Summer 2012

Dynamic full-body interactions in virtual reality : understanding effects of display and locomotion modality on perception and action

Timofey Yurievich Grechkin
University of Iowa

Recommended Citation

Grechkin, Timofey Yurievich. "Dynamic full-body interactions in virtual reality : understanding effects of display and locomotion modality on perception and action." PhD (Doctor of Philosophy) thesis, University of Iowa, 2012. https://ir.uiowa.edu/etd/3307.

DYNAMIC FULL-BODY INTERACTIONS IN VIRTUAL REALITY: UNDERSTANDING EFFECTS OF DISPLAY AND LOCOMOTION MODALITY ON PERCEPTION AND ACTION

by

Timofey Yurievich Grechkin

<u>An Abstract</u>

Of a thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Computer Science
in the Graduate College of
The University of Iowa

July 2012

Thesis Supervisor: Professor Joseph K. Kearney

ABSTRACT

Many practical applications of Virtual Reality (VR) technology rely on adequate immersive representations of 3D spaces and support of embodied, dynamic interactions with the virtual world. Evaluation of these properties remains an important research problem.

This thesis aims at developing a method of conducting user evaluations of dynamic, full-body interactions in VR systems based on using support for perception and action coupling as a criterion for comparison. The thesis has three main components.

First, the thesis starts by presenting an experimental study of distance perception in real and virtual environments. The results indicate that the choice of the method to report perceived distances may have a significant effect on the outcome of a study. We argue for the need to develop an approach to VR evaluation that holistically considers both perception and action.

Second, we propose a theoretical framework to conduct such user evaluations based on the notion of affordances. The thesis presents a second experimental study that explores perception of affordances in a complex, realistic task of bicycling across two lanes of opposing traffic in a VR simulator. This experiments highlights a methodological approach to studies of user's perception of dynamic affordances.

Finally, we present an experimental study that builds on theoretical and methodological frameworks developed in the thesis to explore the effects of display type and locomotion modality on user performance in a dynamic VR task that involves synchronization of self-motion with the motion of virtual objects. The results inform our understanding of the trade-offs involved in selecting major components of the VR system.

Abstract Approved: ______________________________
Thesis Supervisor

____
Title and Department

____
Date

DYNAMIC FULL-BODY INTERACTIONS IN VIRTUAL REALITY: UNDERSTANDING EFFECTS OF DISPLAY AND LOCOMOTION MODALITY ON PERCEPTION AND ACTION

by

Timofey Yurievich Grechkin

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Computer Science
in the Graduate College of
The University of Iowa

July 2012

Thesis Supervisor: Professor Joseph K. Kearney

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Timofey Yurievich Grechkin

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Computer Science at the July 2012 graduation.

Thesis Committee: ______________________________
Joseph Kearney, Thesis Supervisor

____
Jodie Plumert

____
James Cremer

____
Juan Pablo Hourcade

____
Alberto Maria Segre

____
Geb Thomas

ABSTRACT

Many practical applications of Virtual Reality (VR) technology rely on adequate immersive representations of 3D spaces and support of embodied, dynamic interactions with the virtual world. Evaluation of these properties remains an important research problem.

This thesis aims at developing a method of conducting user evaluations of dynamic, full-body interactions in VR systems based on using support for perception and action coupling as a criterion for comparison. The thesis has three main components.

First, the thesis starts by presenting an experimental study of distance perception in real and virtual environments. The results indicate that the choice of the method to report perceived distances may have a significant effect on the outcome of a study. We argue for the need to develop an approach to VR evaluation that holistically considers both perception and action.

Second, we propose a theoretical framework to conduct such user evaluations based on the notion of affordances. The thesis presents a second experimental study that explores perception of affordances in a complex, realistic task of bicycling across two lanes of opposing traffic in a VR simulator. This experiments highlights a methodological approach to studies of user's perception of dynamic affordances.

Finally, we present an experimental study that builds on theoretical and methodological frameworks developed in the thesis to explore the effects of display type and locomotion modality on user performance in a dynamic VR task that involves synchronization of self-motion with the motion of virtual objects. The results inform our understanding of the trade-offs involved in selecting major components of the VR system.

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
INTRODUCTION

Virtual Reality (VR) technology [10, 90] provides a valuable human-computer interface for many practical applications of computing. It is characterized by interactivity, muti-sensory feedback, and mental immersion of the user into the simulated virtual world. A formal definition of VR offered by Sherman and Craig [90] reads as follows (p. 13):

> *Virtual reality* is a medium composed of interactive computer simulations that sense the participant's position and actions and replace or augment the feedback to one or more senses, giving the feeling of being mentally immersed or present in the simulation (a virtual world).

From its early development stages in the 1960s VR strived to achieve the quality of the interaction that Dourish [21] later described (in a somewhat different context) as *embodiment*, the sense of embodied presence in a computer-mediated world. One of the pioneers of virtual reality technology Morton Heilig, who developed an early prototype [38] (Figure 1.1, left) of today's head-mounted display, placed particular emphasis on creating immersive experiences for the user. Although he primarily viewed virtual reality through the prism of video recordings, Heilig had a revolutionary vision of using multi-sensonsory feedback to enhance the sense of immersion into simulated world. His visual arcade "Sensorama" [39] (Figure 1.1, right) featured stereoscopic, wide-field-of-view visual feedback, "motion, color, stereo sound, aromas, wind effects (using small fans placed near the user's head), and a seat that vibrated" [10] (p. 3).

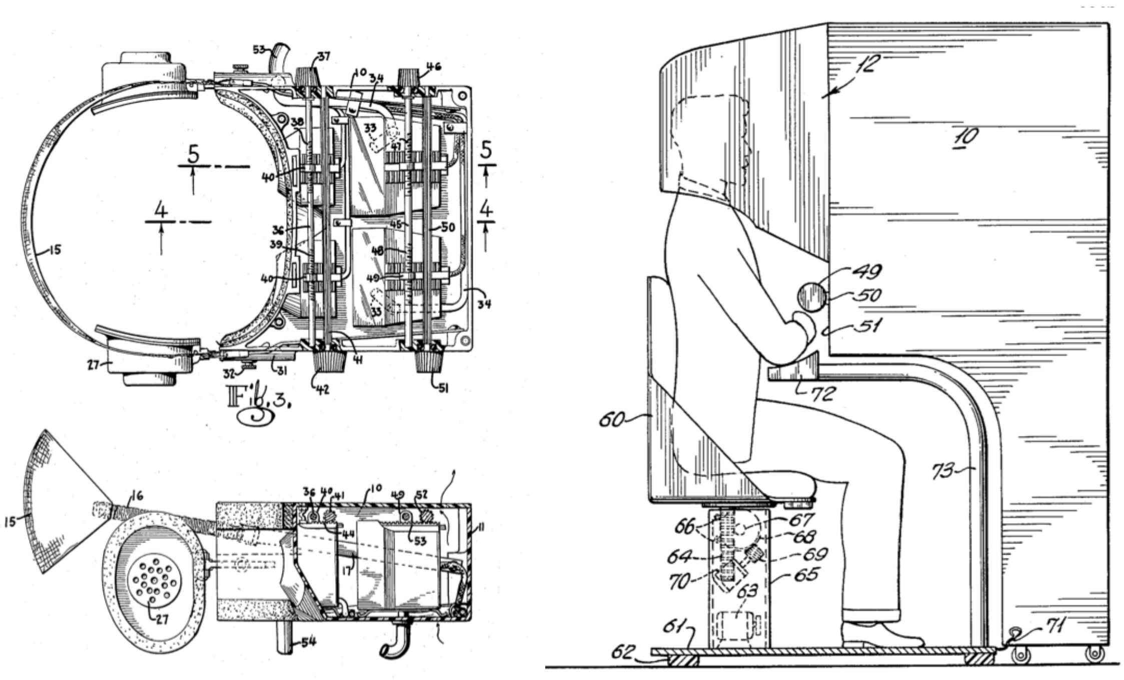

Figure 1.1: Early prototypes of virtual reality technology: 1960 “Stereoscopic-television apparatus for individual use” [38] (left) and 1962 Sensorama simulator [39] (right). Images courtesy of United States Patent and Trademark Office.

The work of Ivan Sutherland [96, 97] marked the shift of VR visuals from video to computer-generated 3D graphics. On the one hand, this development allowed the users of VR systems to experience virtual worlds that do not otherwise exist. On the other hand, with advances in computer graphics it gave VR designers and operators an unprecedented amount of control over increasingly realistic, three-dimentional representations of the real world. This latter aspect opened the door for VR-based computer simulation. Today, two types of physical displays are used most often: head-mounted displays (HMDs) and large projection screens.

Advances in tracking technology liberated the user from maintaining a fixed position in front of a display and, to an increasing extent, allowed the user to explore and interact with the virtual worlds using their own body (through head rotation, changes in posture, and naturalistic locomotion).

Although other types of sensory feedback remained important components of VR experience, the focus of research and engineering attention have historically been on improvement in real-time interactive sensor-based visual feedback [7].

1.1 A survey of practical applications of Virtual Reality

Today's VR boasts a large number of practical applications [15]. This brief survey is meant to highlight some common themes and requirements on VR systems that originate from various practical uses of the technology.

1.1.1 Medical Applications

An increasing number of existing applications [10, 15, 88] shows growing acceptance of VR in medical community.

VR offers great promise for education and training. The Anatomic VisualizeR project [10] (p. 287-289) leveraged a database of realistic 3D models of human organs developed from the Visible Human database to create an immersive visualization dissection room. It allows students to gain knowledge of anatomical structure in a cost-efficient way. In a similar manner, three-dimentional reconstructions of internal organs built using Computer Tomography (CT) and Magnetic Resonance Imaging (MRI) data can assist practicing medics in diagnosis and operation planning [88].

The BDI Surgical Simulator [15] (p. 146-153) is an example of surgical training applications [36] that allow medical students to practice in performing surgical procedures. The user performs an operation using real instruments that are augmented with haptic force-feedback, while viewing a realistically depicted process on a screen. The key aspects that make this application work are the realistic depiction of the operation process and the naturalistic interface resembling real-life manipulations of the surgeon during the operation.

A number of applications used VR as a tool for psychological rehabilitation via exposure treatment [15] (p. 178-187). In these applications the treatment of psychological disorders such as phobias and post-traumatic stress disorder (PTSD) causing overwhelming anxiety and fears is based on gradual exposure to the situations that cause discomfort to a patient. The systems work because patients experience the sense of mental immersion into simulated scenario that helps create realistic experiences and gradually adapt the patient to the situation in question [40]. The treatment specialist needs a way to control the scenario and the grade of exposure (the level of realism of the situation). In this application multi-sensory feedback plays central role, giving greater role to haptic and auditory feedback [7] (p. 26).

Finally, medical rehabilitation studies [15] (p.157-164), such as the work of Strickland [94] report on success in using VR for teaching special needs children the skills required to complete dangerous every-day tasks such as crossing a road with traffic. Using a similar approach Whitney et al.[111] report on a pilot study that explored the use of a VR setup for rehabilitation treatment of patients with vestibular dysfunctions, who experience difficulties in visually complex environments. For all these applications immersive, realistic simulation and the ability to create tightly controlled, safe environment for the patient once again plays a critical role in determining the success of the application.

