

2.3.4.2 Experimental Setup

Participants were pre-screened for a height which would fit into our medium or large Xsens MVN motion capture suit (see Figure 2.1). Participants' height ranged from 1.62 to 1.84 meters (mean= 1.72, SD =0.07). Participants' height, arm span and foot length were also measured for the Xsens MVN motion capture suit software to more accurately estimate the position and orientation of the body and, in the case of height and arm span, for appropriate scaling of the self-avatar for use when calculating the accuracy of the foot on the stepping stones.

Participants wore the Xsens Moven motion capture suit and the pelvis tracker object. Unlike Experiment 1a and Experiment 1c, the participants wore nothing on their head (no HMD and no limiting field of view goggles) and therefore in the real world study their field of view was not limited. For this task, we wanted to be able to make a clear comparison between everyday real-world locomotion and locomotion in a virtual environment. They completed the stepping stone task four times and were not given any practice trials. Between trials they walked back to the start of the stepping stone task with vision, but did not walk on the stepping stones.

2.4 Results

The analysis of Experiment 1a includes 31 subjects. For the distance estimation task, the percent distance walked for each participant's distance judgment was calculated; for the object interaction task, the speed at which the task was completed was measured; while for the stepping stone locomotion task, the speed and accuracy (cm from the center of the stones) at which the task was completed was measured. Also, in the learning phase of all three tasks, the amount of time spent viewing the animation was recorded. During the task performance phase, we calculated how long the participants looked down at the location of their body and/or saw their avatar before and during task performance.

The analysis of Experiments 1b, 1c, and 1d were done in a similar manner to Experiment 1a, with the exception that the real world control experiments were set up as individual experiments and so participants did not perform each task. Experiment 1b (real world egocentric distance estimates) had 16 participants. Experiment 1c (real world object interaction task) had 10 participants. Finally, Experiment 1d (real world stepping stone task) had 8 participants. The task completion time and/or the accuracy of the task performance were measured as a control comparison for Experiment 1a.

2.4.1 Summary of Descriptive Results

The times for exploration during the learning phase are out of a 66 second total time possible, since this time is how long the animations played during this phase (see Table II.1). During the learning phase, participants viewed the animation for the majority of the time regardless of condition. During the distance estimation task, participants viewed the distance animation an average of 47.81 seconds (SD= 12.99). During the object interaction (OI) task they looked at the animation an average of 58.06 seconds (SD= 4.13). Finally, for the stepping stone (SS) task they viewed the animation an average of 49.78 seconds (SD= 7.74).

During the task performance phase, we calculated how long the participants

Table II.1: Head-gaze data during the learning phase: The mean time per condition that participants spent looking at the animation during the learning phase with one standard error in parentheses.

Condition	Distance Estimation	OI Task	SS Task
Cone/no-self	48.35s (2.26)	56.96s (0.79)	46.25s (3.23)
Character/no-self	39.41s (5.02)	58.96s (1.89)	49.40s (2.48)
Cone/self	45.54s (5.64)	57.69s (1.28)	47.01s (2.17)
Character/self	59.40s (0.90)	58.7s (1.98)	57.41s (0.84)

viewed the self-avatar, or the location where the self-avatar would have been in the no self-avatar condition (see Table II.2). There was no limit put on the time that the participants could look at themselves before performing the tasks. Before the distance estimation task, participants viewed their self-avatar (physical body location) an average of 32.45 seconds. For the object interaction task, they looked at their self-avatar (location) both before and during task execution an average of 62.32 seconds. Finally, for the stepping stone task participants viewed their self-avatar (location) both before and during task execution an average of 40.27 seconds.

Table II.2: Head-gaze data during the task performance phase: By condition, the mean time subjects looked at the location of the self-avatar during the task performance phase of the experiment with one standard error in parentheses.

Condition	Distance Estimation	OI Task	SS Task
Cone/no-self	24.25s (2.79)	63.38s (5.61)	19.03s (1.48)
Character/no-self	21.38s (1.5)	65.13s (3.72)	29.98s (3.46)
Cone/self	41.88s (6.03)	58.25s (1.69)	47.86s (5.97)
Character/self	43.71s (5.74)	62.57s (4.68)	67.63s (6.88)

In Experiment 1a (HMD), participants estimated distance by walking without vision 77.89% (SD= 13.44) of the actual distance to the previously seen targets. Participants performed the object interaction task in an average of 17.08 (SD= 2.54) seconds. Participants performed the stepping stone task in an average of 5.60 (SD= 1.03) seconds. In Experiment 1b (real world), participants estimated distance by walking without vision 90.11% (SD=17.96) of the actual distance to the previously seen targets. In Experiment 1c real world), participants performed the object interaction task in an average of 13.05 (SD= 0.81) seconds. Finally, in Experiment 1d (real world), participants performed the stepping stone task in an average of 2.97 (SD = 0.45) seconds (see Table II.3).

Table II.3: The HMD and the real world experimental results listed for the average distance walked for the distance estimation task (expressed as a percentage of the true egocentric distance), as well as the average times to completion for the object interaction and stepping stone tasks with one standard error in parentheses.

Condition	Distance Estimation	OI Task	SS Task
Cone/no-self	75.14% (7.74)	19.12s (0.78)	6.65s (0.42)
Character/no-self	76.88% (4.08)	16.25s (0.77)	5.10s (0.19)
Cone/self	80.09% (2.59)	15.51s (0.53)	5.45s (0.29)
Character/self	76.66% (3.86)	17.50s (1.10)	5.15s (0.25)
Experiments 1b,1c & 1d, Real World Controls	90.12% (7.33)	13.04s (0.25)	2.97s (0.16)

2.4.2 Statistical Analysis

For the statistical analysis we focus on each task separately. In Experiment 1a, we analyze both task performance and learning with respect to the two factors, self-avatar and character animation. We additionally do a statistical comparison between the task performance in the HMD (Experiment 1a) and the real world (Experiments 1b, 1c, and 1d).

2.4.2.1 Experiment 1a: HMD Egocentric Distance Estimation Results

A pre-recorded character animation and a self-avatar had no impact on egocentric distance estimates as measured by direct blind walking. Regardless of condition, participants significantly underestimate egocentric distances in the direct blind walking task (Mean= 77.89, SD = 13.44), as is indicated by a one-sample t-test; t(30) = -9.16, $p < 0.001$ (see Figure 2.8). A repeated measures ANOVA confirmed that for the distance estimation task neither of the between subject factors, character animation and self-avatar, were significant for percent distance walked, $F(1,26) = 0.291$, $p = 0.594$ and $F(1,26) = 0.112$, $p = 0.740$, respectively. The interaction between character animation and self-avatar was also non-significant, $F(1,26) = 0.130$, $p = 0.721$.

To assess learning, we investigated the effect of time (i.e., different trials) on percent distance walked in a mixed ANOVA with trials as a within subject factor and animation and self-avatar as between subject factors. We found a significant effect of trial on the percent distance walked $F(8,208) = 2.249$, $p = 0.025$ ($\eta p^2 = 0.08$). This effect of trial on completion time was not modulated by animation, avatar, or a particular combination of avatar and animation as indicated by non-significant

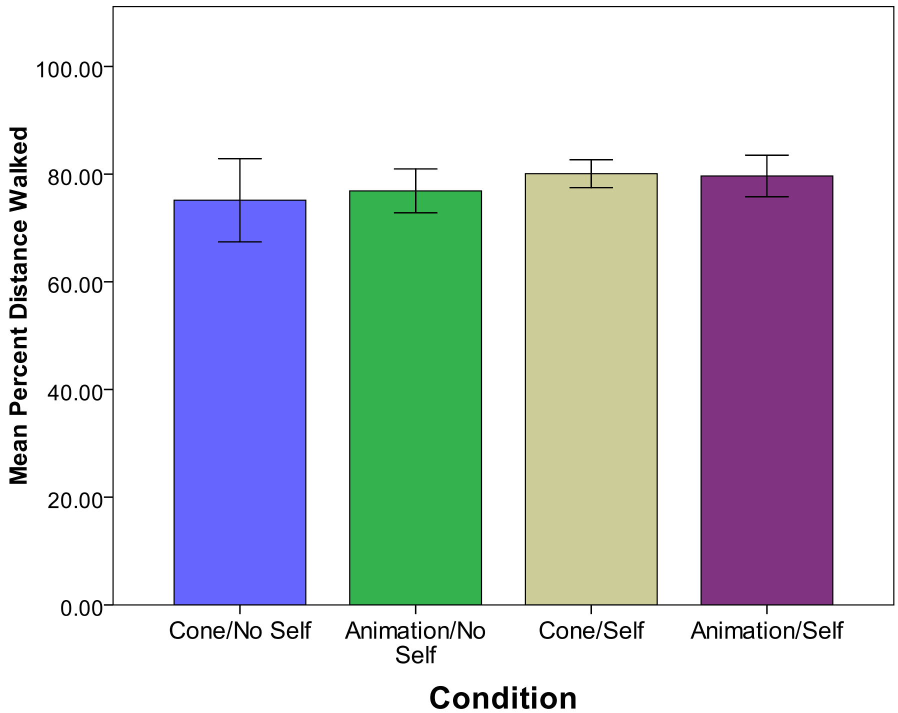

Figure 2.8: The average percent distance walked in the distance estimation task in the HMD shown for each of the four conditions separately as well as the real world. 100% would indicate veridical walking to the true egocentric distance. Error bars indicate +/- 1 standard error.

interactions between trial and animation $F(1{,}26) = 0.291$, $p = 0.594$, trial and self-avatar $F(1{,}26) = 0.112$, $p = 0.740$, and trial, self-avatar, and animation, $F(1{,}26) = 0.130$, $p = 0.721$. Specifically, the percent distance walked increased over time 74.07 to 81.36%. This finding suggests that the learning effect was statistically similar in all conditions.

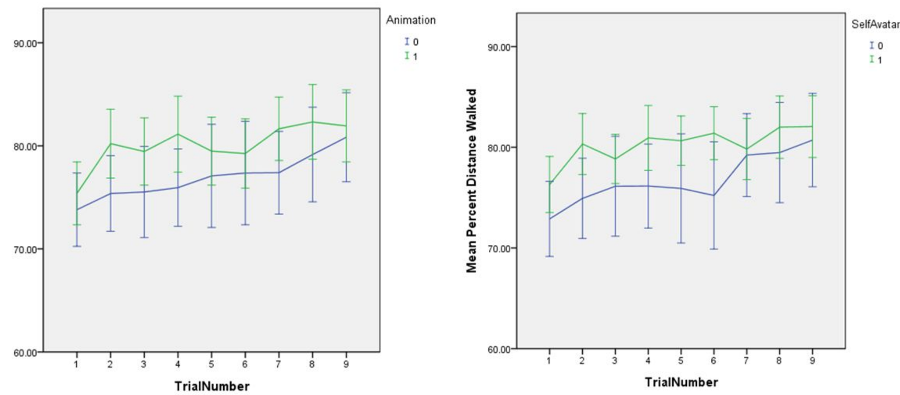

Figure 2.9: The average percent distance walked in the distance estimation task in the HMD shown over the nine trials. The left graph shows the factor of character animation during the learning phase, while the right graph shows the factor of self-avatar during the task performance phase. Error bars indicate +/- 1 standard error.

2.4.3 Experiment 1a: Object Interaction Results

The time to complete the object interaction task is shown for each of the four experimental conditions separately in Figure 2.10. In a repeated measures ANOVA, we find a significant main effect of self-avatar, $F(1,26) = 5.284$, $p = 0.03$, but no significant impact of animation on object interaction task performance, $F(1,26) = 0.1235$, $p = 0.277$. However, the interaction of self-avatar and animation was significant, $F(1,26) = 7.532$, $p = 0.011$. The interaction indicates that the significant effect in this task was that the self-avatar was beneficial when the animation was not used in the learning phase; otherwise the self-avatar had no impact on task performance.