1.1.2 Industrial Prototyping, Visualization, and Architectural Walkthroughs

Leveraging the ability to create immersive 3D representations of objects and environments that do not yet exist or are not easily accessible, a number of applications use VR for prototyping and visualization. The use of VR representations of future products for evaluation of industrial prototypes is common in auto industry [7](p. 22), [15] (p. 84-92) and a wider field of engineering [22], [15] (p.71-83). In a similar fashion VR can serve as a useful visualization tool for decision support in urban planning [50], exploration of natural resources [87], and architectural walk-throughs [8, 9]. All these applications rely on VR's potential to offer insight into the 3D structure of the objects and the environments as well as relationship between their constituent elements.

Reconstruction of historical environments and artifacts opens new doors for educational and scientific applications and puts an interesting twist on the same idea, re-creating places and objects that no longer exist. For example, Cremer et al. [19, 18] describe a reconstruction project depicting Cedar Rapids, Iowa circa 1900 for an exhibit in a local museum. In this and similar applications [35, 47, 79] VR provides a unique, immersive, first-person experience.

1.1.3 Simulation, Training, and Entertainment

Simulators of complex engineering systems such as driving [80] and flight [53] simulators are among the early success stories in VR. Describing his experience in a British Airways Boeing 747 simulator [7] (p. 20-22), Brooks characterized it as "a stunningly good illusion - the best VR I have ever experienced". In part this is due to the fact that the key elements of real-life experience associated with using hardware controls can be replicated almost exactly and then supplemented with visual feedback. The ability to closely replicate the experience of using various types of complex equipment also underlines a wide-spread use of VR for military training applications [15] (p. 266-272) and [10](p. 328-342).

VR has a great promise for entertainment. One of the early successes is using for immersive storytelling experiences was Disney's Aladdin exhibit [73]. There have also been some experiments in using VR for entertaining theatrical audiences [55]. However, because of the high cost of VR equipment these systems have not yet reached home entertainment market. The recent emergence in low-cost tracking systems such as Nintendo Wii and Microsoft Kinect as well as low-cost consumer-grade stereoscopic displays opens a great promise for VR at home. In all these applications interactivity and the sense of immersion are critical for maintaining user's engagement.

1.1.4 Common themes and system requirements

The survey uncovers a number of recurring themes and common requirements to VR systems:

1. The sense of embodied presence in the virtual world is a key distinguishing characteristic for many applications. Examples have shown that this is a complex phenomenon that can be attributed to many elements of the VR experience: realistic interactions, good quality of graphics, mult-sensory feedback, etc. Understanding how the sense of presence can be affected through combination of these characteristics is one of the requirements for the VR system development.
2. Many applications rely on VR's ability to provide an immersive 3D experience in a virtual space. The key underlying requirement is to represent virtual spaces in such a way that they are adequately perceived by the users. In particular, when virtual environments replicate real spaces and objects, it is desirable that

the user's perception of virtual space is similar to that of the modeled real space.

3. In a similar manner it is often important to provide naturalistic way of interacting with the virtual environment reminiscent of the interaction that occur in the real world. It also seems to be the case that naturalistic interactions might play a role in inducing the sense of presence.
4. The surveyed systems reflect a wide range of technical implementations used in VR setups. This variety leads to the need to make informed decisions in selecting technical implementation that would well match the practical problem.

1.2 Virtual Reality as a laboratory for studies of human behavior

The survey presented above neglected one of the most promising applications of virtual reality as a research tool in psychology [56], primarily because it deserves special consideration.

The use of VR as an experimental laboratory to study human behavior stands among the most important applications of the technology. Virtual reality promises a unique opportunity to conduct research in an environment close in realism to that of the real world, but with control over events, replicability of trials, and a level of safety intrinsic to laboratory studies. The safety aspect is particularly important, because it opens the door for investigations of potentially risky situations that are very difficult to study in the real world. For example, the problems with both precise control over the experimental conditions and the level of safety of the participants

Figure 1.2: Virtual reality setups for behavioral studies with large projection screens (left) and HMD (right) used in Hank virtual reality laboratory.

make real-world studies of crossing a traffic-filled road all but impossible. Yet the potential insight into the mechanisms driving the behavior is extremely valuable for training, incident prevention, and so on.

Psychology applications place general requirements on the VR systems that closely resemble common requirements outlined above. Because in VR studies virtual environments are used as a substitute for the real world, the issue of ecological validity is of special importance. In other words, the circumstances reconstructed in an experiment and the behaviors demonstrated by the participants should be close approximations of the real-life situations and behaviors under investigation.

Another important problem is the comparison between various experimental VR setups [56]. Different laboratories and even different studies within the same laboratory use different implementations of the VR technology (Figure 1.2), which may be a confounding factor in an experiment.

Cross-comparison between VR technologies and with respect to the real world in terms of their impact on user behavior is an important research problem. Interestingly, it is also a part of a positive feedback loop. Evaluations of human behavior in order to better understand and improve VR systems lead to better tools for studying human behavior.

1.3 Key research objectives of this dissertation

The research described in this dissertation is primarily focused on evaluating how well can VR systems support dynamic, full-body interactions with immersive virtual environments. These kinds of interactions are critical for behavioral studies and other practical applications of VR technologies that we discussed above. The psychological studies in particular need quantitive evaluations that show any potential differences between user's behavior in virtual reality and the real word. This dissertation attempts to develop an approach to evaluate users' behavior in such dynamic interactions.

1.4 Outline of the Thesis

This chapter discussed the defining characteristics, history, and applications of the VR technology. It also outlined key common requirements of virtual reality systems. A special degree of attention was given to applications of virtual reality as a laboratory to study human behavior. Finally, section 1.3 presented the main research goals behind this dissertation. The rest of the thesis is organized as follows:

- Chapter 2 discusses perceptual evaluation of VR technology in comparison with the real world. It focuses on the interconnection between perception and action. Because perception is impossible to study directly, researchers are forced to select an action that can be to used measure perceptual phenomena. The chapter presents an experimental study that explores how the choice of action to express perceived distance may affect the outcome of an experiment comparing distance perception in real and virtual environments. The results supply experimental evidence that shows that the choice of a response action can significantly influence the outcomes of perceptual evaluation experiments. In conclusion the chapter suggests the need for holistic approach that looks at perception and action together.

- Chapter 3 focuses on a theoretical framework that can be used to connect perception and action in the context of a computer-mediated environment. In particular it explores the notions of affordance and coupling in psychology and computer science. In the conclusion of the chapter the key experimental objectives of this dissertation are re-formulated in terms of affordances and coupling between perception and action.
- Chapter 4 presents an experiment that can serve as a case study describing how complex, dynamic affordances can be explored using virtual reality and discusses implications for the experimental goals of this dissertation.
- Chapter 5 develops the experimental framework more formally. It discusses experimental methods, the choice of conditions, tasks, procedures, and measures for an experimental study comparing the effects of display type and locomotion

modality on perceiving dynamic affordances in full-body, interactive, motion-coordination task in VR.

- Chapter 6. Discusses experimental results of the study conducted using the methodological approach developed in Chapter 5.
- Finally, Chapter 7 presents general conclusions and outlines future research directions.

CHAPTER 2
PERCEPTUAL EVALUATION OF VIRTUAL REALITY SYSTEMS AND THE NEED FOR HOLISTIC APPROACH TO PERCEPTION-ACTION COUPLING

2.1 Evaluation of distance perception in virtual reality systems

Many practical applications rely on user's perception of virtual environment as an adequate representation of some real-world environment. One common way to assess perception of virtual spaces is to compare perception of distances to target objects in the real world and in its virtual representation. However perception of object heights (i. e. scale) [117], relationship between distance and scale [64], and travelled distances [65] have also been studied.

Perception can not be assessed directly and, therefore, in order to be measured, it has to be expressed through an appropriate action. A wide range of actions have been used to measure perceived distances [23, 57], for example:

- verbal report, where observers verbally report perceived distances in terms of some measurement unit;
- visually directed action, such as walking blindfolded towards a previously seen target;
- expression of imagined action, such as timing an imagined walking towards the perceived location of the target.

Importantly, the selected response action is performed in an open-loop manner in the absence of visual feedback, because humans continuously adjust scaling of perceived distance in terms of the selected corresponding action based on the visual feedback they gather from the environment.

There has been a number of studies that compared distance perception in the real world and in virtual environments [3, 5, 16, 45, 23, 63, 101, 114, 115, 120]. These studies have shown that people significantly underestimate distances in virtual environments even though they can accurately estimate distances up to 20 meters in the real world. The reasons for distance underestimation in virtual environments are not well understood, although several factors have been considered as potentially contributing to this phenomenon.

The display type is, perhaps, the most important component affecting visual perception and the properties of the specific display type may have a significant impact on the distance estimates. In the case of a head-mounted display, for example, users have to deal with restricted field-of-view, significant weight of the helmet, and a tipping torque created by the displacement of the weight in the HMD relative to the head's center of mass. Willemsen et al. [114] found that the head-mounted display (HMD), by itself, can contribute to underestimation of distances relative to the real world. Reduced field-of-view was also found be detrimental for distance judgments in the real world [116] and in setups with large immersive screens [23].

The surrounding context was shown to influence people's judgments of real-world distances [52]. In VR studies, the visual settings vary from indoor hallways [115, 101] to outdoor lawns [77, 120, 23]. An inconsistency in visual settings might make it difficult to compare results between studies.

The quality of rendering has been raised as yet another factor to potentially impact distance estimation in virtual environments. Loomis and Knapp [57] hypothesized that photorealistic rendering of virtual environments can lead to more accurate perception of distance. However, so far this assertion has not been experimentally supported. Thompson et al. [101] found significant underestimation of distances in a photo-based panoramic environment displayed in the HMD. Significant underestimation of distances was also reported in an HMD environment that showed images from a head-mounted video camera [60]. Klein et al. [23] found significant distance underestimation relative to the real world in a large screen display system using a photo-based panorama of an outdoor background in conjunction with a rendered virtual ground and target.

Interestingly, Richardson and Waller [81] have shown that after a brief interaction (unrelated to distance estimation task) with the virtual environment participants became significantly more accurate in distance estimation task. This seems to be consistent with the notion of the perception-action system as a continuously adjusted equilibrium sensitive to available feedback and scaling clues. There is also evidence that a body-scaled avatar provides sufficient clues to improve distance estimated in virtual environments [62]. Together these studies raise a question of whether the distance compression in the virtual environment is a stable perceptual phenomenon or simply a consequence of insufficient calibration of the measurement method to the virtual environment. On the other hand, Interrante et al. [42] put forward a hypothesis that distance perception in virtual environments may be related to the sense of embodied presence in the virtual world rather than to the perception-action calibration. If that hypothesis is true, some of the results discussed here can be explained in terms of increased sense of presence resulting from the interaction or the presence of the self-avatar.

One question that has not received sufficient attention is whether the outcome of a distance estimation task may depend substantially on the choice of the response action. It is possible that some aspects of virtual reality experience may affect the response action rather than perception. If so, the choice of the measurement protocol may prove to be a confounding factor.

The next section will examine the effects of the measurement protocol on the outcome of the distance estimation studies more closely.

2.2 Effects of measurement protocol on distance perception in real and virtual environments

What effect can a measurement protocol have on the results of a distance estimation study? If the choice of a protocol does have an effect, the first indication will be a disagreement between the results obtained using different protocols. Research studies generally show good agreement between various protocols for assessing distance perception in the real world. Philbeck and Loomis showed agreement between verbal reports and blindfolded walking towards the target [74]. Plumert et al.[77] reported close agreement between timed imagined walking and blindfolded walking to the target. Klien et al.[23] found very close agreement between verbal report, timed imagined walking, and triangulated walking. The overall agreement between measurement protocols in the real world appears to indicate that the choice of the protocol is unlikely to affect the outcome of a distance estimation study.

However, when the elements of virtual reality systems are introduced into experiments the picture begins to change. For example, Klein et al. [23] report differences in results obtained in two different setups with with large immersive screens: verbal report and timed imagined walking were not in agreement with triangulated walking. This suggests that measurement protocols can vary in their sensitivity to properties of the VR system.

In this section we will discuss the results of a study (originally published in [32] and also discussed by one of my colleagues, Tien Dat Nguyen as part of his dissertation [65]) that provides additional evidence that the measurement protocol can have a significant effect on distance estimation in real and virtual environments. The study investigated two commonly used measurement protocols: blindfolded walking and timed imagined walking. The participants were asked to estimate the distance to a pair of poles located 6 to 18 meters in front of them in a hallway setting (Figure 2.1). Each participant viewed the same hallway environment in one of the following six presentation methods:

1. Real: unrestricted real-world view of hallway
2. Real+HMD: real-world view of hallway seen through an HMD
3. Virtual+HMD: virtual model of hallway viewed in an HMD

Figure 2.1: Photo of the real (left) and view of the rendered virtual (right) hallway and targets used for distance estimation experiments (originally reported in [32]).

4. Virtual+LSID: virtual model of hallway viewed on multiple large screens in a CAVE-like arrangement (we refer to this display type as a large screen immersive display or LSID)
5. Photo+LSID: photo-based presentation[1] of hallway viewed on multiple large screens
6. AR: augmented-reality presentation of virtual target objects superimposed on a real hallway seen through HMD.

The primary aim of this study was to compare a number of different visual presentation methods using two measurement protocols, while keeping the setting, targets, distances, visual model, and the methods constant. While the study attempted

---

[1] Participants viewed a photographic panorama of the real hallway and real targets. In order to create perspectively-correct displays for subjects, we hired a professional photographer to capture the scene for eye heights ranging from 55 through 71 inches, in one-inch increments, at each target distance.

to answer a number of research questions (see [32] for details), this discussion will primarily focus on the role of measurement protocol in comparing the effects of two systems on distance estimation.

The study consisted of two experiments. One experiment compared the first five presentation methods using the timed imagined walking protocol suitable for assessing distance estimation in both LSID and HMD systems. The other experiment compared non-LSID presentation methods (conditions 1 through 3) along with the AR condition (condition 6) using a blindfolded walking protocol. Finally, we investigated the effects of measurement protocols by comparing the results in experimental conditions present in both experiments.

2.2.1 Experimental design and methods

2.2.1.1 Experimental Design

Both experiments used a between-subjects design. Each participant viewed the environment in one of the five presentation conditions in Experiment 1 or in one of the four presentation conditions in Experiment 2. Participants in Experiment 1 made their judgments using the timed imagined walking protocol, whereas participants in Experiment 2 used the direct blindfolded walking protocol. The targets were placed at a distance of 6, 9, 12, 15 or 18 meters. Each participant was presented with three random permutations of these five distances for a total of 15 trials.

2.2.1.2 Apparatus and Materials

In the HMD conditions, we used an nVIS nVisor ST head-mounted display system with optical see-through functionality. The HMD contains two small LCOS displays each with resolution of 1280 x 1024 pixels. Stereoscopic display was used in all HMD conditions. Convergence distance for our HMD was set by the manufacturer to 10m. The field of view is 40.5 degrees vertical and 49.5 degrees horizontal. The optical see-through functionality enables both a virtual-environment-only presentation and an augmented reality presentation, in which virtual objects are superimposed on a view of the real environment.

An Intersense Vistracker IS-1200 6 DOF optical tracker was mounted on the HMD to measure participants' position and orientation in the hallway. A black cloth was used to block out light around the sides and back of the HMD and a black piece of foam was attached underneath the HMD lenses to prevent participants from seeing the floor or their feet. The cloth had a flap that could be lifted in some conditions to allow participants to view the real environment.

In the LSID conditions, the VE was displayed on three 10-feet wide x 8-feet-high screens placed at right angles relative to one another, forming a three-walled room. The room's floor was one foot above the bottom of the screens, so the effective screen height was seven feet. Participants stood eight feet from the front screen, midway between the side screens. Three Projection Design F1+ projectors were used to rear project high-resolution graphics (1280 x 1024 pixels) onto the screens, providing participants with approximately (depending on the participant's height)

224 degrees horizontal and 46 degrees vertical FOV of nonstereoscopic, immersive visual imagery. The viewpoint of the scene was adjusted for each participant's eye height, but motion parallax was not available.

In the Real conditions, a simple blindfold was used to block out light during the blindfolded walking task. A laser rangefinder was used to measure distances. The targets were a pair of cylindrical poles. The poles were 0.30 m in diameter and 1.219 m tall. The distance between the centers of the poles was one meter. Virtual targets were a faithful representation of the real ones.

The virtual hallway model was built to match the real hallway as closely as as possible. We determined the HMDs field of view by positioning the HMD so that known hallway features aligned with the outer boundary of the scene visible through the HMD. To ensure proper registration of the virtual and real hallways, we needed to determine the relationship between the coordinate frame of the tracking camera mounted on top of the HMD and coordinate frame of the HMD's display screens. We measured the translational distances between the tracking camera mount and HMD screens by hand. To obtain a rotation matrix for the relative orientation, we placed the HMD on a person's head and used a see-through mode to visually align the rendered wireframe view of the doorway at the end of the virtual hallway and its with the real doorway.

2.2.1.3 Procedure

The experiments were carried out either in the immersive large screen environment in our laboratory or in the hallway outside the lab. To minimize exposure to the environments immediately prior to the experiment, participants met with the experimenter in the lobby of the building and were escorted to the hallway or the lab blindfolded. For HMD conditions, setup and calibration of the HMD were performed prior to reaching the hallway. The goal of the calibration process was to ensure that participants eyes were centered on the HMD display screens. We displayed a test-pattern-style image (cross hairs in the middle and a nested set of colored thin rectangular rings at the outer portion of the image) and first directed participants to use the top and back HMD fit adjustment knobs so that device was snug and they could see the same color at the extreme top and bottom of their view. Next, participants adjusted the HMD's eyepieces to center each eye horizontally on its screen. On the nVisor ST, each eyepiece has its own IPD (inter-pupillary distance) adjustment knob. Participants were told close one eye and use the corresponding IPD knob to adjust the eyepiece so that the same color was visible at the right and left edges of that eyes display. This was then repeated for the other eye.

Before each trial, we positioned the participant at the appropriate starting location. Participants had the opportunity to view the target for four to five seconds before the experimenter told them to close their eyes. Then the blindfold was replaced or the screens were turned off and participants were instructed to make their distance judgment via either blindfolded walking or timed imagined walking. No feedback was given at the end of a trial.

For timed imagined walking, participants started the stopwatch when they imagined starting to walk and stopped the stopwatch when they imagined reaching the target (without ever looking at the stopwatch). The experimenter then recorded the stopwatch time. Headphones were not used in the imagined movement conditions due to low ambient noise in the environment and the fact that participants remained stationary. This also minimized the encumbrance. After each participant completed all 15 trials, the experimenters obtained an estimate of that person's average walking speed by measuring how long it took him or her to walk 18 meters at a comfortable walking speed.

For blindfolded walking, participants walked until they thought they had reached the target. The experimenter then recorded the distance walked and escorted participants back to the starting position blindfolded. For blindfolded walking with the HMD, participants wore headphones that transmitted a constant white noise (with the exception of the Real condition for which participants wore no headphones). The headphones served a dual purpose of both blocking out some of the ambient noise of the hallway and as a guide for walking without vision. If participants deviated too far from the center of the hallway to the right, the white noise in the right ear headphone would get progressively louder. If participants moved too far to the left, the white noise in the left ear headphone would get louder. Differences in amplitude of white noise in the two ears allowed participants to self correct their travel direction while walking with eyes closed. The experimenter demonstrated the white noise guiding tool to the participants at the beginning of the session, instructing them to step left and right to experience the amplitude changes in each ear.

2.2.2 Experiment 1. The effect of presentation condition on distance estimation using timed imagined walking

Experiment 1 compared five presentation conditions using the timed imagined walking protocol. This measurement protocol was selected because it was suitable for both HMD and LSID presentation conditions. The conditions were as follows:

1. Real: unrestricted real-world view of hallway ($N = 12$)
2. Real+HMD: real-world view of hallway seen through an HMD ($N = 12$)
3. Virtual+HMD: virtual model of hallway viewed in an HMD ($N = 13$)
4. Virtual+LSID: virtual model of hallway viewed on multiple large screens ($N = 15$)
5. Photo+LSID: photo-based presentation of hallway viewed on multiple large screens ($N = 12$)

Specifically, the experiment attempted to answer the following experimental questions :

1. What is the impact of the HMD encumbrance (weight, FOV, etc.) on distance perception? Can HMD encumbrance cause distance compression observed in the virtual environments as shown earlier by Willemsen et al. [114] ?
2. How does the accuracy of distance estimation in HMDs compare to that in LSID systems?

3. Can visual models of higher rendering quality substantially improve distance perception in an LSID system?

2.2.2.1 Participants

We recruited 64 undergraduate and graduate students to participate. The participants received either course credit or monetary compensation. There were 39 males and 25 females. There were five additional participants who completed the task, but were excluded from the analysis due to the difficulties they experienced with the measurement protocol.

2.2.2.2 Measures

The primary measure for the timed imagined walking conditions was *Time To Target*. We used each participant's average walking speed to convert each *Time To Target* measurement into an estimate of *Distance Walked*. We used distance walked to calculate *Judged Percentage of True Distance* for each trial, which was expressed as a ratio between the distance walked and the true distance to the target. *Judged Percentage of True Distance* can be used as a measure of the accuracy of a participant's distance estimates:

$$Judged\ Percentage\ of\ TrueDistance = \frac{Distance\ Walked}{True\ Distance}. \tag{2.1}$$

We assessed the precision of the distance estimates by computing *Variable Error*, which is expressed as a coefficient of variation for distance walked. Specifically, *Variable Error* for a given value of true distance and a given participant is the ratio