To assess learning we investigated the effect of time (i.e., performance over the four trials) on object interaction completion time in a mixed ANOVA with trials as a within subject factor and animation and self-avatar as between subject factors. We found a significant effect of trial on the object interaction completion time $F(3,78)$

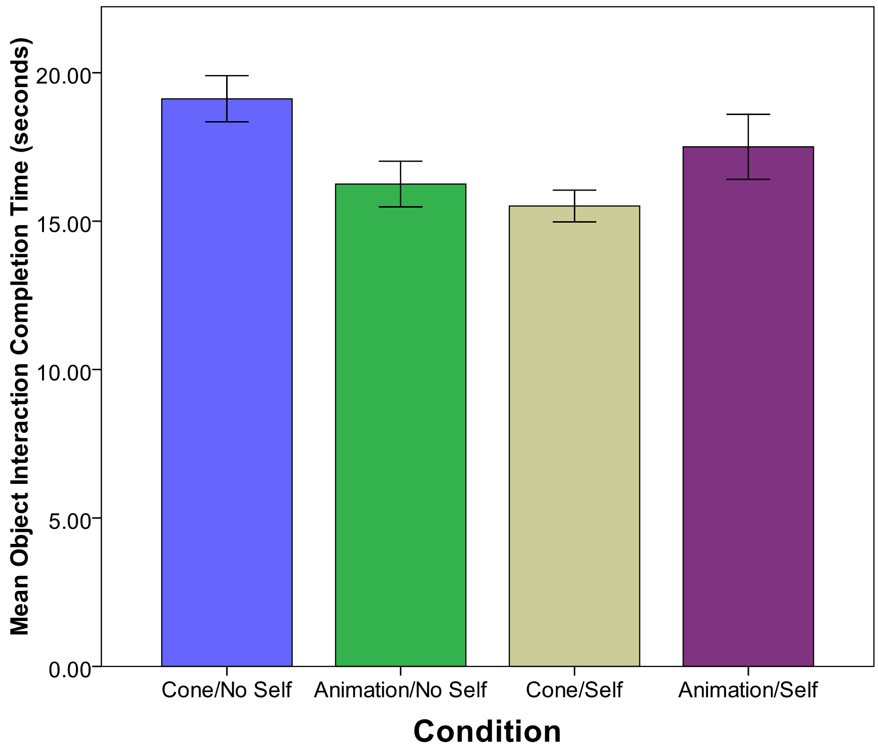

Figure 2.10: The mean completion time in seconds of the object interaction task in the HMD shown for each of the four conditions separately. Error bars indicate +/- 1 standard error.

$= 53.298$, $p < 0.01$ ($\eta p^2 = 0.672$). This effect of trial on completion time was not modulated by animation, avatar, or a particular combination of avatar and animation as indicated by non-significant interactions between trial and animation $F(3{,}78) = 1.905$, $p = 0.136$, trial and self-avatar $F(3{,}78) = 2.203$, $p = 0.094$, and trial, self-avatar, and animation $F(3{,}78) = 0.684$, $p = 0.565$. Specifically, the completion time decreased over time from 20.065 to 14.909 seconds (see Figure 2.11).

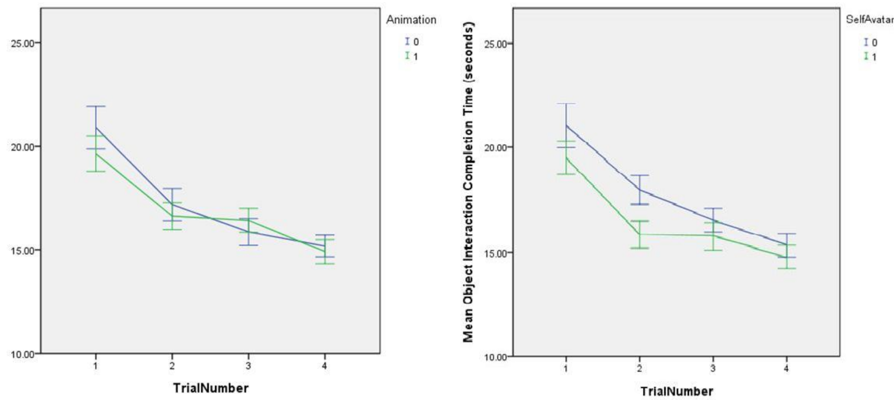

Figure 2.11: The mean completion time in seconds of the object interaction task in the HMD shown over the four trials. The left graph shows the factor of character animation during the learning phase, while the right graph shows the factor of self-avatar during the task performance phase. Error bars indicate +/- 1 standard error.

2.4.4 Experiment 1a: Stepping Stone Results

The completion time results of the stepping stone locomotion task are shown in Figure 2.12. Both the animation and the self-avatar have an impact on stepping stone completion time. However, this effect does not appear to be additive. The repeated measures ANOVA shows a significant main effect of self-avatar on time to complete the stepping stone task, $F(1,24) = 4.262$, p =0.05, and a significant effect of animation, $F(1,24) = 5.754$, $p = 0.025$. The interaction of self-avatar and animation was not significant, $F(1,24)$ =2.414, $p = 0.133$. For the stepping stone task, it appears that any human character animation (character or self) was sufficient for significantly improving task performance time.

To assess learning we investigated the effect of time (i.e., performance over the four trials) on stepping stone completion time in a mixed ANOVA with trials as a within

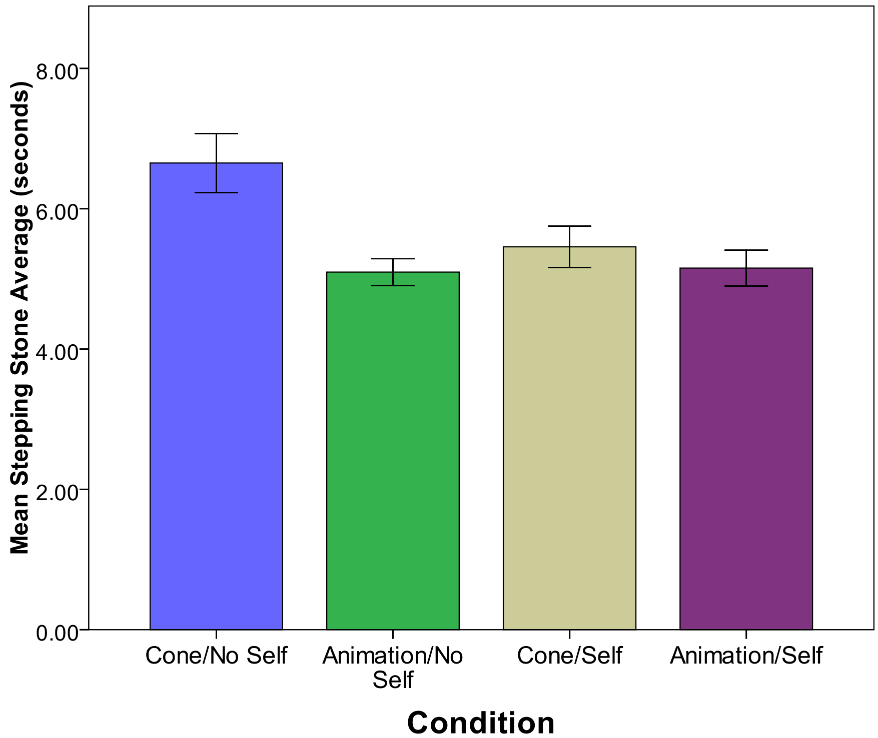

Figure 2.12: The mean completion time in seconds in the stepping stone task in the HMD shown for each of the four experimental conditions separately. Error bars indicate +/- 1 standard error.

subject factor and animation and self-avatar as between subject factors. We found a significant effect of trial on the stepping stone completion time $F(3{,}72) = 10.522$, $p < 0.001$ ($\eta p^2 = 0.305$). This effect of trial on completion time was not modulated by animation, avatar, or a particular combination of avatar and animation as indicated by non-significant interactions between trial and animation $F(3{,}72) = 2.064$, $p = 0.113$, trial and self-avatar $F(3{,}72) = 0.330$, $p = 0.804$, and trial, self-avatar, and animation $F(3{,}72) = 0.256$, $p = 0.857$. Specifically, the completion time decreased over time from 6.36 to 5.16 seconds (see Figure 2.13).

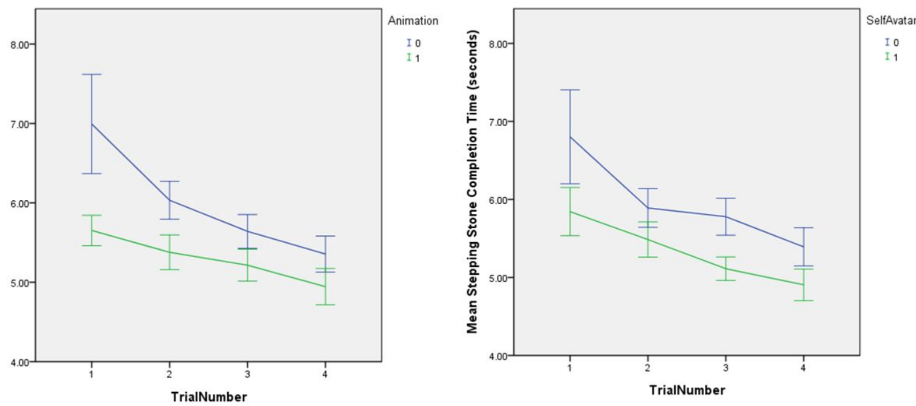

Figure 2.13: The mean completion time in seconds in the stepping stone task in the HMD shown over the four trials. The left graph shows the factor of character animation during the learning phase, while the right graph shows the factor of self-avatar during the task performance phase. Error bars indicate +/- 1 standard error.

Additionally, we analyzed how accurately participants performed the stepping stone task, as defined by the average distance (cm) that a foot came to landing on the center of each stone per trial (see Figure 2.14). A repeated measures ANOVA showed a significant effect of animation $F(1{,}24) = 18.932$, $p < 0.01$, $\eta p^2 = 0.441$ and self-avatar $F(1{,}24) = 129.049$, $p < 0.01$, $\eta p^2 = 0.843$ on task performance and a significant interaction between animation and self-avatar $F(1{,}24) = 560.528$, $p = 0.022$, $\eta p^2 = 0.2$. For the stepping stone task, it appears that both the character animation and the self-avatar significantly improved the accuracy of the participants. The interaction indicates that accuracy is improved more with the addition of a self-avatar when the learning phase did not contain an animation.

To assess learning we investigated the effect of time (i.e., performance over the four trials) on stepping stone accuracy in a mixed ANOVA with trials as a within subject

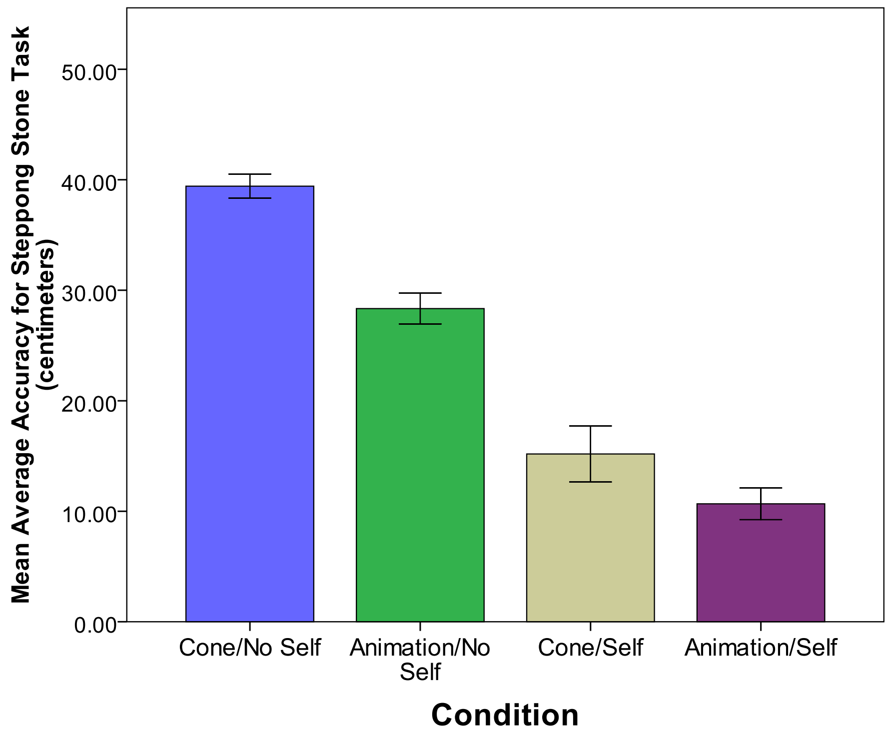

Figure 2.14: The mean accuracy (centimeters) of the stepping stone task in the HMD shown for each of the four experimental conditions separately. Error bars indicate +/- 1 standard error.

factor and animation and self-avatar as between subject factors. We found a significant effect of trial on the stepping stone accuracy $F(2.269,56.725) = 3.269$, $p < 0.05$, $\eta p^2 = 0.116$. This effect of trial on accuracy was not modulated by animation, avatar, or a particular combination of avatar and animation, as indicated by non-significant interactions between trial and animation $F(2.269,56.725) = 2.362$, $p = 0.078$, trial and self-avatar $F(2.269,56.725) = 0.399$, $p = 0.698$, and trial, self-avatar, and animation $F(2.269,56.725) = 0.221$, $p = 0.113$. Specifically, the accuracy decreased over time from 23.147 to 21.159 centimeters (see Figure 8). Mauchly's test indicated

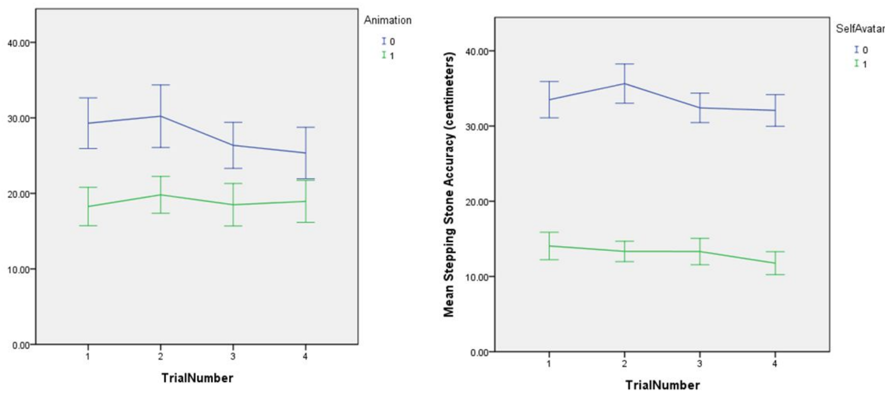

Figure 2.15: The mean accuracy in seconds of the stepping stone task in the HMD shown over the four trials. The left graph shows the factor of character animation during the learning phase, while the right graph shows the factor of self-avatar during the task performance phase. Error bars indicate +/- 1 standard error.

that the assumption of sphericity had been violated (chi-square = 14.735, $p < .05$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.756$).

2.4.4.1 Experiment 1b: Real World Egocentric Distance Estimation Results

Participants' average percent distance walked in the real world to previously seen egocentric targets was 90.12% of the actual target distance with a SD = 7.33 (see Figure 2.16). A repeated measures ANOVA with repetition as a factor and the dependent variable as the mean percent distance walked in the real world was performed. The results revealed no significant influence of repetition on mean percent distance walked, $F(1.347,20.202) = 0.985$, $p = 0.359$ ($\eta p^2 = 0.062$). Mauchly's test indicated that the assumption of sphericity had been violated (chi-square = 9.291, $p < .05$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.67$).

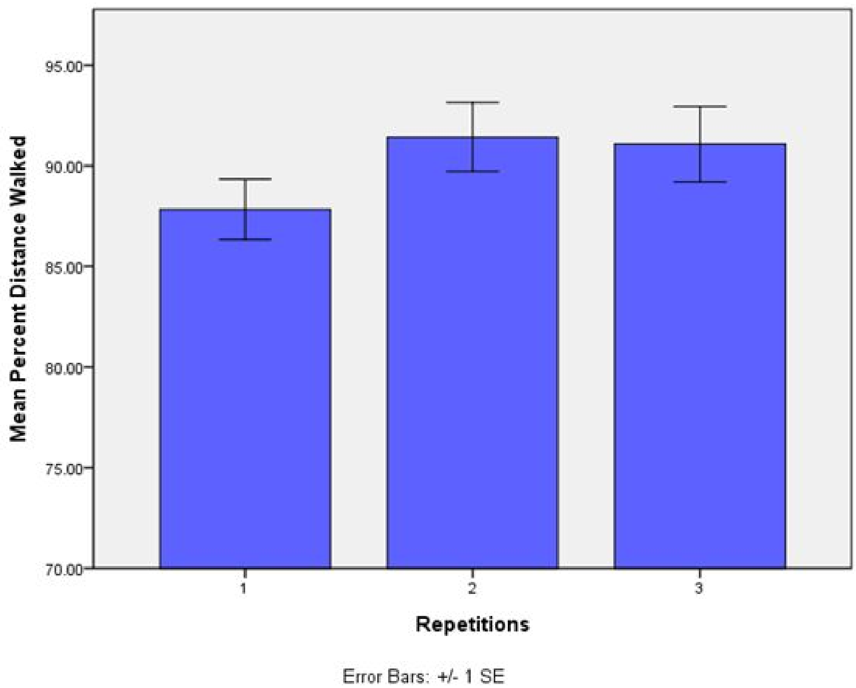

Figure 2.16: The mean percent distance walked to the nine different targets distances in the real world room shown over the three blocked repetitions. Error bars represent one standard error of the mean.

2.4.4.2 Experiment 1c: Real-world Object Interaction Results

Mean completion time for the object interaction task in the real world was 13.04 seconds with a SD = 0.81 (See Figure 2.17). The repeated measures ANOVA with trial as the factor and completion time as the dependent variable revealed a significant influence of trial on completion time in the object interaction task $F(3{,}27) = 6.8577$, $p < 0.001$, $\eta^2 = 0.432$. The mean in the first trial was 13.56 seconds with a SD =

1.034 and in the last trial was 12.49 seconds with a SD = 0.91.

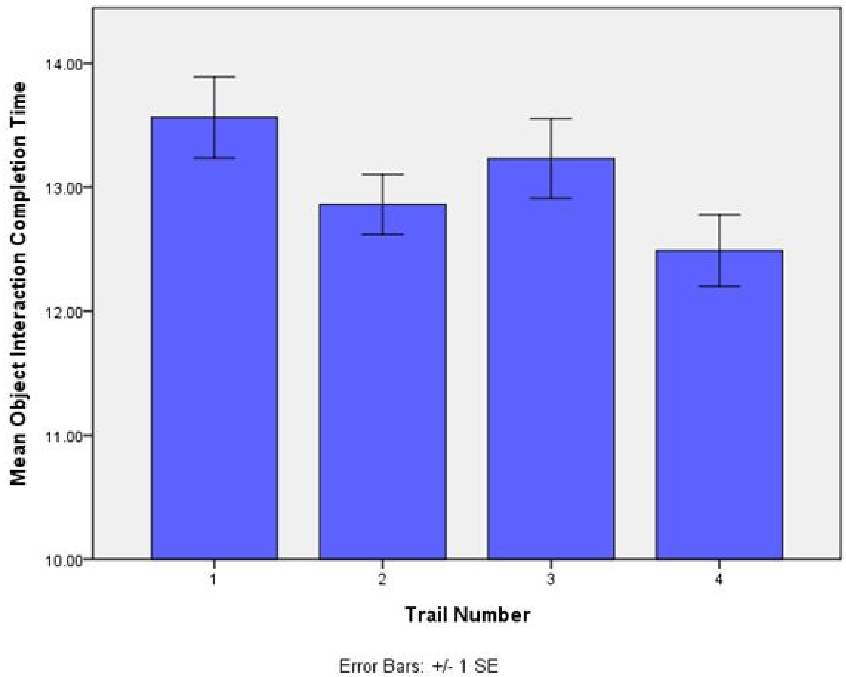

Figure 2.17: Mean completion time (in seconds) for the object interaction task in the real world across the four trials. Error bars indicate +/- 1 standard error.

2.4.5 Experiment 1d: Real-world Stepping Stone Results

Participants average time to complete the stepping stone task in the real world was 2.97 seconds with a SD = .45. Their average accuracy was 8.16 centimeters with a SD=2.81 (see Figure 11). A repeated measures ANOVA with trial as the factor and accuracy and completion time as the dependent variables revealed that there was no significant influence of trial on completion time $F(1.791,12.538) = 2.150$, p

$= 0.160$, $\eta^2 = 0.235$, however there was a significant influence on mean accuracy in the stepping stone task $F(3,21)$=6.636, $p < 0.05$, $\eta^2 = 0.487$. This suggests that people immediately perform the stepping stone task as quickly as possible in the real world and learn to perform more accurately over time (see Figure 2.18). For completion time, mauchlys test indicated that the assumption of sphericity had been violated (chi-square = 13.2, p ¡.05), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.60$).

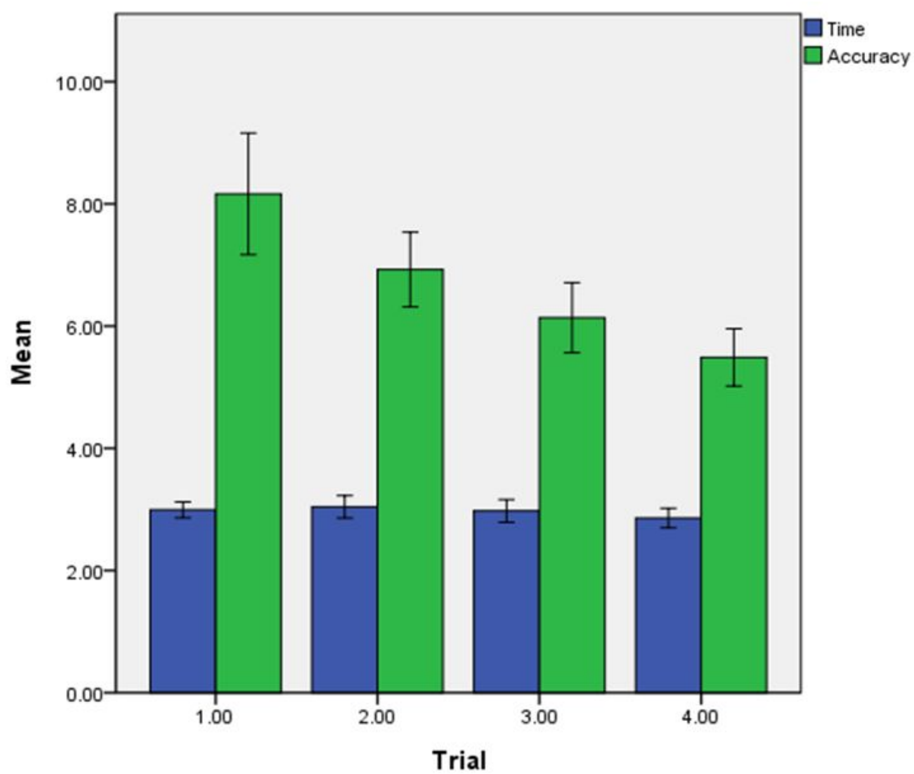

Figure 2.18: As indicated by the legend, the blue bars indicate the average completion time in seconds for the real-world stepping stone task across four trials. The green bars indicate the average accuracy, i.e., distance from the center of each stepping stone in centimeters, across the four trials. Error bars indicate +/- 1 standard error.

2.4.5.1 Statistical comparison between real and virtual world task performance

Independent sample t-tests were performed to determine if performance was different in the real and virtual worlds between Experiment 1 and Experiments 1b, 1c, and 1d. These t-tests revealed a significant difference between real (M = 90.12%, SD = 17.96) and virtual world (M = 77.88%, SD = 13.44) distance estimation task performance; t(23.93) = -2.399, p = 0.025, a significant difference between real (M = 6.68cm, SD = 1.62, M = 5.60s, SD = 1.03) and virtual world (M = 23.39cm, SD = 12.35, M=2.97s, SD=0.45) stepping stone accuracy; t(37)= 3.917, $p < 0.001$ and completion time; t(37) = 7.005, $p < 0.001$), and a significant difference between real (M =17.08s, SD = 2.54) and virtual world (M = 13.04, SD = 0.81) completion time for the object interaction task; t(39) = 4.916, $p < 0.001$).

2.4.6 Summary of Results

In Experiment 1a, we found no significant impact of a character animation or a self-avatar on distance estimates in an HMD VE. We saw a significant impact of a character animation and a self-avatar in a stepping stone task. For the stepping stone task we see an increase in accuracy and decrease in completion time for both the character animation and the self-avatar. For this task we also see an interaction between the animation and the self-avatar on performance accuracy, indicating that the self-avatar helps accuracy more when task learning was not conducted with a character animation. For the object interaction task, we also saw an interaction between the character animation and the self-avatar. This interaction indicates a finding consistent with the prior interaction, that is, that the self-avatar was beneficial when learning was not conducted with the character animation, but otherwise offered no further benefit to task performance. We note that learning occurs in HMD VEs regardless of whether an animation or a self-avatar has been provided.

In Experiments 1b, 1c, and 1d we investigated real world performance of task performance as a control for Experiment 1a. We recorded the baseline task completion time for the stepping stone and object interaction tasks with no obstructed field of view and little to no technological equipment encumbering action. We recorded the baseline performance in the replica real world room for direct blind walking percent distance walked, and further recorded the accuracy of the foot placement in the stepping stone task.

Finally, in a comparison between real and virtual world task performance we found a significant difference between real and virtual world task performance, where the real world participants performed the tasks more accurately and/or faster.

2.5 Discussion

In this chapter, we investigated whether seeing character animations of relevant tasks being performed and/or seeing a self-avatar during task performance help people to perform tasks better in an HMD VE. Additionally, we investigated the rate of learning in an HMD VE for these tasks. Our results show that the impact of character animations and self-avatars on task performance is task-dependent and that learning takes place for each task, but there is no influence of character animation or self-avatar on the rate of learning.

The question of whether providing a self-avatar improves distance estimates in a VE remains open; in contrast to previous research [Mohler et al. 2010b; Mohler et al. 2008; Ries et al. 2008], we found no influence of a self-avatar on a distance estimation task. This result is interesting because the reasons why an animated or static self-avatar may contribute to distance estimation is still not well understood, and the robustness of the effect is not known. Mohler et al. [2010b] suggest that either familiar size and/or perceptual-motor information about the impact of body movements over time may play a role. Ries et al. [2008; 2009] suggest increased presence in the virtual world as an additional possible cause.