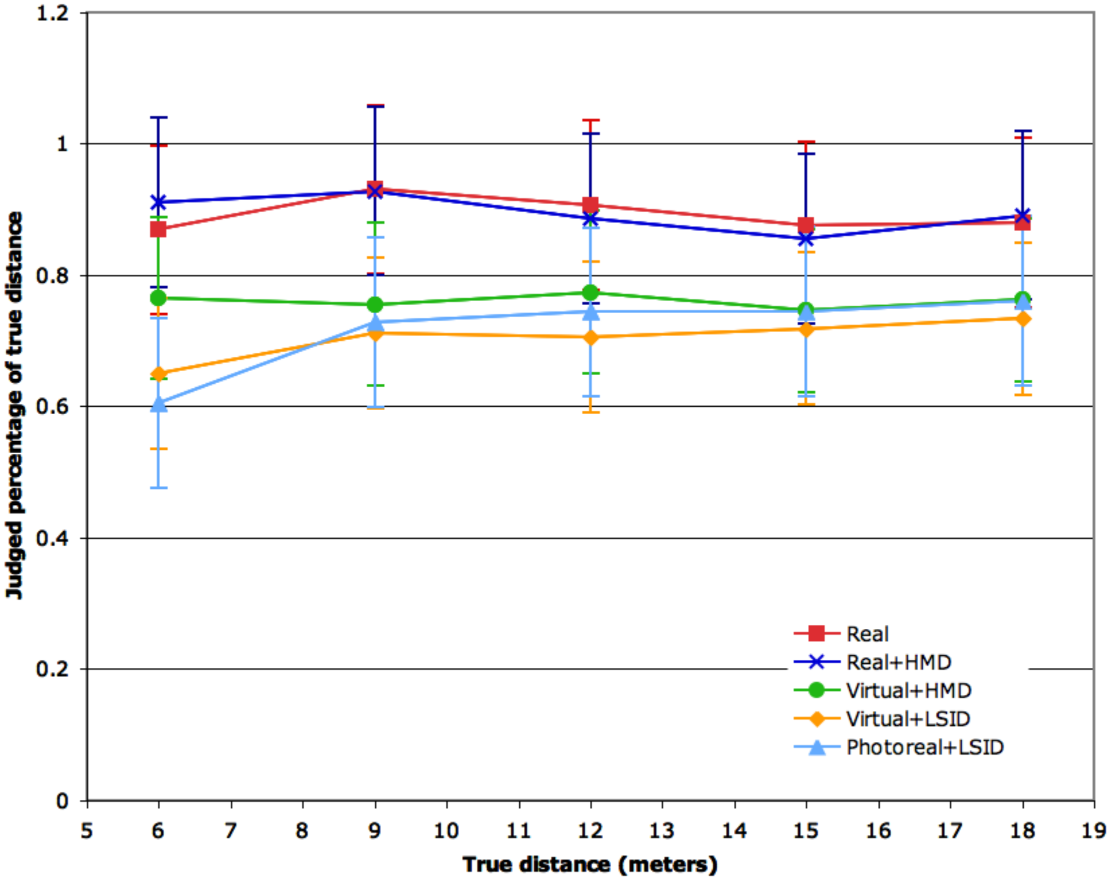

Figure 2.2: Mean percentages of true distance using timed imagined walking. Error bars represent 95% confidence intervals.

of standard deviation of distance walked to the mean distance walked:

$$Variable\ Error = \frac{SD(Dist.\ Walked)}{Mean(Dist.\ Walked)}. \tag{2.2}$$

2.2.2.3 Results

Figure 2.2 shows mean judged percentage of true distances for each observed distance in all experimental conditions. These means were obtained by first finding the mean of the three observations for each participant at a given distance and then by averaging across all participants in each condition. The figure suggests that the performance of the participants remained fairly stable across all five distances, though the Photo+LSID and Virtual+LSID conditions showed a noticeable decrement in performance for the 6m distance.

To simplify the analyses of accuracy, we estimated a single value of expected *Judged Percentage of True Distance* for each participant over the whole range of observed distances. To do this, we simultaneously fit a linear regression line for each participant with true distance to target as predictor and distance walked as a dependent variable. The intercept was fixed at zero. In this model the estimated linear slopes correspond to the expected judged percentages of true distances for each of the participants. This method yields results very similar to finding a simple mean percentage of judged true distance. However, it assumes stronger relationship between observations from the same participant over the range of distances and treats the whole distance range as a continuous interval. Overall, we found that the model fit the data well ($R^2 = 97\%$). Figure 2.3 illustrates the resulting linear fit for each condition to the mean distances walked at each true distance. Table 2.1 shows estimates of the mean judged percentage of true distance for each condition.

We compared the mean accuracy of the participants in each condition using a one-way ANOVA with experimental condition as the predictor and *Expected Judged Percentage of True Distance* as the dependent variable. The overall ANOVA F-test was not significant, $F(4, 59) = 1.927, p = 0.118$. However, Figure 1 suggests that participants in the Photo+LSID, Virtual+LSID and Virtual+HMD conditions

Table 2.1: Means of expected judged percentages of true distances using imagined walking

Condition	**Mean**	**95% Confidence interval**
Real	0.888	[0.767, 1.008]
Real+HMD	0.885	[0.764, 1.006]
Virtual+HMD	0.759	[0.643, 0.875]
Virtual+LSID	0.719	[0.611, 0.827]
Photo+LSID	0.743	[0.623, 0.864]

underestimated true distances more than did participants Real and Real+HMD conditions. Therefore, we grouped participants who observed a virtual world or photographic images of the real world (Virtual+LSID, Virtual+HMD, and Photo+LSID conditions) and compared them to participants who observed the real world (Real and Real+HMD conditions). A one-way ANOVA with Group (Real, Virtual) as a predictor and expected *Judged Percentage of True Distance* as the dependent variable confirmed our hypothesis, $F(1, 62) = 7.775, p = 0.007$.

We also compared mean performance of participants in Real condition with ideal performance (i.e., 100% accuracy in judged distance). A one-sample t-test showed that participants did not significantly underestimate true distances: $t(11) = -1.74, p = 0.11$.

Variable errors were analyzed in a $Condition(5) \times TrueDistance(5)$ repeated measures ANOVA with the first factor as a between-subjects variable and the second as a within-subjects variable. We found a significant main effect of True Dis-

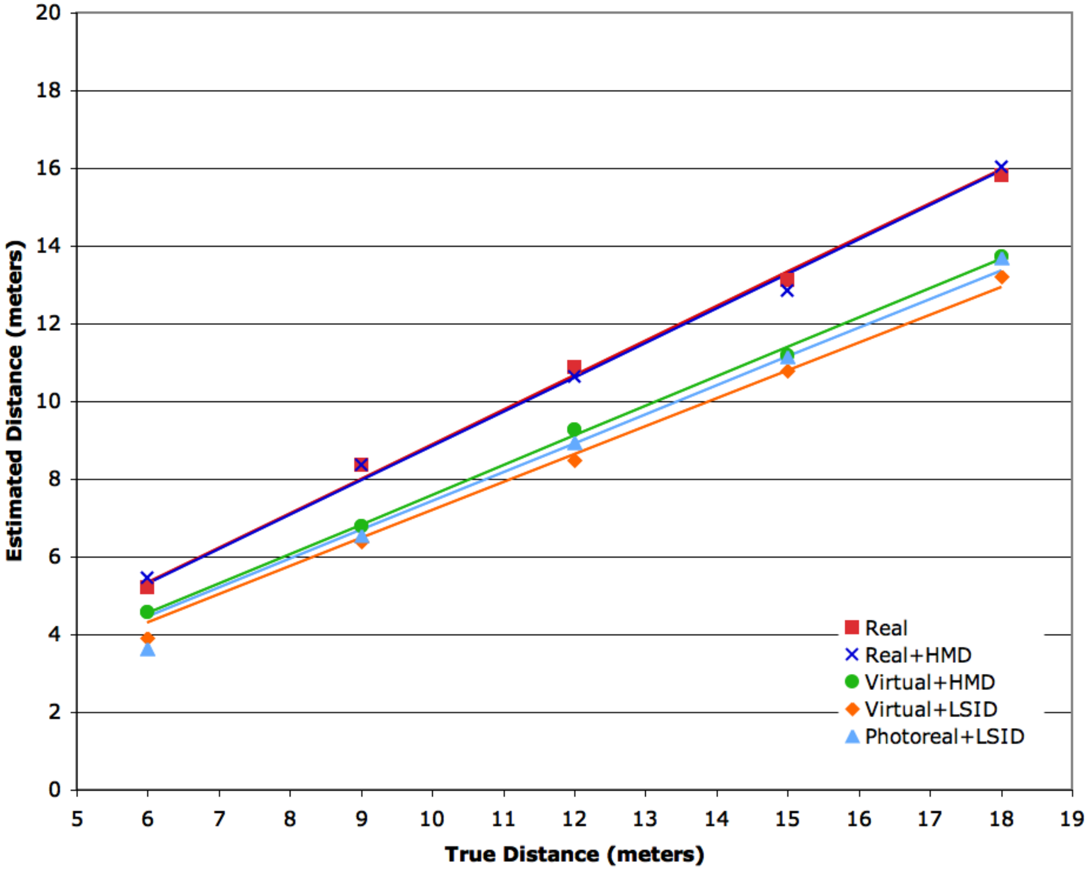

Figure 2.3: Mean accuracy for distance estimation using timed imagined walking. Points represent mean judged distances. The slopes of the lines correspond to expected judged percentages of true distance for each condition.

tance, $F(4, 236) = 4.86, p < 0.001$. The main effect of condition ($p = 0.226$) and the $Condition \times TrueDistance$ interaction ($p = 0.55$) were not significant. Variable errors tended to decrease with distance. To estimate the linear trend over the range of observed distances, we repeated the ANOVA analysis with Distance as a continuous predictor. There was a statistically significant negative trend for the error which tended to decrease on average by 0.5% per meter of true distance, $F(1, 251) = 15.70, p < 0.001$. The overall mean Variable Error was estimated at 16.7% ($SE = 0.6\%$).

2.2.2.4 Discussion

These results indicate no difference in accuracy of distance estimation between the participants in Real and Real+HMD conditions, with means of the judged percentages of true distance being almost identical (88.8% and 88.5% respectively). We conclude that the combination of HMD weight and limited FOV are not likely to cause substantial underestimation of distances as measured by the timed imagined walking protocol.

At the same time, our data shows that participants in the Virtual+HMD, Virtual+LSID, and Photo+LSID conditions exhibited similar levels of distance compression relative to the real world. The similarity across the three conditions is especially notable in light of substantial differences between display systems (HMD vs. LSID) and presentation methods. In particular, the lack of difference between judgments in Virtual+LSID and Photo+LSID conditions implies that the improving the quality of virtual model rendering would not improve distance estimation, at least in non-stereoscopic, non-parallax LSIDs systems.

2.2.3 Experiment 2: The effect of presentation condition on distance judgments using direct blindfolded walking

Experiment 2 compared four experimental conditions using direct blindfolded walking protocol. Compared to the previous experiment we added the AR condition and excluded two conditions with LSID presentation, where the use blindfolded walking protocol was not feasible. The conditions were as follows:

1. Real: unrestricted real-world view of hallway ($N = 11$)
2. Real+HMD: real-world view of hallway seen through HMD ($N = 14$)
3. Virtual+HMD: virtual model of hallway viewed in HMD ($N = 10$)

(6) AR: augmented-reality presentation of virtual target objects superimposed on real hallway seen through HMD ($N = 8$)

Experiment 2 specifically focused on the impact of the HMD on distance perception (Question 1 from experiment 1). Planned comparisons were designed to examine the following:

1. how Real+HMD, Virtual+HMD, and AR conditions compare to the control Real condition
2. a three-way comparison between Real, Real+HMD, and Virtual+HMD conditions.