Our method with regard to the distance estimation task differs slightly from prior work in some potentially significant ways. In our experiment, participants were allowed to explore their self-avatar for as long as they liked before each of the nine distance estimation trials. We provided participants with a mirror where they could view their virtual avatar as well as encouraged them to look towards their physical body to view the self-avatar. Providing this mirror meant that participants viewed the avatar also from the front and therefore saw their avatar's face, which was not true in the research performed by Mohler et al. [2010a] or Ries et al. [2009]. This methodological difference could have impacted the participants' feeling of embodiment in the VE and therefore decreased the effectiveness of the cues provided by the self-avatar that might lead to improved task performance. The freedom to explore their self-avatar provides another difference from prior work. Participants only looked down or in the mirror an average of 42 seconds when they had a self-avatar, and only an average of 23 seconds when they were provided no self-avatar. In Mohler et al.'s research experiment, participants were instructed to look down at their self-avatar or forward to their non-collocated avatar for five minutes [Mohler et al. 2010a]. In Ries et al.'s research experiment, participants only looked down at their avatar for a short time while stepping forward [Ries et al. 2008]. Ries et al. had participants perform a pre-test and post-test (within subject design), while we wanted to avoid any carry-over effects and conducted our experiment in a between-subject design, as in Mohler et al. [2010b].

Finally, we used a different animation technique than Mohler et al [2010b]. They used inverse kinematics (which ensured that the positions of the hands and feet were at the exact locations as the physical counterparts, but did not necessarily provide accurate angles at the elbows and knees). In the present research we used full-body animation from an inertial measurement suit (which instead ensured the opposite). Differences in calibration protocol (calibration of the self-avatar to the human user) could have resulted in the differences in distance estimation from prior work as well. Although these differences are small, it is an open question how calibration of the self-avatar to the human effects task performance, embodiment, and ownership of the self-avatar. Given that a person's calibration to their own body is based on a considerable amount of experience, small differences may be significant. These differences could likely be trained out through a process of recalibration [Welch and Sampanes 2008] but our training was likely not long enough to do that. The limit to which a self-avatar can be calibrated is likely another of the general constraints that HMD-based VEs place on achieving performance equal in the VE to performance in the real world, much like the constraints that limited field-of-view, weight, and inertia of the HMD place.

Consistent with recent research we found no influence of a pre-recorded character animation [Mohler et al. 2010a] on egocentric distance estimation judgments, even though our character animation was performing a locomotion task, unlike Mohler et al's previous research which showed a pre-recorded character animation performing a self-exploration task. This difference is consistent with Mohler et al's [2010a] argument that when seeing a self-animated avatar people are likely learning the coupling between the location of their physical body and the visual representation of their body in the virtual world.. This coupling would not occur with a character that is not self-animated (i.e. the pre-recorded animation).

For more complex tasks that involved the on-line visual control of actions, we found a significant impact of character animation or self-avatar. In both the stepping-stone and the object interaction tasks, the presence of a self-avatar was more beneficial on performance when learning was conducted without the character animation. Thus, the presence of an animated component, either in learning or during performance (in the form of an animated self-avatar), is helpful in task performance. In the object interaction task, this effect is the only one we find. In the stepping stone task, we saw significant influence on accuracy and completion time when participants saw the character animation in the learning phase and when they had a self-avatar as well. Overall, we believe participants performed well at both the object interaction task and the stepping stone task. Participants were able to perform the task nearly as quickly as the character animations (which were of the first author performing the task as quickly as possible in the real world). On average, participants performed the object interaction task in 17.08 seconds, compared to 15.3 seconds in the real world. For the stepping stone task, participants were traveling at an average speed of 1.57 m/s (8.8 meters/5.6 seconds) while jumping from one stone to the next. We were surprised that participants were comfortable performing this task in an HMD VE so quickly.

For this experiment we had participants perform the three tasks multiple times so that we could investigate the learning effect for these tasks in the HMD VE. We find that for all tasks, regardless of condition, participants improve at the tasks over time. In the case of distance estimates, this finding is consistent with some real world results, despite the fact that we do not provide any feedback [Kuhl et al. 2006]. It would be interesting to further evaluate their walking speed and stride length in order to determine if it has something to do with the participants walking more naturally over time (as is suggested by Philbeck et al. [2008]). In contrast to the virtual world learning results, we find no improvement in task performance over time in the real world. Participants are performing optimally in the real world; in the virtual world, however, various factors hinder their performance but repetition lends itself to learning and improvement.

Overall, we found that in comparison to real world task performance participants still performed the tasks more quickly and/or more accurately in the real world compared to the virtual world. However, we were encouraged by the performance in the virtual world and the rapid rate of learning of the participants in the virtual world given animations as feedback. Our results suggest that providing a character animation, a self-avatar and/or allowing participants to perform the task at least a few times could be used as a method of training and obtaining real world performance. Overall, these research results should provide encouragement to those interested in having similar performance in virtual and real worlds. This research suggests that adding animations of avatars performing tasks that are relevant to the environment and to the tasks which the user's of the VE must perform will yield more accurate task performance. For complex tasks that involve interaction with elements in the virtual world, a self-avatar will improve performance, but in most cases not more than a character animation. This news is encouraging since a self-avatar still requires special and expensive hardware to render accurately and character animation is much easier and more cost-effective to provide to the users of virtual reality. Finally, regardless of whether avatars (self or character) were given to the participants, learning occurred for all tasks over time. Therefore, if it is possible to provide the users of virtual reality with a training phase, task performance should improve rather rapidly in HMD VEs.

Chapter III

UNDERSTANDING PERCEPTION AND ACTION WHEN MAKING JUDGMENTS ON ERRORS IN TRAJECTORIES WHILE THROWING IN AN IMMERSIVE VIRTUAL ENVIRONMENT

3.1 Abstract

The trend in immersive virtual environments is to include the users in a more active role by having them interact with the environment and objects within the environment. Studying action and perception in virtual environments, then, becomes an increasingly interesting and important topic to study. We performed three experiments in an immersive virtual environment that investigated a user's ability to judge errors in self-produced and observed motion as well as the roles of visual and motor feedback while making these judgments. Specifically, we asked subjects to make judgments about the trajectories and endpoints of a ball thrown underhandedly. We chose throwing because it is a fairly common activity that many people have experience with and also so that we could model our first study on previous research [Reitsma and Pollard 2003; Reitsma et al. 2008]. All experiments were within subject and used a two-alternative forced-choice procedure. Experiment 2a had users judge the fidelity of the trajectory of their throws. We changed the trajectories of their throws with three perturbations by altering both the vertical and horizontal components of the ball's initial velocity and also by changing the magnitude of the gravity acting on the ball

over the course of its flight. Experiment 2b took the first step towards discriminating which cues subjects were using in order to make their judgments from Experiment 2a. In this second experiment, we hid the trajectory of the ball and showed only its landing point, along with a second ball representing the endpoint of a trajectory altered by one of the same perturbations from Experiment 2a; the user was asked which ball was theirs. We chose to investigate the role of the endpoint because the restricted field of view (FOV) of the head mounted display makes it difficult for users to view the complete trajectory of the ball, making the endpoint one of the more consistent cues available in the virtual environment. Experiment 2c was aimed at investigating the importance of visual and motor feedback and used a similar procedure to Experiment 2b. The user, however, played the role of both an actor and an observer. As an actor, the subject had both motor and visual feedback about their throws as they watched an animated self-avatar perform the task. As an observer, the subjects were deprived of the motor information associated with each throw as they watched another human throwing the ball. All three experiments determined the discrimination thresholds of the subjects. The results of Experiments 2a and 2b showed that there is no difference between a user's ability to perceive alterations in the vertical and horizontal velocities of a ball's motion, but there is a difference when detecting increases versus decreases in the magnitude of gravity acting on the ball. Through Experiments 2b and 2c, we conclude that when deprived of either visual or motor information, the remaining feedback is good enough to make a correct judgment; however, when given both visual and motor feedback there is no additive effect, or increase in performance.

3.2 Introduction

Users in immersive virtual environments (VEs) can often interact with their environments, but, other than walking, such interaction does not often invoke strong proprioceptive cues. A VE must provide some method of interaction, such as a tracking system for proprioceptive locomotion, e.g., Worldviz's PPT, an articulated data glove for gesture recognition, e.g., Cyberglove's Cyberglove, or a joystick. Such interaction methods have been studied extensively for VEs [Poupyrev et al. 1998]. Much prior work on perception in VEs has focused on perception, comprehension, and integration into memory of a virtual environment, and how orientation within this space is maintained during locomotion. The main goal of such short-term perception, comprehension, integration, and memory is navigation and wayfinding—mastering the layout of the VE itself and moving about within it.

In this chapter, we examine a different kind of distance and orientation dependent task, throwing a ball. Similar to much of the prior research, our goal is to understand users' perception of their actions in VEs and relate them to behavior in the real world. The work reported in this chapter represents a first step towards understanding the perception of throwing in VEs. The study here is divided into three experiments, investigating users' perception of their own actions as well as the actions of others and the roles that visual and motor information play when observing these actions. In each of the three experiments, users either performed or observed an action, namely throwing, and were asked to make a judgement to discriminate between truth and various alterations added to either the trajectory or endpoint of the thrown ball. We chose throwing because it is an action that acts over a distance and yields immediate feedback to the user. Additionally, we were motivated by a prior set of experiments performed by Reitsma et al. [2003; 2008] in which users were asked to watch trajectories of both a human character and a sphere in ballistic motion from a third person perspective and to make judgments on three types of errors in the trajectories. These perturbations affected the vertical and horizontal velocities of the ball during the flight phase of the trajectory or the gravitational constant governing the trajectory of the ball. Experiments 2a and 2b, instead of placing the subject into the role of passive observer, used the subject as the source of the motion; Experiment 2c had the participants play both roles.

The mirror-neuron system offers a direct matching between actions observed from a third person perspective and actions executed from a first person perspective [Rizzolatti and Craighero 2004]. This matching system suggests that perceptual effects of actions (such as the perceived flight path of a ball from a throw) are coded similarly during self execution and observation of the same action from profile. However, there are also some important differences between being the source of an action and being the passive observer of a similar action. In the case of self-produced actions, sensory effects (such as the proprioceptive cues and the visual feedback from the flight path of the ball) are canceled out based on specific sensory predictions of one's own action. Thus, it becomes an interesting question of whether motor or visual information is dominant when perceiving an action.

Separating motor and visual feedback in the real world is a difficult task requiring complicated experimental setups and it is much easier to think about performing this separation in the virtual world. The laws of physics do not apply in the virtual world and a person's viewpoint can be manipulated instantly. We took advantage of these aspects of virtual reality in this study so that we could better understand the roles of the visual and motor systems separately. We were able to alter both the path of the trajectory of a subject's throw as well as it's visibility and placed subjects at different viewpoints. All of these alterations allowed us to control which information was available to a subject at any point in time.

Section 3.3 introduces related work that has helped to motivate our current study. Section 3.4 explains the method used in order to accurately, and in real-time, simulate the trajectories of the throws of our subjects. Section 3.5 describes Experiment 2a which focuses mainly on the user's ability to detect one of three errors introduced into the flight path of the ball. Building upon Experiment 2a, Section 3.6 outlines Experiment 2b which investigated the role of the endpoint of the ball when detecting the errors used in Experiment 2a. The final experiment is described in Section 3.7 and examined the importance of visual and motor feedback during perception. Section 3.8

provides a discussion of our conclusions which can be made using the results of all three experiments.

3.3 Related Work

Most prior work on judgments of errors in ballistic motion placed the subjects in the role of an observer. For example, Reitsma et al. [2003; 2008] performed a series of experiments dealing with the effect of character type and preparatory motion on perception of errors in ballistic motion. They found that a short period of preparatory motion may be important when judging subsequent errors in the trajectory. Our study replaces this preparatory motion, observed from a third person perspective, with user-based action. While it has been suggested that humans have a poor grasp on ballistic motion and tend to be forgiving for large errors in ballistic trajectories of thrown balls [Hecht and Bertamini 2000], these results once again placed the users in a role in which they observed from the third-person. Placing the subject into an active role changes how we perceive the observed action. Gibson [1979] asserts that perception and action are inseparable and theories such as common-coding theory suggest that the same representations which code action effects also allow one to perceive the intended action goals [Hommel et al. 2001]. In other words, when performing an action, the same system is involved in both producing and perceiving the action. In accordance with the common-coding approach, altering action effects (such as perturbing the visual feedback associated with an action) can be an efficient way to study action-perception linkages. By creating a sensory discrepancy between action goals and their perceivable effects, one can break the functional linkage between action and perception, and can therefore examine how stable the sensori-motor coupling is with respect to the altered sensory effect.

Because throwing is an action that is fundamental to many sports, there exists a considerable literature examining the components of throwing accuracy, e.g., [Jardine and Martin 1983; Van Rossum and Bootsma 1989], and the mechanics of throwing, particularly underarm, as done in this chapter, are reasonably well understood [Bartlett 2008; Dupuy et al. 2000]. Throwing has been found to be a good indicator of distance perception in VEs [Sahm et al. 2005], and a simulated version of it, similar to what was done in this chapter, has been used in the context of a virtual handball game [Bideau et al. 2003].

3.4 Throwing in a Virtual Environment

The VE that was used in all three experiments consisted of a brick floor and 12 pillars, 6 equally spaced, 3m apart on each side of the participant, along with a red, square box approximately 4m away from the participant (see Figure 3.1). The texture on the floor, the pillars and the box were all used to add depth cues to the scene. The VE was presented by a full color stereo NVIS (Reston, VA) nVisor SX Head Mounted Display (HMD). The resolution of the HMD is 1280 x 1024 (SXGA) pixels per eye, a manufacturer's specified Field Of View (FOV) of 60° diagonally, and a

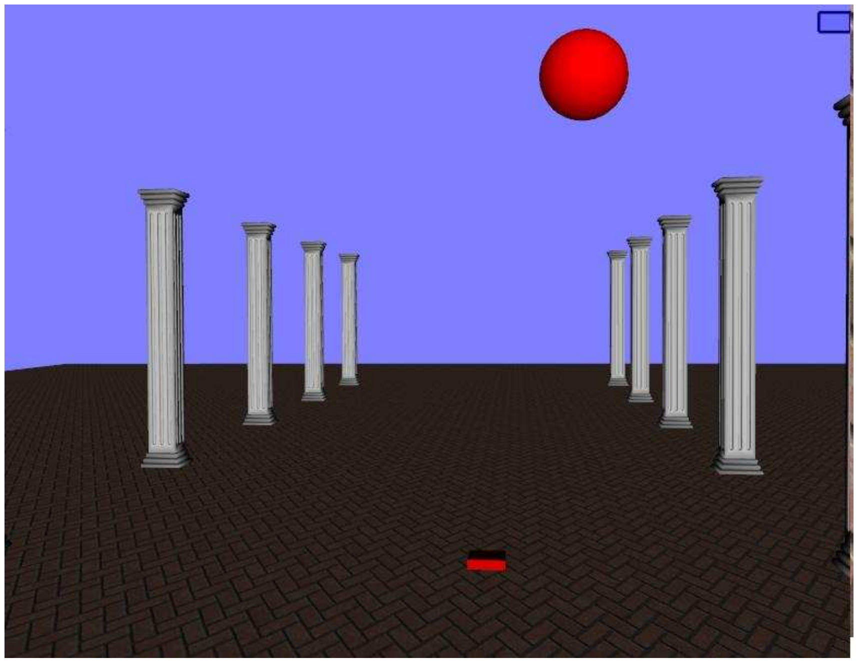

Figure 3.1: This figure shows the VE used in the throwing experiments. The red box on the ground is approximately 4m away, but was not used as part of the task in any of the experiments. The figure shows a ball in flight.

frame rate of 60Hz. An interSense IS-900 precision motion tracker is used to update the participant's rotational movements around all three axes. Position is updated using 4 optical tracking cameras working with 2 LED lights, and an update rate of 60Hz.

Two objects were tracked during each of the three experiments, a novelty tennis ball, 3 inches in diameter, and a wrist object (see Figure 3.2). During Experiment 2c there were 4 additional tracked objects which allowed for the rendering of a self-avatar. These objects were tracked using eight Vicon (Los Angeles, CA) MX-F40 cameras and Vicon Tracker software. The tennis ball had five markers attached to it in a specific configuration (see Figure 3.2) and the center of the ball was tracked by

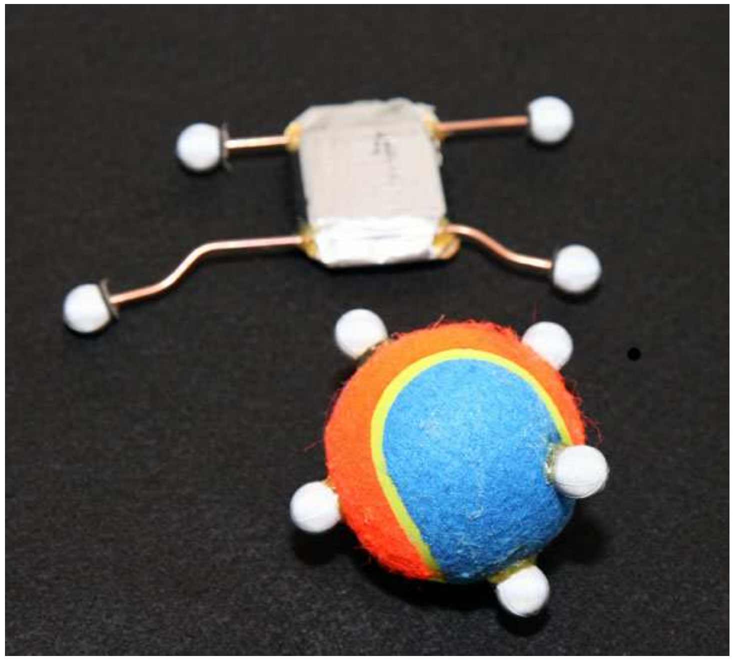

Figure 3.2: This figure shows the markered ball and wrist marker used in all three experiments.

averaging the position of four of the markers that were placed around the diameter of the ball. This calculation was used to position the virtual ball in the VE. In Experiment 2a we rendered the ball's shadow using an overhead light source, as did Reitsma et al. [2008]. The position of only the subject's right hand was tracked in the first two experiments and the hands, waist, feet and head were also tracked during Experiment 2c (see Figure 3.10). A collision detection algorithm, between the right hand's wrist makers and the ball markers, was used in order to detect the subject's release of the ball. When the subject was holding the ball, the algorithm fixed the position of the ball relative to the hand (see Figure 3.3). Once the ball was released, the position was sampled six times, once every other frame, given a frame

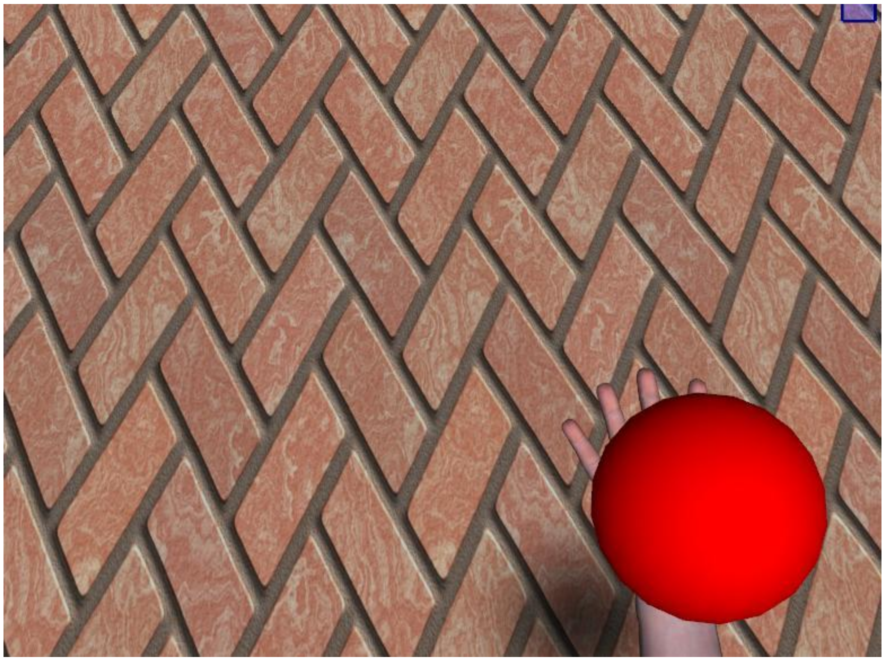

Figure 3.3: How subjects might see their hand if they looked down during Experiment 2a.

rate of 60Hz, and the velocity between each of these points was calculated. The initial velocity was chosen to be the median of these five values; any values of 0.0 m/s were omitted, as this suggested latency in the VE system. The sampling lasted for the first 0.17 seconds of the ball's trajectory, after which the rest of the trajectory could be calculated and simulated, either with or without alterations.

3.4.1 Alterations

Three types of alterations were added to the ball in Experiments 2a and 2b, in the positive and negative directions; both the horizontal and vertical velocity of the ball, along with gravitational constant were altered independently on different trials. Maximum and minimum noticeable magnitudes for the perturbations were determined using verbal reports based on the noticeability of a range of alterations used during a pilot study (see Table III.1). The velocity, v, of a projectile is described in Newtonian physics and separated into both a vertical and a horizontal component. These components are described by

$$v_{vert} = v_{0vert} - gt \tag{3.1}$$

$$v_{horiz} = v_{0horiz} \tag{3.2}$$

where g is standard gravity, t is time and v_0 is the initial velocity. Following the experimental parameters used by Reitsma et al. [2003], both vertical and horizontal velocity perturbations were smoothly added to the initial component velocities over a .25 second period at the beginning of the ball's trajectory, after the initial 0.17 seconds during which the ball's initial velocity was calculated. This was accomplished by using the sine function

$$v_{0altered} = v_0 + \Delta v_{\max} \sin\left((t - 0.16727)\left(\frac{\pi}{2}\right)\left(\frac{1}{0.25}\right)\right) \tag{3.3}$$

where $\Delta v_{\max}$ is the signed magnitude of the perturbation and v_0 is the initial vertical or horizontal velocity component used in Equations 3.1 and 3.2. After the .25 seconds of acceleration, the full magnitude of the alteration had been added to the initial velocity and this value remained constant for the rest of the trajectory. The gravitational constant was altered at the beginning of a perturbed trajectory and remained constant throughout the entirety of the ball's flight.

Table III.1: Maximum and minimum values across all three perturbations.

Direction	Vertical Velocity	Horizontal Velocity	Gravitational Constant
positive	0.1-2.0m/s	0.1-2.0m/s	0.1-4.0m/s^2
negative	0.1-2.0m/s	0.1-2.0m/s	0.1-2.5m/s^2

The effect of the perturbations on a sample trajectory can be seen in Figure 3.4. Horizontal velocity changes affect only the horizontal distance traveled by the ball; vertical velocity alterations, on the other hand, affect both the vertical and horizontal displacements of the ball. Adding or subtracting the same magnitude perturbation to the ball in both the vertical and horizontal directions, on independent throws, results in the same horizontal displacement from the starting position. Altering the gravitational constant, however, affects both the vertical and horizontal displacement differently than the other two perturbations. It should also be noted, while not evident in Figure 3.4, that altering either the vertical velocity of the ball or the gravitational constant will have an effect on the timing of the ball's trajectory, while altering the horizontal velocity only changes the horizontal displacement, and not the time it takes for the ball to hit the ground.

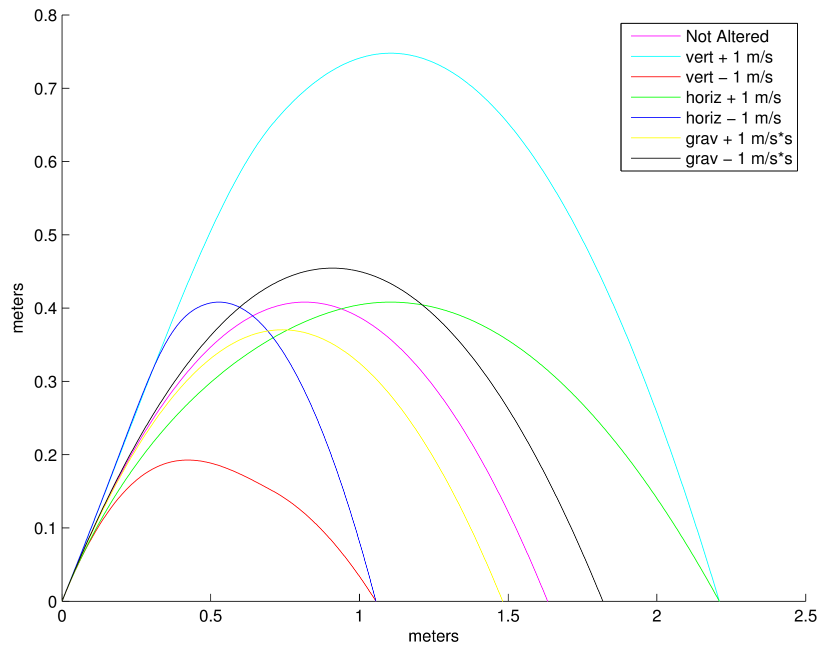

Figure 3.4: This figure shows how the trajectory of a throw is altered by perturbations of different types, but of the same magnitude.

3.5 Experiment 2a: Perceiving Alterations in a Trajectory while Throwing in a Virtual Environment

3.5.1 Participants

Twelve participants, 6 male and 6 female, participated in Experiment 2a. Their age ranged from 19 to 25 (mean = 20.83, SD = 1.53). All participants were student volunteers from the university and were compensated $10 per hour for their participation.