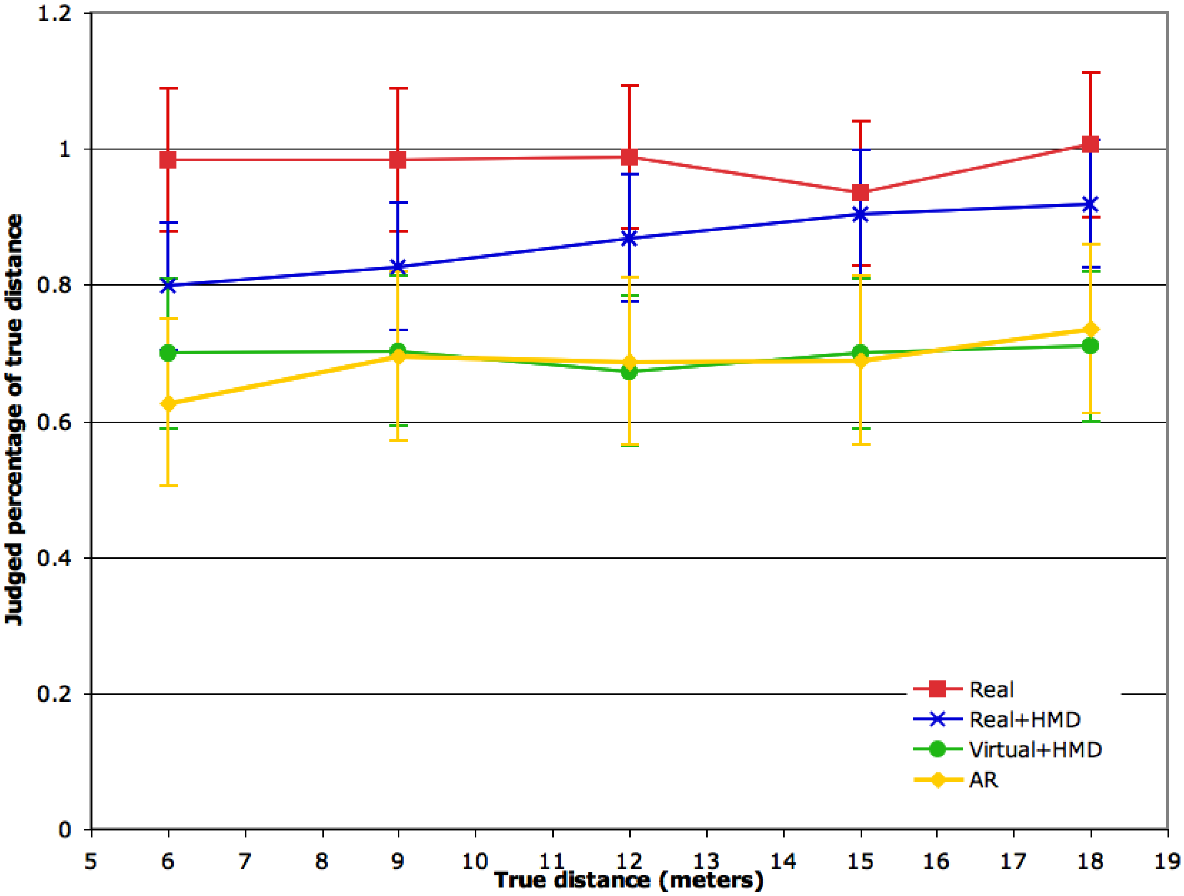

Figure 2.4: Mean percentages of true distance using blindfolded walking. Error bars represent 95% confidence intervals.

2.2.3.1 Participants

We recruited 43 undergraduate students to participate for course credit. There were 21 males and 22 females. Two additional participants in the AR condition completed the task, but were excluded from the analysis due to apparent difficulties with the blindfolded walking protocol.

Table 2.2: Means of expected judged percentages of true distances using blindfolded walking

Condition	Mean	95% Confidence interval
Real	0.979	[0.875, 1.084]
Real+HMD	0.891 0.825 [a]	[0.798, 0.98] [0.741, 0.909]
Virtual+HMD	0.699	[0.589, 0.809]
AR	0.706	[0.583, 0.829]

[a] Excluding two overestimating participants

2.2.3.2 Measures

We used *Distance Walked* as a primary measure for on each trial and then computed *Judged Percentage of True Distance* as discussed earlier.

2.2.3.3 Results

Figure 2.4 shows mean *Judged Percentage of True Distance* for each observed distance in all experimental conditions. These means were again obtained by first finding the mean of the three observations for each participant at a given distance and then by averaging across all participants in each condition.

As in Experiment 1, we estimated a single value of expected *Judged Percentage of True Distance* for each participant by fitting a linear regression line with true distance to target as the predictor and distance walked as the dependent variable. The intercept was fixed at zero. The model fit the data well ($R^2 = 97\%$). Figure 2.5 illustrates the resulting linear fit to the mean distances walked at each true distance

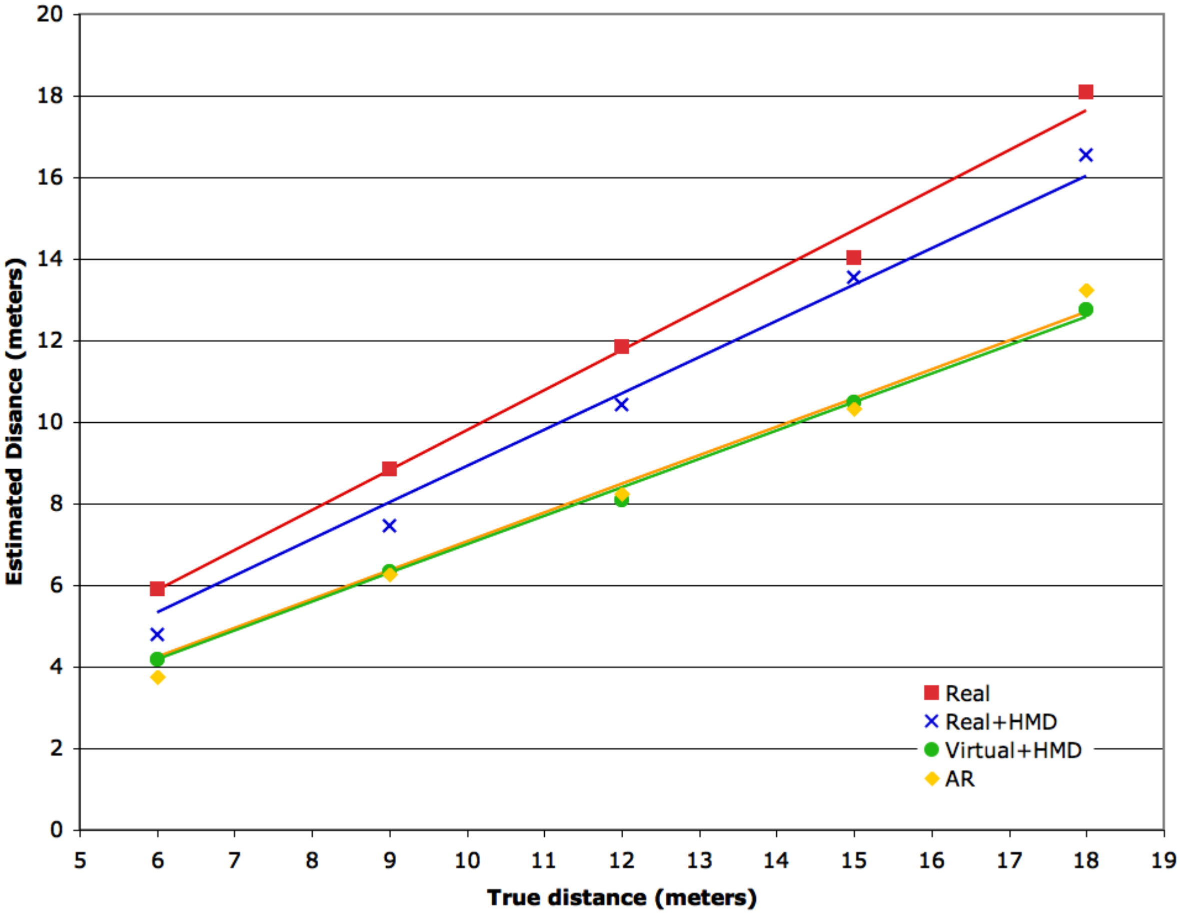

Figure 2.5: Mean accuracy for distance estimation using blindfolded walking. Points represent mean judged distances. The slopes of the lines correspond to expected judged percentages of true distance for each condition.

for each experimental condition. Table 2.2 shows estimates of the mean judged percentage of true distance for each condition. See also Figure 2.6 that summarizes the mean judged percentages of true distances for both experiments.

A one-way ANOVA with experimental Condition (4) as the predictor and expected *Judged Percentage of True Distance* as the dependent variable compared the mean accuracy of the participants in each condition. We found a significant effect of experimental condition, $F(3, 39) = 6.68, p < 0.001$.

Planned comparisons between conditions were carried out using a Bonferroni-Holm adjustment. We found significant underestimation relative to Real condition in the Virtual+HMD ($p = 0.002$) and AR ($p = 0.004$) conditions. There was significant underestimation in the Virtual+HMD condition relative to Real+HMD ($p = 0.02$). The difference between Real and Real+HMD conditions ($p = 0.21$) was not significant.

We observed that performance in Real+HMD condition was substantially influenced by two participants, which on average overestimated true distances by 38.9% and 18.5% respectively. A classical Grubb's test suggested that both these values were potential outliers ($p = 0.013$). We repeated the analysis without the data from these two participants. The planned comparisons showed significant underestimation of distances relative to Real condition in the AR ($p < 0.001$), Virtual+HMD ($p < 0.001$), and Real+HMD ($p = 0.023$) conditions. The difference between the Real+HMD and Virtual+HMD conditions was also significant ($p = 0.048$).

We also compared mean performance of participants in Real condition with

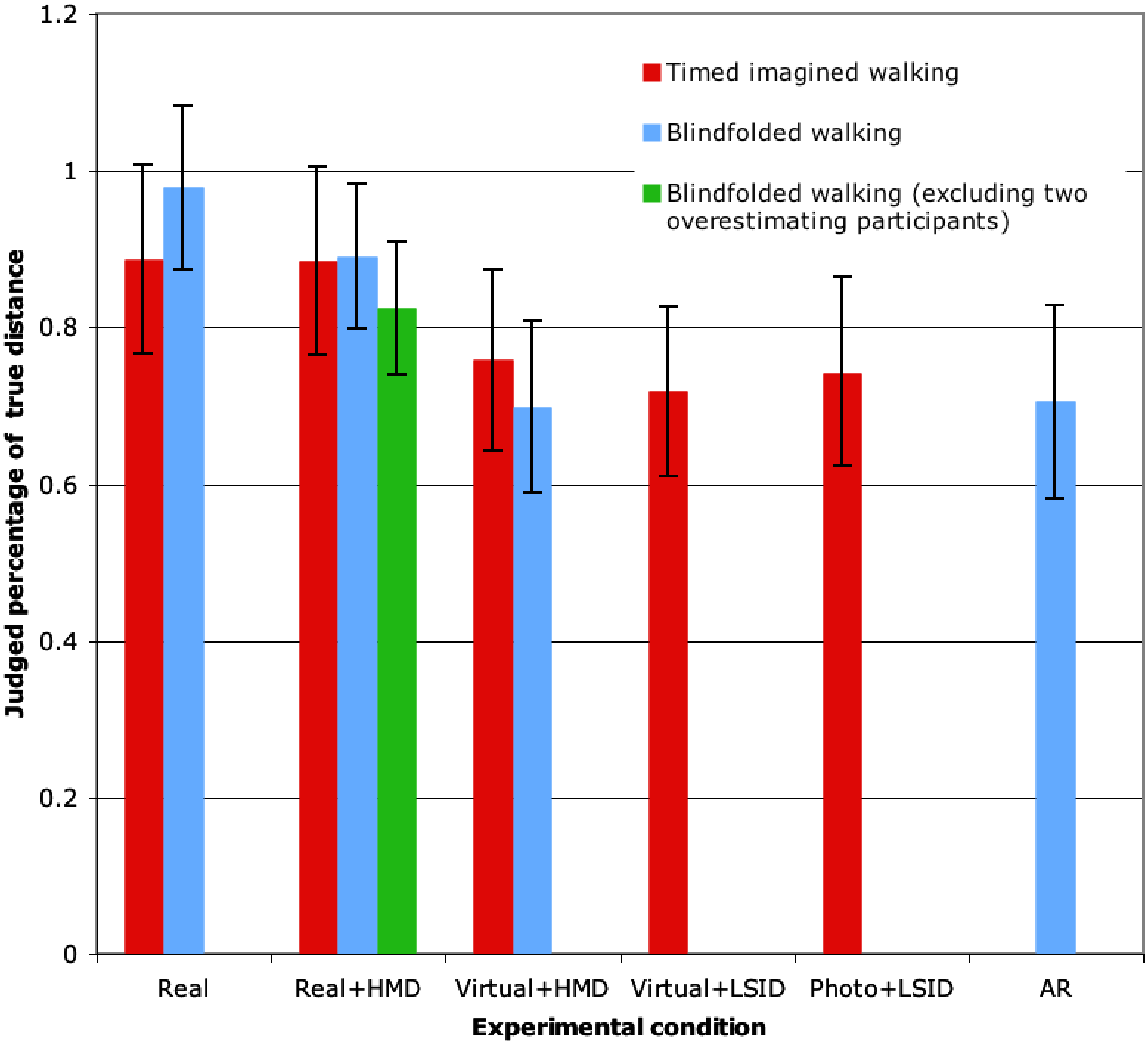

Figure 2.6: Mean accuracy for distance estimation using blindfolded walking and timed imagined walking. Error bars represent 95% confidence intervals.

ideal performance (i. e. 100% accuracy in judged distance). A one-sample t-test showed that participants did not underestimate true distances: $t(10) = -0.41, p = 0.69$.