3.5.2 Method

Before the task phase of the experiment, each participant went through a training phase in which they threw 20 underhanded throws, none of which were altered. This training phase was used as a way to both familiarize the subjects with the system and provide a baseline for what an unaltered throw should look like. After the training phase, a two-alternative forced choice procedure was used in which each participant was asked to throw the tennis ball underhanded and make a judgment after every pair of throws as to whether the trajectory of the first or the second throw had been altered. Judgments were made for 180 pairs of throws; there were 30 trials for each of the three perturbations in both positive and negative directions, for a total of 360 throws per participant. Every participant took a 5 minute break after 90 trials were completed and were told that they could take additional breaks if necessary; only 2 participants chose to do so. The length of the experiment ranged from 50 to 90 minutes, depending on how long the subject took to throw the ball again once it had been returned. The experiment required two people, one to sit at the computer to reset the tracking of the ball and record the response for each trial, and another to catch the ball after each throw. The ball was caught so as to eliminate any sound cues produced by the ball hitting the floor, which may have conflicted with the collision of the virtual ball with the brick floor in the environment when altering the trajectory. The order in which the 6 perturbations were applied to the ball was randomized along with the presentation of whether or not the first or the second trajectory in each pair was altered. In order to determine a discrimination threshold for each perturbation, we used a maximum-likelihood stimulus procedure [Grassi and Soranzo 2009]. The values for the midpoint, slope and false alarm rate were all chosen in a manner similar to that of Wu et. al. [2009] in a pilot experiment. This procedure calculates the magnitude of the next perturbation by optimizing over candidate psychometric functions and eventually converges to a threshold value. An example of this process is shown in Figure 3.5.

3.5.3 Results

Threshold values were computed for each subject and for each condition by averaging the last four trials of the maximum likelihood method, e.g., Figure 3.5. Figure 3.6 shows the mean threshold values for all subjects with standard errors of the mean for perturbations in both the horizontal and vertical velocity components and Figure 3.6 shows the the same for gravitational perturbations. In Figure 3.6, vUP, hUP means there was an increase in the magnitude of the initial velocity in the respective direction, whereas vDOWN, hDOWN is a decrease. In Figure 3.6, gUP is an increase to the standard gravity term g (9.8 m/s^2), whereas gDOWN is a decrease. Note that the base threshold for horizontal and vertical perturbations are expressed as a velocity, in m/s, although they were phased in over a short time window and are therefore technically an acceleration.

Subject sensitivity scores for the vertical and horizontal velocity alterations were

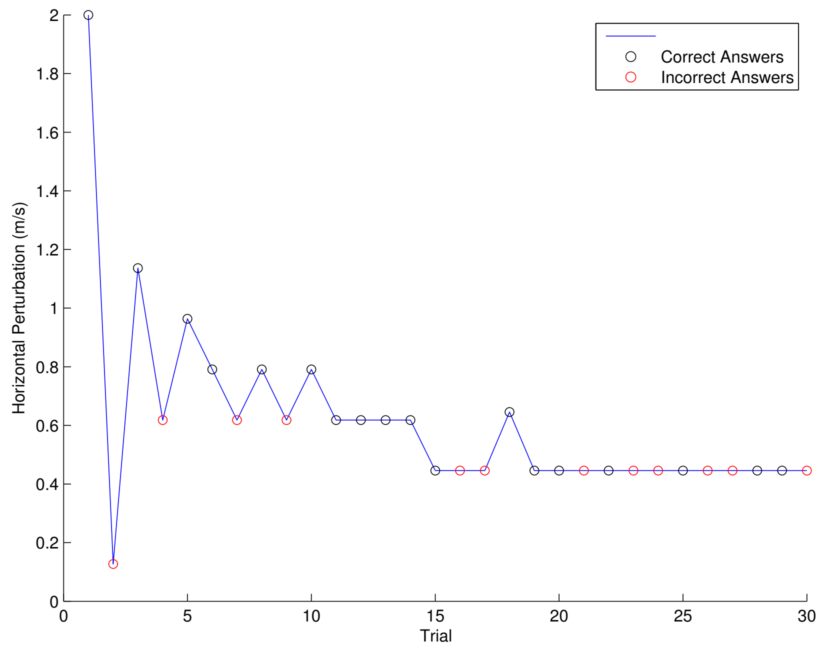

Figure 3.5: An example of a result. Blue circles indicate that the subject gave a correct response as to which trajectory was perturbed, red indicates an incorrect response. The final discrimination threshold for this subject was 0.45 m/s.

analyzed with 2×2 repeated measures analysis of variance (ANOVA) where the factors were error type (vertical velocity, horizontal velocity) and error direction (positive, negative). The ANOVA revealed no significant main effect of error type, F (1, 11) = 0.99, p = 0.34; or error direction F (1, 11) = 1.22, p = 0.3; and no interaction, F (1, 11) = 1.27, p = 0.1. As also can be seen in Figure 3.6, results indicated that participants did not differ in their sensitivity to the two different types of perturbations on velocity and to the two different directions of added errors. Subjects were equally good at detecting changes in the horizontal and vertical velocity, and in either direction. Results of the current analysis contrast with those of Reitsma et al. [2003; 2008], who found that errors in horizontal velocity are easier to detect than errors

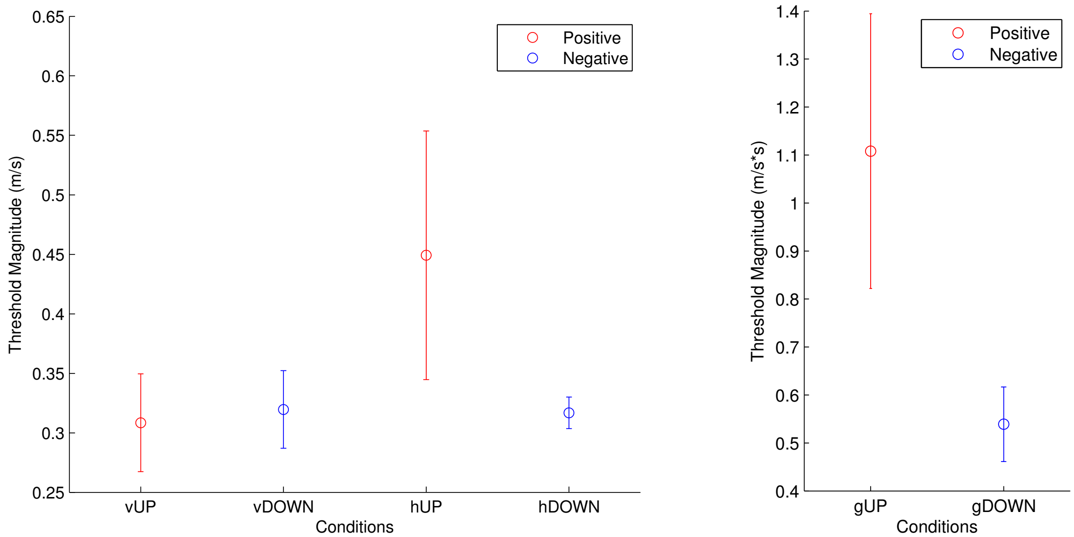

Figure 3.6: (left) Threshold magnitudes for the four velocity conditions. (right) Threshold magnitudes for the gravity conditions. The error bars show one standard error of the mean.

in vertical velocity, and that positive errors (accelerations) are easier to detect than negative errors (decelerations).

Subject sensitivity scores for the gravity alterations were analyzed with a paired-sample t-test to see if subjects differed in their sensitivity to errors in the positive vs. negative direction. Results revealed a marginally significant effect of direction, t (11) = 1.85, $p = 0.09$. Thus, it was slightly easier to detect decreases in gravity over increases in gravity; although this finding needs to be verified with future studies and with larger sample sizes. This result is in line with Reitsma et al.[2003; 2008], who found lower gravity changes easier to detect than higher gravity changes; however it is likely that our study, with a small sample size and increased within-subject variability in lower gravity changes (M = 1.10, SD = .99) compared with higher gravity changes

(M = .53, SD = .27), was not able to reliably detect the effect of error direction on gravity.

Additional analysis was done to see if there was an effect of gender on sensitivity scores. Consistent with Reitsma et al. [2003; 2008], the current study also resulted in no significant effect of gender ($p = .46$).

3.5.4 Discussion

This study determined the thresholds for perception of alterations in the trajectory of a ball thrown by the user of a VE. We found no significant differences between vertical and horizontal velocity perturbations, regardless of direction. Subjects threw the ball with an average initial velocity of 3.94 m/s (SD=0.64), and the mean detectable velocity threshold, collapsed over both directions and types, was 0.35 m/s (SD=0.21), so people can detect a velocity perturbation of roughly 9% using underarm throwing in the ranges we tested. For gravity perturbations, subjects could detect a downward acceleration (increase in gravity) of 1.1 m/s^2 (SD=0.99), 11% of a standard gravity, and an upward acceleration (decrease in gravity of 0.54 m/s^2 (SD=0.27), 5.5% of a standard gravity.

When judging whether or not a trajectory is altered, we may hypothesize about what cues subjects were using to make their decisions. Because the FOV of the HMD is restricted, the subjects did not see the full trajectory in a majority of the trials; this happened when the subjects were looking down at their hand or the target to begin the throw and missed the peak of the trajectory. It makes sense, then, that the judgment on the fidelity of the trajectory may have been based on the ending point of the ball, which led to our next experiment as described in the next section.

3.6 Experiment 2b: Eliminating Visual Feedback while Throwing in a Virtual Environment

3.6.1 Participants

Twelve participants, 7 male and 5 female, participated in Experiment 2b. Their age ranged from 18 to 57 (mean = 28.58, SD = 12.06). All participants were volunteers from the university and were compensated $10 per hour for their participation.

3.6.2 Method

There was a training phase in which each participant was asked to throw 20 unaltered underhand throws. The training phase was followed by the task phase which consisted of 180 trials of one throw per trial. Only once the ball hit the floor (in the VE that is, as simulated, since the ball was caught in the real world to eliminate any sound cues) was the subject able to see the ball again after its initial release. Once the ball landed, two balls (one green and one blue, with color randomized on each trial) were shown to the user. One was positioned at the true endpoint of their throw and one was positioned at a point representing the endpoint of a trajectory that had been altered by one of the six perturbations used in Experiment 2a. The users were then

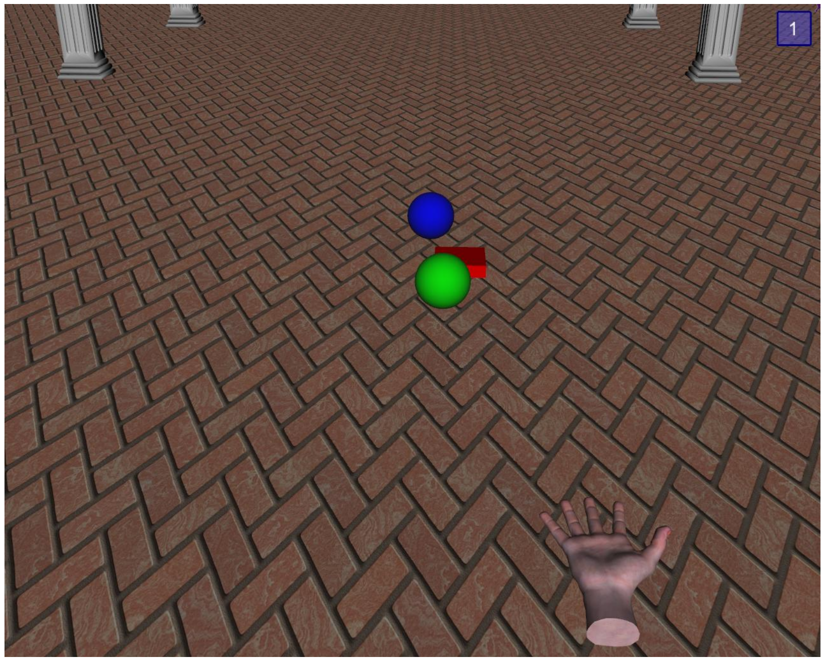

Figure 3.7: This figure shows what the subjects saw after throwing the ball. They had to choose which ball was the one that they had thrown.

asked to choose which ball they had thrown. The same maximum likelihood function was used as in Experiment 2a to determine the perturbations added to the altered ball on each trial and so that a final discrimination threshold could be determined for each perturbation for each user.

3.6.3 Results

A 2 (error type: vertical velocity, horizontal velocity) $\times$ 2 (error direction: positive, negative) repeated measures ANOVA showed no significance with regard to error type or error direction. There was also no interaction. A paired-sample t-test showed that subjects could more easily distinguish a decrease in the gravitational constant

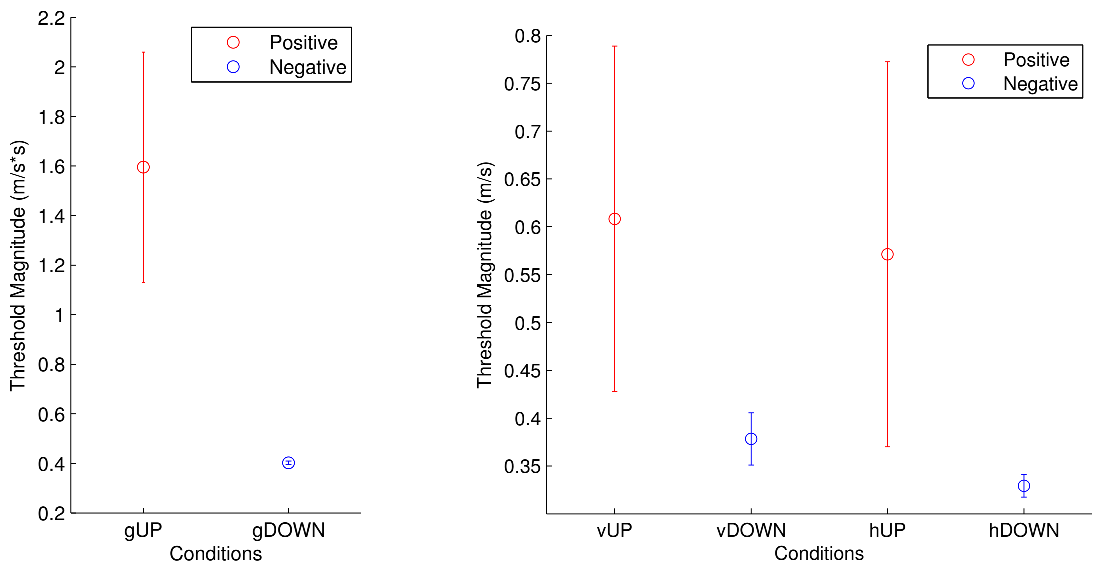

Figure 3.9: Threshold magnitudes for velocity perturbations.

over an increase ($p = .026$). Repeated measures ANOVAs were also used to measure the sensitivity thresholds for velocity/gravity with the trajectory information (full trajectory, just endpoint) as the between subjects factor. Again, there were no significant results in the case of velocity perturbations, but the effect on gravity error direction was significant ($p = .004$).

3.6.4 Discussion

Experiment 2a found only a marginal significance of gravitational error direction while the current study found a significant effect of error direction. The thresholds across all conditions are higher than those from Experiment 2a, but not significantly so. There are many factors that could play into a subject's ability to judge errors in their own throws. These include acceleration profiles and coordinates from the trajectory, proprioceptive cues from the act of holding and throwing the physical ball, timing of the trajectory from release to landing and the final resting point of the ball. Because the results outlined in this experiment are not completely conclusive, we cannot make a strong statement as to how important the endpoint is when judging alterations. While we believe that the endpoint is one of the discriminating factors, we do not believe it is the only factor or that it is the most important factor.

3.7 Experiment 2c: The Importance of Visual and Motor Information while Throwing in a Virtual Environment

3.7.1 Participants

Sixteen subjects, 8 male and 8 female, participated in Experiment 2c. Their age ranged from 18 to 48 (mean = 22.69, SD = 6.94). All participants were volunteers and were compensated $10 per hour for their participation.

3.7.2 Setup

An avatar was rendered using Autodesk MotionBuilder 2010 software and was set to appear slightly transparent so that the users could see the motion of the arm as the ball was thrown (see Figure 3.11). The avatar was scaled to each individual's height and gender matched. Participants were asked to wear 6 tracking objects with reflective markers, see Figure 3.10, so that their movements could be tracked and

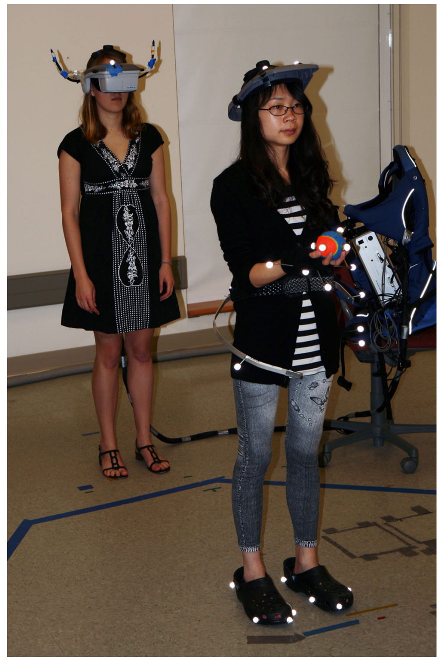

Figure 3.10: A participant acting as an observer in the experiment. The assistant is wearing the 6 tracking objects on her feet, hands, waist, and head so that a character can be created for the participant to view in the VE.

mapped to an avatar during the experiment.

3.7.3 Method

The subjects in this experiment were asked to play the role of both an actor and an observer. The participants were put into one of two conditions: playing the role of the actor followed by the observer or vice versa. Condition was counterbalanced across all participants. As the actor, the participant was placed 4m behind their self-avatar in the VE. They were asked to throw 20 underhanded throws in the environment as a training and adjustment period before the task portion of the experiment. During these 20 throws the ball was visible for the entirety of its trajectory. As an actor

Figure 3.11: Example viewpoint of a subject in Experiment 2c after throwing the ball. Male subjects saw the avatar shown here and female subjects were provided with a female avatar.

during the task phase, subjects were asked to watch themselves underhand throw a ball 40 times. After release, the ball disappeared and at the moment it hit the ground two balls appeared in front of the participant. One ball represented the placement of the ball that had just been thrown, and one ball was displaced in either the positive or negative x-direction. Just as in experiment 2b, one ball was green and one was blue, with color randomized on each trial. The placement of the second ball, however, was not determined using the six different alterations as in Experiments 2a and 2b. Instead, we decided to use the maximum likelihood function to determine a distance away from the true endpoint with which to place the altered ball. The maximum likelihood function gave us a value between 0.1 and 3.0m. The 6 different alterations were not used because the purpose of this experiment was not to judge a user's perception of alterations, but instead to further investigate the importance of visual and motor feedback. The subject was asked to respond with a 'farther' or 'closer' response as to which ball they believed had just been thrown. As the observer, the participant was placed 4m behind an avatar of another person who was in the lab (and gender matched to the participant) as they watched the assistant underhand throw the ball 40 times and make the same judgment. Also, there was a similar training phase in which the participant watched the assistant underhand throw 20 times while watching the entirety of the trajectory. Figure 3.11 shows what a subject would see after the two balls appeared.

3.7.4 Results

The mean threshold value for each participant in each role was computed by averaging the displacement of the ball during the last 6 trials of the experiment, see Figure 3.5 for an example output from the maximum likelihood function. The mean threshold computed for the observer role condition was 0.321m (SD = 0.246) and the mean threshold value for the actor role was 0.406m (SD = 0.541).

A 2 (viewpoint: third-person, first-person) $\times$ 2 (gender: male, female) repeated measures ANOVA showed no significant main of effect of viewpoint F (1, 14) = 0.4, $p = 0.537$ or gender F (1, 14) = 1.622, $p = 0.224$. There was also no interaction F (1, 14) = 0.849, $p = 0.372$.

3.7.5 Discussion

We found no significant difference between users' ability to make judgments about the landing point of ball which they threw versus the landing point of a ball which was thrown by someone else. This result is surprising because in one case the subject has both their own proprioceptive cues as well as the visual information from their body available to them and in the second case only the visual information is present.

One explanation is that the visual information from watching a representation of our own body is not usually available to us and may have caused conflict when making the judgment. While this point of view is not completely foreign to us (this is a common viewpoint in video games), it could have caused the endpoint of the ball as determined through the user's initial action plan and the endpoint of the ball as viewed by the user in the VE to conflict. A subject may have expected the ball to land 3 meters in front of them, but instead saw it land 7 meters ahead of them in the VE because of the displacement of the viewpoint from the self-avatar in the scene. In the case of the observer, however, there was no motor information with which the visual information had to be resolved. A second explanation is that visual information alone is so dominant over any other sensory input that adding other feedback has no effect.

3.8 General Discussion

We conducted three experiments investigating perception and action in an immersive VE. Throughout the three experiments we asked users to make judgments on the fidelity of the flight path of a thrown ball; in some cases depriving them of visual or motor feedback.

In Experiments 2a and 2b, the reason why vertical and horizontal errors were similar in terms of their detectability is not obvious. Reitsma et al. [2008] found that subjects could detect horizontal velocity errors easier than vertical velocity errors; we note that it is difficult to quantitatively compare our study with Reitsma et al. because of methodological differences. One difference between their study and the current study is that they had subjects make judgments on the ballistic motion of a cannonball from an observer's point of view, whereas in the current study subjects judged the actions that they performed themselves and from their own first person point of view. Self-initiated actions are associated with sensory predictions based on the initial motor intention of the performer; so, in the case of self-throwing, sensory expectations from the action plan, along with acquired motor skills based on past experience with throwing, are also parts of the judgment. Thus in Experiments 2a and 2b, the additional information from the motor plan itself might have been useful to compensate for the discrepancy between vertical/horizontal velocity judgments that Reitsma et al. [2003] reported for observed actions. Thus, people might be equally sensitive to vertical and horizontal velocity changes when they are the ones performing the action to be judged.

Behavioral differences associated with gravity perturbations may be because gravity is ubiquitous and special. Gravity is a world invariant and the human conception of it is resistant to modifications [Zago et al. 2004]. Because we have no experience with alterations in gravity's affect on everything we come in contact with, it is difficult to perceive changes to gravity. Another explanation for perceptual differences with gravity could be due to the fact that velocity, with units of m/s, is the first derivative of position, while acceleration, with units of m/s^2, is the second derivative. It is interesting to note, then, that although there is a large body of research that tells us that both acceleration and jerk (the derivative of acceleration) play a large role in whole-body motion perception [Benson et al. 1986; Grant and Haycock 2008; Naseri et al. 2008], when visually judging the effects of one's motion, people are less sensitive to changes in acceleration. In future work, we could look at altering the jerk experienced by the ball and see if the same decrease in sensitivity occurs.

Using the data available to us after conducting Experiment 2a we were able to make some assertions about the importance of the endpoint when making a judgment about errors introduced into a trajectory. Taking the mean values of the initial velocities across all throws by a subject in Experiment 2a, the standard deviations in the vertical and horizontal directions are .7164 and .6128, respectively. Both of these values are far larger than the discrimination thresholds of vertical and horizontal velocity, .3141 and .3830, respectively and collapsing across direction (as there is no difference between perception of increases or decreases in either perturbation). The variation in the throws of each subject, then, exceeds the amount of variance introduced by the alterations, and leads us to believe that while the ending point may have been one factor that was considered when making a judgment, it was not the only deciding factor. This idea is also supported by Figures 3.4, 3.6, and 3.6. The magnitudes of the sensitivities for the increased horizontal velocity and decreased gravity conditions from Figures 3.6 and 3.6 are similar (note the base difference in units, m/s and m/s^2, however), but this should probably not be the case if endpoints are being used as a deciding factor, as the ending point of a trajectory experiencing increasing horizontal velocity is much farther from the true endpoint than the stopping point of the ball in the decreased gravity condition (see Figure 3.4). Experiment 2b further investigated these claims using a procedure in which the trajectory of the ball was hidden, leaving the endpoint as the only visual cue with which to make a judgment. The threshold values from Experiment 2b, while slightly higher, do not vary significantly from those of Experiment 2a. These results, while indicating that the endpoint is enough information without the trajectory to make a judgment, does not rule out other deciding factors such as the proprioceptive cues available during the performance of the action as well as the timing of the trajectory from beginning to end.

Experiment 2b also led us to conclude that motor information alone is enough to differentiate between the ending point of a throw and the ending point of an altered throw. The subjects in Experiment 2b had limited visual information with which to make their judgments as the trajectory of the ball was eliminated leaving only the movement of the virtual hand (which was occluded for the majority of trial do to the limited FOV of the HMD), the VE and the endpoint of the ball. Experiment 2c built on these conclusions and further clarified the roles of visual and motor information in the perception of actions. We saw that visual information obtained from watching a person throw a ball is enough to accurately discriminate between a true and an altered flight path. The results of Experiments 2b and 2c suggest that, from a multi-sensory perspective, either motor or visual information alone is sufficient when making judgments, and there is no additive or synergistic effect when both modes of information are available.

When performing an action, an initial action plan is created which predicts the outcome of our action. Throughout the course of the action this plan may change, but in the case of Experiments 2b and 2c, with the trajectory of the ball remaining hidden, there was no information which could have been used to update the action plan. Therefore, even though it may seem that the added motor information available when playing the role of the actor should allow for better judgments to be made, because the mirror-neuron system offers a direct mapping between observed actions and performed actions [Rizzolatti and Craighero 2004], our initial action plan predicting the endpoint of the ball may be unaffected by the lack of motor information when acting as an observer.

Studying the perception of throwing in a VE as we did allowed us to decouple components of motion that could not be decoupled easily in the real world. This work thus falls within a category of work in perception and action in VEs that is of considerable significance for learning and training, since such studies can provide realistic or modified feedback about the consequences of actions taken with tools or artifacts. Furthermore, the ability to systematically manipulate feedback about the consequences of actions is important for discovering the degree of fidelity to real world consequences that is required for effective training and transfer of training to occur.