Variable errors were analyzed in a Condition (4) x Distance (5) repeated measures ANOVA with the first factor as a between-subjects variable and the second as a within-subjects variable. Neither of the main effects were significant, $F(3, 39) = 1.88$, $p = 0.149$, and $F(4, 156) = 1.57$, $p = 0.185$, respectively. However, the linear trend for Variable Error to increase with distance was significant, $F(1, 168) = 4.69$, $p = 0.032$. On average, the error tended to increase by 0.7% per meter of true distance. The overall mean variable error was 12.2% ($SE = 0.6\%$).

2.2.3.4 Discusison

Together, our analyses of accuracy indicate that participants in both Real+HMD and Virtual+HMD conditions showed significant underestimation of distances. However, the significant distance compression in Virtual+HMD condition relative to Real+HMD condition indicates that the HMD encumbrance alone cannot account for the degree of distance underestimation observed in the virtual environment. Our initial analysis suffered somewhat due to the large variability in performance of individual participants in Real+HMD conditions, which made the differences between Real and Real+HMD condition less apparent. We were able to clarify this difference, by excluding two participants with somewhat unusual level of overestimation of distances.

We also found that the performance of participants in AR condition, who observed virtual targets in the real visual settings, was similar to that of participants in Virtual+HMD condition, who observed virtual targets in the virtual visual settings. However, it is difficult to draw conclusions based on a relatively small sample of AR participants in our analysis.

2.2.4 Effect of measurement protocol

The impact of the measurement protocol on distance perception can be examined by comparing the performance of participants in Real, Real+HMD, and Virtual+HMD conditions between the two experiments. Figure 2.7 shows the mean expected judged percentages of true distances for both direct blindfolded walking and timed imagined walking in each of the three conditions. The mean value for Real+HMD condition for direct blindfolded walking does not include two overestimating participants that were identified in the earlier analysis.

We compared the performance across the measurement protocols using a $Protocol(2) \times Condition(3)$ two-way ANOVA. We found that both the main effect of the $Protocol$ ($F(1, 64) = 0.043$, $p = 0.837$) and interaction between $Protocol$ and $Condition$ ($F(2, 64) = 1.083$, $p = 0.345$) were not significant. The main effect of condition was significant ($F(2, 64) = 5.63$, $p = 0.006$). These results suggest that participants performed similarly in each of the three conditions whether they used the imagined walking or direct blindfolded walking. However, the timed imagined walking protocol yielded somewhat less accurate estimates in the Real condition.

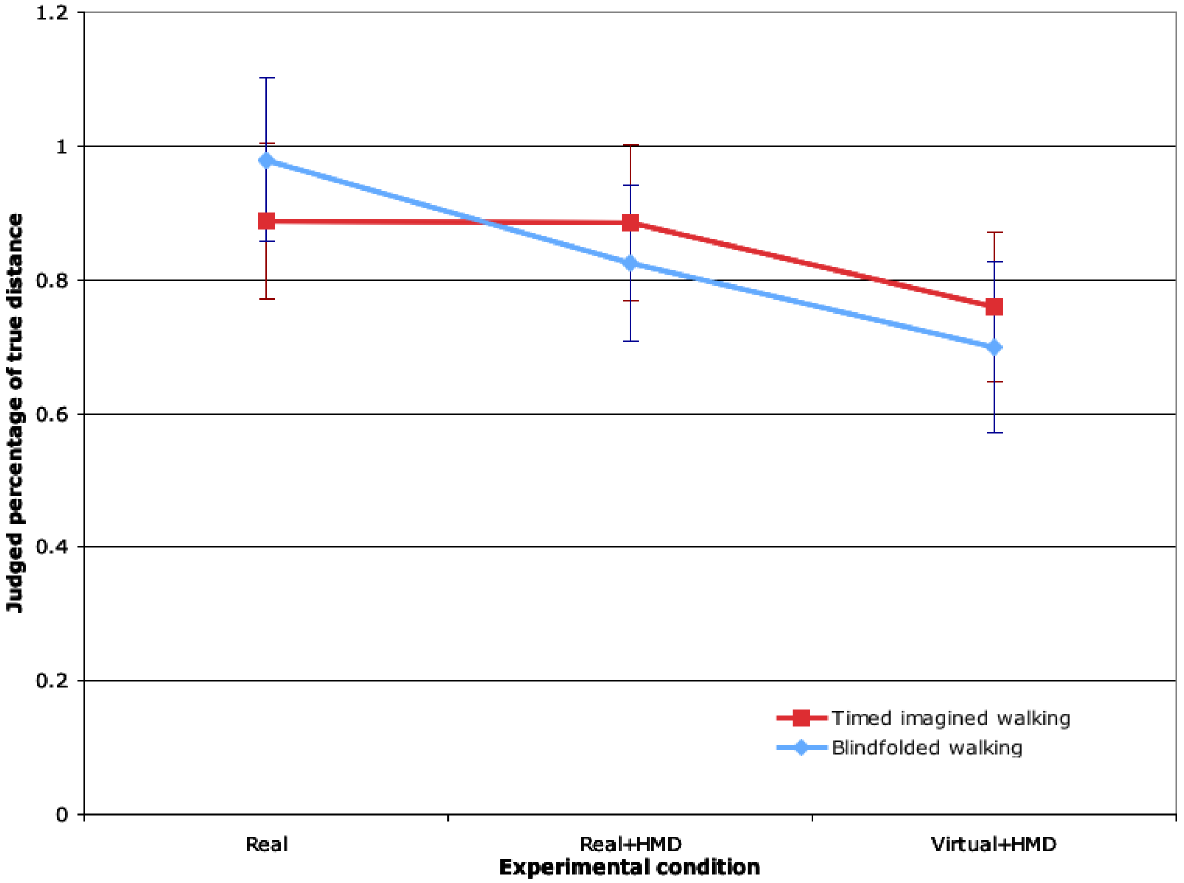

Figure 2.7: Comparison of measurement protocols: Mean expected judged percentages of true distances by experimental condition. Error bars correspond to 95% confidence intervals.

The most important differences were in the relative performance of participants across the three conditions within each protocol. Participants significantly underestimated distance with both measurement methods in the Virtual+HMD condition relative to real world estimates. However, the difference between Real+HMD and Real conditions was only found with the blindfolded walking protocol. When participants viewed the targets in the real world through the HMD but imagined moving to the targets while standing in place, there was no difference between the Real and the Real+HMD conditions. When participants viewed the targets through the HMD and then actually walked to the targets, they underestimated distance in the Real+HMD condition relative to the real condition. Thus, the effect of the HMD encumbrance only influenced distance estimates when the participants were required to physically walk to the target. This may be related to a greater effect of the tipping torque or pull from the cables when walking as opposed to when standing still. Thus, it appears that the effect of wearing an HMD while viewing the real environment depends on the measurement protocol used to study distance estimation.

A potential limitation to the above direct comparison between the two protocols is associated with the potential presence of a systematic bias in the distance estimates produced by timed imagined walking protocol due to conversion of a directly measured value of time into an indirect measure of distance. The estimate of participant's speed required for this conversion was obtained by timing participant's sighted walking over a fixed distance. However, the data collected by Kunz, Creem-Regehr and Thomson [51] indicates that timed imagined walking produces systematically shorter estimates of time required to reach a target than the actual sighted walking. One explanation for this phenomenon is that the implied speed for the timed imagined walking is higher than the actual walking speed. Consequently our method for converting time to distance would lead to systematically shorter implied distances for timed imagined walking protocol. At the same time, such a bias would not affect our key conclusions based on comparisons between experimental conditions, which were carried out within experimental protocol.

Overall, the results of this study demonstrate that the choice of a particular measurement protocol may affect the outcome of the experiment. Because the choice of the protocol can prove to be a potential confound, perception and action should be considered together, in a holistic approach.

2.3 Additional evidence in support of holistic approach to perception and action

A recent study by Ziemer et al. [119] provides further evidence that the choice of the measurement protocol may affect the outcome of a perceptual experiment. They studied the effect of re-calibration that participants experience when the familiar links between perception and action are somehow altered. As mentioned earlier, people continuously calibrate the relationship between perception and action based on the feedback they receive. If participants experience an environment, where, for example, their walking speed remains the same but the optic flow is faster than normal, they might adjust their estimation of how far they need to walk in order to reach objects in the environment.

In this study participants were asked to perform two sets of distance estimates in the real world: pretest and post-test. Participants reported perceived distances using either blindfolded waking or timed imagined walking protocol. Between the two sets of estimates participants walked for a fixed amount of time on a treadmill through a virtual environment where either the walking speed or the speed of the optic flow was manipulated.

Experiments using blindfolded walking showed that participants adjusted their estimates when either speed or optic flow were manipulated in the virtual environment. Experiments using timed imagined walking showed the recalibration effect when the optic flow was manipulated, but not in the case when the walking speed was manipulated.

These results indicate that different kinds of alterations between real and virtual environment did not affect different kinds of response actions in the same manner. Futhermore, these results illustrate that relative calibration between perception and action undergoes continuous change based on person's experience. The effect of the prior experience on perception can be somewhat unexpected.

An earlier study by Ziemer and colleagues [120] compares distance estimates in real and virtual environments the two familiar measurement protocols: timed imagined walking and blindfolded walking. In this study participants performed two sets of distance estimation tasks. Some participants performed the first and the second sets of distance estimates in the same type of environment (i.e real - real or virtual -

virtual), while others changed the type of environment for the second set of estimates.

The results show that the first set of estimates was more accurate in real environment compared to the virtual environment. The second estimates performed in the same environment as the first estimates were also more accurate in the real world compared to the virtual environment. However, the second real world estimates performed by participants who originally estimated distances in the virtual environment were not significantly different from the second estimates in the virtual environment performed by participants who originally estimated distances in the real world.

This suggests that perception-action calibration obtained as a result of a prior experience even without pronounced manipulation of perceptual information can be carried over to a different environment and affect the results of the distance estimation task.

2.4 Conclusions

As we have seen in this chapter perception-action coupling needs to addressed in validation of virtual reality systems, even when the validation is primarily directed toward perceptual measures. The intrinsically inseparable nature of perception and action makes it likely that the results of perceptual evaluation might be affected by the artifacts of sensitivity or insensitivity of a given measurement protocol to some aspects of the evaluated VR system.