Chapter IV

CONCLUSION

We performed two studies which investigated human performance as well as human perception and action within an immersive VE. From the first study we learned that adding a human representation, either in the form of a self-avatar or another character, can help to improve performance on visually-based complex tasks. From the second study we learned both about human ability to perceive errors introduced into self-produced actions, as well as the roles of visual and motor feedback while making judgments about the outcomes of these same actions. A preliminary version of the results in Chapters 1 and 2 were presented at the conference on Applied Perception in Graphics and Visualization 2011 [McManus et al. 2011; McManus et al. 2011; McManus et al. 2011], and journal versions of both chapters are in preparation.

Taken together, we can use these findings to change our approach to designing VEs. From both studies, it seems apparent that, as would be expected, more closely modeling the real-world will elicit better performance. It is rare to find oneself in a situation without any other humans with whom to interact and even more rare to find oneself without vision of one's own body. Adding other characters to the environment as well as a self-representation makes the virtual world more accessible to the user and provides familiar cues that the user can learn from and use while performing tasks. We also find that while observing other characters performing an action we are able to make judgments about the outcome of an action as well as if we were performing the action ourselves, which only furthers the argument for adding human representations to the VE.

The available sensory feedback is also another important aspect to consider during the design of a VE. The rubber hand illusion is a perfect example showing the dominance of vision over other senses [Botvinick 1998] and our results further emphasize these findings. The mirror-neuron system helps to explain how we learn to perform a new task simply through observation and when deprived of motor feedback how we can still make judgments about the outcome of another person's actions [Rizzolatti and Craighero 2004]. It is necessary, then, to pay attention to all of the sensory feedback available in a VE, keeping visual feedback as a priority.

One of the fourteen grand challenges for engineering, as proposed by the National Academy of Engineering [2012], is to enhance virtual reality. We believe that the results from this paper provide a stepping stone along the path to meeting this challenge by offering a look at improving human performance and understanding perception and actions within an immersive VE.

REFERENCES

2012. The national academy of engineering, grand challenges for engineering : Enhance virtual reality. `http://www.engineeringchallenges.org/cms/8996/9140.aspx`.

Bard, C., Hay, L., and Fleury, M. 1985. Role of peripheral vision in the directional control of rapid aiming movements. *Canadian Journal of Psychology 39*, 151–161.

Bartlett, R. 2008. Principles of throwing. In *Biomechanics in Sport*. Blackwell Science Ltd, Oxford, UK, 365–380.

Benson, A., Spencer, M., and Stott, J. 1986. Thresholds for the detection of the direction of whole-body, linear movement in the horizontal plane. *Aviation, Space, and Environmental Medicine 57,* 11, 1088–1096.

Bideau, B., Kulpa, R., Ménardais, S., Fradet, L., Multon, F., Delamarche, P., and Arnaldi, B. 2003. Real handball goalkeeper vs. virtual handball thrower. *Presence: Teleoperators and Virtual Environments 12,* 4, 411–421.

Bodenheimer, B., Meng, 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 *Proceedings of the 4th symposium on Applied perception in graphics and visualization*. APGV '07. ACM, New York, NY, USA, 35–40.

Botvinick, M. 1998. Rubber hands' feel'touch that eyes see. *Nature 391*, 756.

Di Luca, M. 2010. New method to measure end-to-end delay of virtual reality. *Presence: Teleoperators and Virtual Environments 19,* 6, 569–584.

Dodds, T., Mohler, B., and Blthoff, H. 2010. A communication task in hmd virtual environments: Speaker and listener movement improves communication. *Proceedings of the 23rd Annual Conference on Computer Animation and Social Agents (CASA 2010)*, 1–4.

Dupuy, M. A., Mottet, D., and Ripoll, H. 2000. The regulation of release parameters in underarm precision throwing. *Journal of Sports Sciences 18*, 375–382.

Durlach, N. and Slater, M. 2000. Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators and Virtual Environments 9,* 1, 214–217.

Gibson, J. J. 1979. *The ecological approach to visual perception.* Houghton Mifflin, Boston, MA.

Grant, P. and Haycock, B. 2008. Effect of jerk and acceleration on the perception of motion. *Journal of Aircraft 45,* 4, 1190–1197.

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,* 4 (July), 26:1–26:18.

Hecht, H. and Bertamini, M. 2000. Understanding projectile acceleration. *Journal of Experimental Psychology: Human Perception and Performance 26,* 2, 730–746.

Henry, D. and Furness, T. 1993. Spatial perception in virtual environments: Evaluating an architectural application. *Virtual Reality Annual International Symposium*, 33–40.

Hommel, B., Msseler, J., Aschersleben, G., and Prinz, W. 2001. The theory of event coding (tec): a framework for perception and action planning. *Behavioral and Brain Sciences 24,* 5, 849–878.

JARDINE, R. AND MARTIN, N. G. 1983. Spatial ability and throwing accuracy. *Behavior Genetics 13*, 331–340. 10.1007/BF01065771.

JAYARAM, S., CONNACHER, H. I., AND LYONS, K. W. 1995. Virtual assembly using virtual reality techniques.

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. ACM, New York, NY, USA, 9–14.

KUHL, S. A., CREEM-REGEHR, S. H., AND THOMPSON, W. B. 2006. Individual differences in accuracy of blind walking to targets on the floor. *Journal of Vision 6,* 6, 726.

LOK, B. C. 2004. Toward the merging of real and virtual spaces. *Communications of the ACM 47,* 8 (August), 21–46.

LOOMIS, J. AND KNAPP, J. 2003. Visual perception of egocentric distances in real and virtual environments l. j. hettinger and m. w. haas, eds. erlbaum, mahwah, nj, ch. 2. *Virtual and Adaptive Environments*, 21–46.

MCDONNELL, R., LARKIN, M., HERNÁNDEZ, B., RUDOMIN, I., AND O'SULLIVAN, C. 2009. Eye-catching crowds: saliency based selective variation. *ACM Trans. Graph. 28,* 3 (July), 55:1–55:10.

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. ACM, New York, NY, USA, 37–44.

MCMANUS, E. A., ERDEMIR, A., BAILEY, S. W., RIESER, J., AND BODENHEIMER, B. 2011. Using endpoints to judge alterations in self-produced trajectories in an immersive virtual environment. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*. APGV '11. ACM, New York, NY, USA, 122–122.

MCMANUS, E. A., LIN, Q., ERDEMIR, A., BAILEY, S. W., RIESER, J., AND BODENHEIMER, B. 2011. Perceiving alterations in trajectories while throwing in a virtual environment. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*. APGV '11. ACM, New York, NY, USA, 23–28.

MOHLER, B., CAMPOS, J., WEYEL, M., AND BÜLTHOFF, H. H. 2007. Gait parameters while walking in a head-mounted display virtual environment and the real world. *Proceedings of the 13th Eurographics Symposium on Virtual Environments and 10th Immersive Projection Technology Workshop (IPT-EGVE 2007)*, 85–88.

MOHLER, B., CREEM-REGEHR, S. H., THOMPSON, W., AND BLTHOFF, H. H. 2010a. The impact of an animated avatar on egocentric distance perception in an immersive virtual environment. *European Conference on Visual Perception 39 ECVP Abstract Supplement*, 51.

MOHLER, B. J., BLTHOFF, H., THOMPSON, W., AND CREEM-REGEHR, S. 2008. A full-body avatar improves egocentric distance judgments in an immersive virtual environment. *ACM Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization (APGV 08)*, 194–197.

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. Symposium on Applied Perception in Graphics and Visualization*. ACM, 9–14.

Mohler, B. J., Creem-Regehr, S. H., Thompson, W. B., and Blthoff, H. H. 2010b. The effect of viewing a self-avatar on distance judgments in an hmd-based virtual environment. *Presence: Teleoperators and Virtual Environments*.

Naseri, A., Grant, P. R., and Dufort, P. 2008. Modeling the perception of acceleration and jerk using signal detection theory. In *AIAA Modeling and Simulation Technologies Conference and Exhibit*.

O'Sullivan, C. and Ennis, C. 2011. Metropolis: Multisensory simulation of a populated city. In *Games and Virtual Worlds for Serious Applications (VS-GAMES), 2011 Third International Conference on*. 1 –7.

Philbeck, J. W., Woods, A. J., Arthur, J., , and Todd, J. 2008. Progressive locomotor recalibration during blind walking. *Percept Psychophys 70,* 8 (Nov), 1459–1470.

Poupyrev, I., Weghorst, S., Billinghurst, M., and Ichikawa, T. 1998. Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques. *Computer Graphics Forum*.

Prinz, W. Perception and action planning. *European journal of cognitive psychology 9*.

Proteau, L., Boivin, K., Linossier, S., and Abahnini, K. 2000. Exploring the limits of peripheral vision for the control of movements. *Journal of Motor Behavior 32*, 277–286.

Reichenbach, A., Thielscher, A., Peer, A., Blthoff, H., and Bresciani, J.-P. 2009. Seeing the hand while reaching speeds up on-line responses to a sudden change in target position. *The Journal of Physiology 587,* 19, 4605–4616.

Reitsma, P. S. A., Andrews, J., and Pollard, N. S. 2008. Effect of character animacy and preparatory motion on perceptual magnitude of errors in ballistic motion. *Comput. Graph. Forum 27,* 2, 201–210.

Reitsma, P. S. A. and Pollard, N. S. 2003. Perceptual metrics for character animation: sensitivity to errors in ballistic motion. In *Proc. ACM SIGGRAPH*. 537–542.

Richardson, A. R. and Waller, D. 2007. Interaction within 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 Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 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 *16th ACM Symposium on Virtual Reality Software and Technology (VRST '09)*, S. N. S. (Ed.), Ed. ACM, ACM, New York, NY, USA, 59–66.

Rizzo, A. A., Klimchuk, D., Mitura, R., Bowerly, T., Buckwalter, J. G., Kerns, K., Randall, K., Adams, R., Finn, P., Tarnanas, I., Sirbu, C., Ollendick, T. H., and Yeh, S.-C. 2006. A virtual reality scenario for all seasons: the virtual classroom. *CNS Spectrum 11*, 35–44.

Rizzolatti, G. and Craighero, L. 2004. The mirror-neuron system. *Annual Review of Neuroscience 27*, 169–192.

Sahm, C. S., Creem-Regehr, S. H., Thompson, W. B., and Willemsen, P. 2005. Throwing versus walking as indicators of distance perception in real and virtual environments. *ACM Transactions on Applied Perception 1,* 3, 35–45.

Saunders, J. and Knill, D. Visual feedback control of hand movements. *Journal of Neuroscience 24,* 13.

Schomaker, J., Tesch, J., Bülthoff, H., and Bresciani, J.-P. 2011. It is all me: the effect of viewpoint on visual-vestibular recalibration. *Experimental Brain Research 213*, 245–256.

Sebanz, N., Bekkering, H., and Knoblich. Joint action: bodies and minds moving together. *Trends in Cognitive Sciences 10,* 2, 70–76.

Slater, M. and Usoh, M. 1994. Body centered interaction in immersive virtual environments. *In N. M. Thalmann & D. Thalmann (Eds.), Artificial life and virtual reality. New York: John Wiley and Sons.*, 125–148.

Spijkers, W. and Spellerberg, S. 1995. On-line control of rapid aiming movements? *Acta Psychologica 90*, 333–348.

Streuber, S. and de la Rosa, S. 2010. The role of body and tool-based information in joint action coordination. *European Conference on Visual Percpetion Conference Abstract*.

Streuber, S., de la Rosa, S., Trutoiu, L.-C. C., Blthoff, H. H., and Mohler, B. 2009. Does brief exposure to a self-avatar affect common human behaviors in immersive virtual environments? *Eurographics 2009: The 30th Annual Conference of the European Association for Computer Graphics, 1-4. (Eds.) Stamminger, M., P. Dutr, Oxford, UK, Blackwell*.

Streuber, S., Knoblich, G., Sebanz, N., Blthoff, H. H., and de la Rosa, S. 2011. The effect of social context on the use of visual information. *Experimental Brain Research 214*, 273–284.

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.

Van Rossum, J. H. A. and Bootsma, R. J. 1989. The underarm throw for accuracy in children. *Journal of Sports Sciences 7*, 101–112.

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. ACM, New York, NY, USA, 53–63.

Welch, R. B. and Sampanes, A. C. 2008. Adapting to virtual environments: Visual-motor skill acquisition versus perceptual recalibration. *Displays 29,* 2, 152 – 158. Health and Safety Aspects of Visual Displays.

Williams, B., Narasimham, G., Westerman, C., Rieser, J., and Bodenheimer, B. 2006. Functional similarities in spatial representations between real and virtual environments. *ACM Transactions on Applied Perception*.

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. ACM, New York, NY, USA, 35–40.

Zago, M., Bosco, G., Maffei, V., Iosa, M., Ivanenko, Y. P., and Lacquaniti, F. 2004. Internal models of target motion:expected dynamics overrides measured kinematics in timing manual interceptions. *Journal of Neurophysiology*, 1620–1634.