In a broader context, the evidence discussed here argues for a wider use of evaluations that treat perception-action systems holistically. This next chapter will focus on formulating a theoretical framework that incorporates both perception and action as components of VR evaluation. In particular, it will discuss the notion of affordances, its use in ecological psychology and HCI, and its application in context of computer mediated environments.

CHAPTER 3
DEVELOPING THEORETICAL FRAMEWORK FOR STUDIES OF PERCEPTION AND ACTION IN VIRTUAL REALITY SYSTEMS

The key aim of this dissertation is a better understanding of full-body, dynamic interactions with a computer-mediated virtual world, which users could treat in much the same way they treat their physical interaction with the real world. In this it follows a grand vision for the virtual reality research offered by Sutherland in his pivotal paper [96]. Comparing digital virtual world to a “mathematical wonderland”, which we observe through “a looking glass” of a computer display, he describes the quintessential version of “the ultimate display”:

> The ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming such a display could literally be the Wonderland into which Alice walked.

In other words, the goal of VR research is to achieve such level of user immersion into the virtual world that from the user’s point of view the virtual world becomes indistinguishable from the real one. As the user achieves a complete sense of being in the the virtual world, the computer interface disappears completely.

The present day VR systems are still far from fulfilling this vision (Figure 3.1). Despite our best efforts to make the computer interface between the user and the virtual world nonintrusive and invisible, it still affects both user’s perception and action capabilities. We have to face the fact that both perception and action in a VR

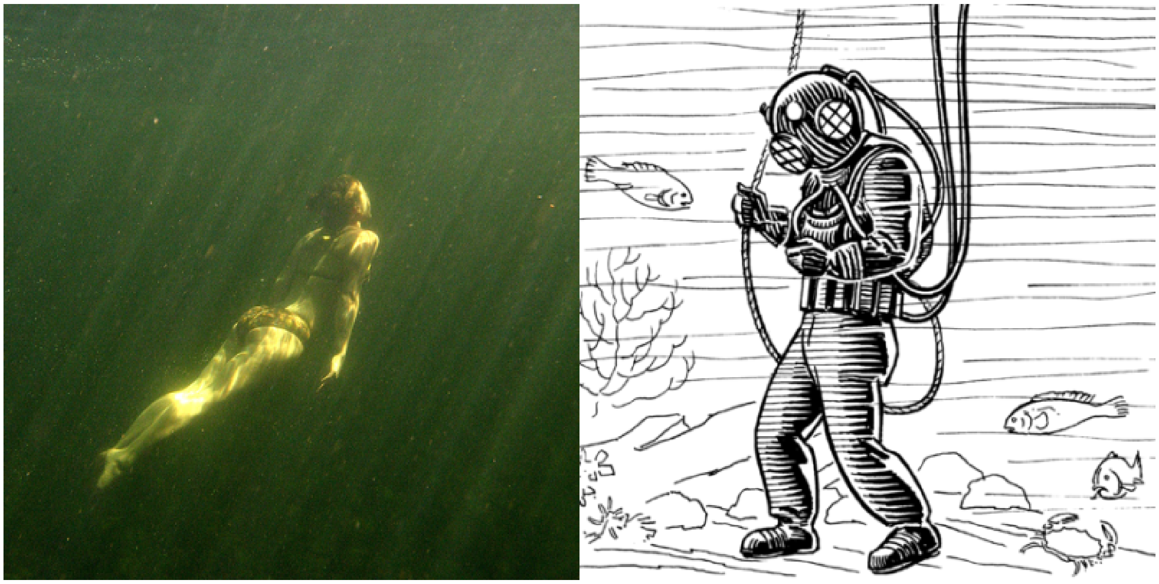

Figure 3.1: Metaphorical representation of VR's vision of unencumbered full immersion into the virtual world (left) and its practical implementation(right). Images courtesy of Gerry Thomasen (under Creative Commons, Attribution 2.0 license via Flickr at http://flic.kr/p/gysVt) and Pearson Scott Foresman (via Wikimedia Commons at http://tinyurl.com/86vmnn5).

system are computer-mediated and can be potentially perturbed by the mediating technology.

The previous chapter established that perception and action are so deeply intertwined, that it is impossible to examine one without the other. The next step is to develop a theoretical framework that relates perception and action in a computer-mediated environment.

One promising candidate for this purpose is the psychological concept of "affordance", which also has a wide-spread use in HCI community. The next two sections of this chapter will provide a detailed account of the concept of affordance in ecological psychology and HCI.

3.1 Theory of affordances in ecological psychology

The notion of affordances was originally introduced by an ecological psychologist James Gibson [31] to describe the opportunities for action that an environment presents to an organism. The concept can be applied very broadly to many different aspects of the environment. For example, with respect to a human actor the terrestrial plane might afford walking, a staircase might afford climbing, and an object in the environment (such as a rock) might afford throwing.

As an evolving concept, the notion of affordances generated a significant debate in the literature [13, 46, 49, 61, 91, 92, 103] with several authors proposing various refinements to the original definition. Despite the fluid nature of the concept, here we will follow the general approach by Fajen et al.[27] and formulate a list of key features that define the notion of affordances:

- Affordances are real and can be associated with invariant physical properties of both the organism and the environment. For example, Warren [108] demonstrated the relationship between a perceived ability to climb a step of stairs and a bio-mechanical model of climbing that related the length of the actor's leg and the height of the step.
- Affordances exist in an organism-environment system whether or not they are being perceived. Organisms, however, may perceive existing affordances accu-

rately or inaccurately by misjudging possibilities for action,

- Affordances are organism-specific. The actions afforded by the environment differ from organism to organism. For example, an object on top of a bookshelf might be simultaneously reachable for a tall person, but out-of-reach for a short one.
- Affordances capture the structural relationship between perception and action. By focusing on what an organism can do, they connect perception of the environment and action capabilities of the organism.
- Affordances are dynamic. In dynamically changing environments the opportunities for action appear, evolve, and disappear over time.
- Affordances allow organisms to prospectively control their behavior [103]. In other words, the organisms can act taking into account the projected future state of the environment, as opposed to simply reacting to its present state.

Two categories of affordances that are studied most often include body-scaled and action-scaled affordances [27]. While the boundary between body-scaled and action-scaled affordances is imprecise, the distinction is generally useful.

Body-scaled affordances represent opportunities for actions that can be primarily expressed in terms of body dimensions. For example, as mentioned earlier the climability of a step of stairs can be expressed in terms of relative height of the step and the leg of the actor [108]. Similarly, the relation between the length of the leg and the height of a seat define sitting affordance [93] (in other words, a chair affords sitting only if its height does not exceed maximum height for a person with a given leg length).

Action-scaled affordances, on the other hand, are expressed primarily relative to the action capabilities of the organism. For example, stopping within a certain distance depends on the maximum deceleration rate within the actor's capabilities [27]. Fajen [25] argues that any arbitrary measurement units are not meaningful in the context of perception and therefore action-scaled affordances have to be expressed in terms of perceived limits on action capabilities (hence the term action-scaled affordances).

Despite some theoretical objections (see Michaels [61]), the concept of affordances has generally been extended to perceiving possibilities for actions of others. Studies show that humans can accurately perceive affordances for other humans. For instance, Stoffregen et al.[93] demonstrated that people can perceive maximum sitting height not only for themselves, but also for the actors they observe.

Humans also detect interpersonal affordances corresponding to the possibilities for joint actions. Richardson et al.[82] investigated the affordances for grasping planks of wood with varying lengths with either one hand, two hands, or jointly with another person. Their results suggest that the type of action selected by participants was associated with the action-scaled ratio that linked the length of the plank to the participant's hand span. The transition between one-hand grasp, two-hand grasp, and two-person grasp action was associated with increasing action-scaled ratio and occurred in variations of the task that involved both judgment-only response and actual grasping action.

Perceptual calibration and experience seem to play an important role in accurate perception of affordances. People seem to be able to rapidly adjust to changes in action-scaled affordances occurring due to changes in action capabilities even with limited feedback. Fajen [26], reports on experiments that involve simulated driving with abrupt breaking, where participants rapidly adjusted how they break after their action capabilities were altered even though the screen turned black shortly after the breaking started and participants were not able to see the outcome of the action.

The perception of affordance for others seems to depend on both the available perceptual information and the level of experience associated with particular action. Weat et al.[109] investigated the role of experience and kinematic information in perceiving affordances for self and others by experts (basketball-players) and non-experts. They found that experienced basketball players were significantly more accurate compared to non-players in judging baskteball-specific, action-scaled reach-with-jump affordances (i.e. the maximum height of a plank that can be reached vertically overhead with a jump) for a model, who did not move. There was no significant difference in judging self-affordances. There were also no significant differences in judging body-scaled affordances not specifically associated with basketball skills for self and for the model. This was true for both standing-reach affordance (i.e. maximum height that can be reached overhead from standing position) and sit affordance (i.e. maximum height of a plank upon which one can sit without lifting heels of the feet from the floor, similar to the affordance discussed in [93]). The researchers also found that exposure to kinematic information about the model by watching the model walk around significantly improved accuracy of judging action-scaled affordance (reach-with-jump) for basketball players, but not for non-players. By contrast, there was no effect of exposure to kinematic information on judging standing-reach affordance and an unexplained degradation of accuracy for judging sitting affordance.

These findings are consistent with the idea that accurate perception of affordances depends on ongoing adjustment to detected changes in action and perception capabilities [27]. This calibration of affordance perception allows organisms to adapt and enables affordance-based control of visually guided actions [25].

3.2 Notion of affordances in human-computer interaction

The introduction of the notion of affordances to HCI community is commonly credited to Donald Norman [66, 68]. However, in Norman's interpretation affordances are no longer defined as relations between an organism and an environment, but instead become properties of an object that can signal what actions can be applied to it. Perceived affordances can depend on internal mental models, prior knowledge, experience, and cultural conventions. Furthermore, affordances can describe possibilities for action that are perceived by the actors, but may or may not be possible. He provides a famous example describing the door with vertical handles to signal that it can be pulled, as opposed to a horizontal bar handle, that signals that the door should be pushed. The actual mode of operation for the door may or may not match the perceived affordances. To distinguish "true" affordances from the possibilities for action that appear to be possible, but do not actually exist, Norman later suggested to re-label his original concept as ”perceived affordances” [67].

The underlying reasons for the divergence between Norman’s and Gibson’s notions of affordance appear to be two-fold.

First, Norman (together with many HCI researchers) and Gibson fall on the opposite sides of a broader philosophical debate in behavioral science. HCI generally follows the traditional, representation-based approach to perception, where as Gibson was a proponent of a rival, information-based approach of direct perception. From the representation-based point of view, perception is considered to be a part of an information processing activity. An organism’s preceptors gather information in order to create internal models (representations) of the environment. Cognitive processing then combines these models with other representations (knowledge-based models of environment dynamics, body dimensions, and dynamics, etc.) to produce an action. In the Gibsonian information-based approach, on the other hand, perception is direct and is tightly coupled with action in the form of an affordance. The affordance-relevant information is gathered directly from the environment. For example, when catching a flying ball, the visual information about ball’s movement (such as the optical rate of expansion) is translated directly into running acceleration needed to intercept the ball. Because the affordances for action are perceived directly, acting on an affordance does not involve cognitive processing (see Fajen[25] for more details).

Second, Gibson formulated his theory primarily in terms of understanding physical actions that naturally emerge towards existing natural objects in the real world. Norman, on the other hand, approaches the concept from a point of view of a designer, who is concerned with creating new objects that have to fulfill some purpose. This approach leads to a need for communicating the intended use of the newly designed object to the user, for making possible actions“visible” through some properties that are embedded in the object itself.

Norman’s approach proved to be attractive for designers of graphical user interfaces (GUI). In the traditional GUI paradigm, users interact with the digital world indirectly, through generic physical input devices such as keyboard and mouse and perceive the digital world as an abstract, external, and intangible representation through a computer display screen and sound system. In an attempt to describe various features of the interface that help the user understand its intended purpose, some researchers in the HCI community proposed to extend Norman’s use of the term “affordance” to describe a number of aspects of the visual interface that are cognitively processed by the user [37]. As a result of this evolution in HCI literature the term “affordance” was they inadvertently misappropriated from its intended meaning in ecological psychology. Moreover, the inconsistent use of its various meanings created a great deal of confusion within the HCI community itself [59, 69, 70, 102], which gives a lot of strength to the argument in favor of reserving the use of the term for its original meaning [102, 70].

The emergence of tangible and ubiquitous computing approaches that place embodied interaction paradigm [21] in the center of interface design make the case for Gibsonian understanding of affordances even stronger.

Ubiquitous computing [1, 110] aims for more natural interaction with digital world by augmenting real environment and people with computational devices. To make the interface with computing systems disappear, these devices strive to maximally leverage the natural methods of interaction that have evolved over millennia of humans interacting with their environment. In particular, they treat human bodies as an integral part of the "interface" [29]. In some sense, ubiquitous computing tries to understand the affordances that exist between humans and the real world and the gradually introduce new possibilities for actions by closely mimicking the existing ones.

Tangible interfaces [89, 104] attempt to create real-world representations for digital objects. In this paradigm the user interacts with the virtual world by manipulating physical objects in the real world. Here again the idea is to leverage the user's knowledge of actions possible with the real world objects to make computer interactions more intuitive.

Jacob at al. [44] build on these commonalities to group virtual reality, augmented reality, mixed reality, tangible computing, and ubiquitous computing under the umbrella term "reality-based" interaction. They identify four themes that connect these interfaces with "reality" interactions. These include leveraging naive understanding of real-world physics, body awareness and skills, environment awareness and associated skills, and, finally social interactions.

Its worth noting that for all these system the notion of affordance, which is close to the original psychological view seems to be most useful.

3.3 Applying the concept of affordance to

mediated interactions

Two previous sections outlined the concept of "affordance" and its evolution in ecological psychology and HCI. Despite some differences in interpretation the term proved useful in both areas of research as concept that links perception and action, actor, and the environment. There remains, however, an important question. How does the concept of affordance apply to environments where perception and action (or both) are mediated through technological interface? This section attempts to answer this question by considering how people adjust to changes in their action capabilities and how new action possibilities arise from using objects in the environment to supplement the actor's body.

Most of the actions that humans perform are embodied (i. e. performed through the use of the body) and embedded into the environment. As mentioned earlier, the affordances that the environment offers to an actor depend on the latter capabilities for action such as body dimensions, strength, skills, etc. Over time, people establish certain statistically supported expectations about connections between their perceptual and action abilities that help to perceive affordances accurately. Thus they establish perception-action systems (or couplings) to perform certain types of action.

The perception-action systems are not static. People effectively calibrate perception-action system to accurately perceive changes in affordances associated with temporary changes in body-related abilities. For example, participants report perceived changes in the ability to climb a slope of the hill, when they are wearing heavy backpack or are fatigued [4]. The alterations in perceptual information have similar effects. Without practice it is hard to control an action while looking at a mirror, which alters the familiar frame of reference by visually switching left and right sides. But most people learn to use mirrors to visually control brushing of their teeth or combing their hair.

Body-related abilities also undergo more permanent changes associated first with development and then with aging. Furthermore, people are cable of learning new types of actions, which requires establishing of new perception-action couplings. These kinds of changes are particularly apparent in infants, who deal with rapidly changing abilities and learn new types of interactions with the real world on a everyday basis (see Adolph [2]).

Often new possibilities for action arise from using am intermediate object to augment the action capabilities of the body. Dourish [21] discusses an example of using a stick to push a rock. While the actor is holding and acting upon a stick, the interaction really occurs with the rock. Furthermore, the stick opens new possibilities for action as some rocks might not be reachable without it. Thus, new affordances can be perceived by coupling the intermediate object with the object to which the action is directed. Conceptually the process of learning this type of action is similar to the process of establishing a new perception-action system for any other type of action.

In computer-mediated environments the interaction occurs with objects in the digital world by coupling these objects with the physical computer interface accessible to the user. Describing affordances associated with technology-mediated environments Gaver refereed to actions that are afforded by a chain of intermediate objects (real or digital) as “nested affordances” [30], because the actions directed towards an object in such a chain are incapsulated (or nested) in actions afforded by the preceding object.

Affordances arising in computer-mediated environments may differ substantially from the familiar affordances associated with the real world interactions. The key problem for the development of user interfaces is understanding what types of action the system will afford to the user and in making sure that these affordances are easily and accurately percieved.

The careful design of the system that leverages familiar perception-action links may be beneficial in making the new interfaces truly useful. Consider, for example, design and implementation of a new laparoscopy system discussed by Voorhorst et al. [107]. The system improves on previous methods for non-invasive surgery is two important ways:

- provides an egocentric frame of reference for visual feedback and enables control of visual viewpoint by head movement, thus supporting more intuitive visual exploration during the operation,
- allows the surgeon to control the apparatus directly, where as previously it had to be controlled indirectly through an assistant.

Most of the new types of interfaces discussed earlier similarly leverage perception-action systems that exist in real world to make computer interfaces more intuitive and transparent for the user.

3.4 Implications for the current research agenda

Virtual Reality systems follow the resent trends in HCI research in attempts to make computer interfaces more intuitive and transparent. The key difference, however, is VR's goal of creating a fully immersive experience. Instead of taking bits of digital into the real world (as do tangible interfaces [43]), VR systems attempt to take the user into the virtual world by fully immersing her into the interactive experience. With the new understanding of the role of affordance as a link between the user and the simulated environment the goals of this research can be now restated in terms of affordances.

The goal of this dissertation is to establish method for comparing the immersive VR interfaces, which uses the perception of affordances by the user as a criterion, and to use this comparison as a way to understand practical implications of using various forms of VR technology. In particular, the aim is to explore the effects of different VR systems on full-body dynamic interactions that involve coordination of motion with objects in the virtual enviormments. Because VR systems are complex, as a first step, this dissertation looks to begin exploring the effect of display type and locomotion modality.

CHAPTER 4
EXPLORING PERCEPTION OF DYNAMIC AFFORDANCES IN VIRTUAL REALITY: A CASE STUDY

4.1 Motivation

Construct an experiment comparing different VR systems on the basis of the theoretical framework developed in the last chapter is a non-trivial problem. Here we will consider a case study that highlights some helpful solutions. Specifically, the following sections present a psychological study that examined how children and adults bicycle across to lanes of opposing traffic in a virtual reality simulator. Several components of the investigation discussed are of interest:

- It is an example of a study that investigates perception of dynamic affordances in a computer-mediated, virtual reality environment.
- It uses a realistic scenario involving a complex perceptual-motor task that involves full-body interactions with the environment. It is particularly important that the task is reminiscent of common every-day activities, because it allows participants to applied well-practiced actions to solve the problem they are facing.
- The study that compares participants in three different age groups, who may differ in their perceptual and motor abilities. This developmental aspect can serve as a helpful proxy for comparing variations in the VR systems affecting participants' perception and action.

This description of the investigation presented here is a modified, shortened version of the work that is originally published as [33].

4.2 Introduction: road-crossing as a perceptual-motor task

Real-world pedestrians and cyclists frequently cross roads with multiple lanes of traffic, often coming from opposing directions. Integrating information about multiple opposing streams of traffic is a challenging perceptual-motor task, and may be especially difficult for children. Children and adolescents between the ages of 5 and 15 are overrepresented in the bicycle crash data, having the highest rate of injury per million cycling trips [83]. The acute complexity of the task is particularly relevant on busy roads, where pedestrians and cyclists are more likely to take small gaps rather than wait for larger gaps [34, 78].

Our approach to understanding the processes underlying road-crossing behavior starts with considering road crossing as a perception-action task with two main components. The first is to determine if a given gap in traffic affords safe crossing. The second component is to coordinate movement through the selected gap without colliding with any vehicles. These two components are intertwined: gap choices constrain crossing actions, and action capabilities constrain gap choices.

4.2.1 Gap selection

For a single lane of traffic a gap affords crossing if the cyclists (projected) crossing time is less than the temporal size of the gap [54]:

$$(T_{Tail} - T_{Lead}) > \frac{d}{v}. \tag{4.1}$$

Here T_{Lead} and T_{Tail} correspond to the arrival times of the first and second vehicles to the planned crossing line, d is the distance to be traversed, and v is the riders average speed. To cross safely, cyclists must accurately judge both the size of the temporal gap and the amount of crossing time required.

Judging gap affordances when crossing multiple lanes of opposing cross-traffic is a considerably more challenging task than crossing a single lane. The rider must select a pair of gaps composed of a near lane gap and a far lane gap that in combination afford safe crossing. This means that each gap in the selected pair must individually be sufficiently large to allow safe crossing of the corresponding lane. Furthermore, because the rider must be able to move from the near lane into the far lane of traffic while both gaps are open simultaneously, the spatio-temporal relationship between the gaps in the pair is critical in determining whether a safe crossing is possible. Since gaps approach from opposite directions they cannot be simultaneously observed. This greatly complicates the estimation of how the near and far lane gaps will overlap, because the riders have to integrate visually available and remembered information about the estimated arrival times of the two gaps to judge whether a pair of gaps affords safe crossing.

Using notation similar to Lee's, a gap pair is crossable when the temporal sizes of the near and far lane gaps are larger than the projected crossing times for the corresponding lanes and the overlap is large enough for the rider to move from the near lane into the far lane:

$$\begin{cases} (T_{NearTail} - T_{NearLead}) > \frac{d_{near}}{v} \\ (T_{FarTail} - max(T_{FarLead}, T_{NearLead})) > \frac{d_{far}+d^*}{v} \\ (min(T_{NearTail}, T_{FarTail}) - max(T_{NearLead}, T_{FarLead})) > \frac{d_{rider}}{v} \end{cases}, \tag{4.2}$$

where $T_{NearLead}$, and $T_{FarLead}$ denote arrival times for the rear bumpers of the lead vehicles in the near and the far lane gaps respectively, $T_{NearTail}$ and $T_{FarTail}$ denote arrival times for the front bumpers of the tail vehicles in the near and the far lane, v is the average speed of the rider, d_{near} and d_{far} are the distances the rider must travel to cross the near and far lanes, d_{rider} is the length of the bicycle, and d^* is the distance between the rider and the far lane when the far gap opens.

The first of these inequalities simply requires the near lane gap to afford safe crossing as in the case of the single lane crossing. The third inequality states that the overlap between the two gaps (i.e., the temporal interval when both gaps are simultaneously open for the rider) should be sufficiently large to move the bike from the near lane into the far lane. Finally, when judging whether the far gap is crossable, the rider must account for the additional distance (and therefore the additional time) that has to be covered to reach the far lane. The second inequality calls for the far gap (excluding any unusable portion that precedes the opening of the near gap) to be sufficiently large to afford safe crossing of the far lane and any additional distance d* that has to be covered to reach the far lane.

The value of d^* depends on the temporal offset between the opening of the near and the far gaps ($T_{FarLead} - T_{NearLead}$), providing riders with two qualitatively different opportunities for crossing. When the far gap opens before or with the near

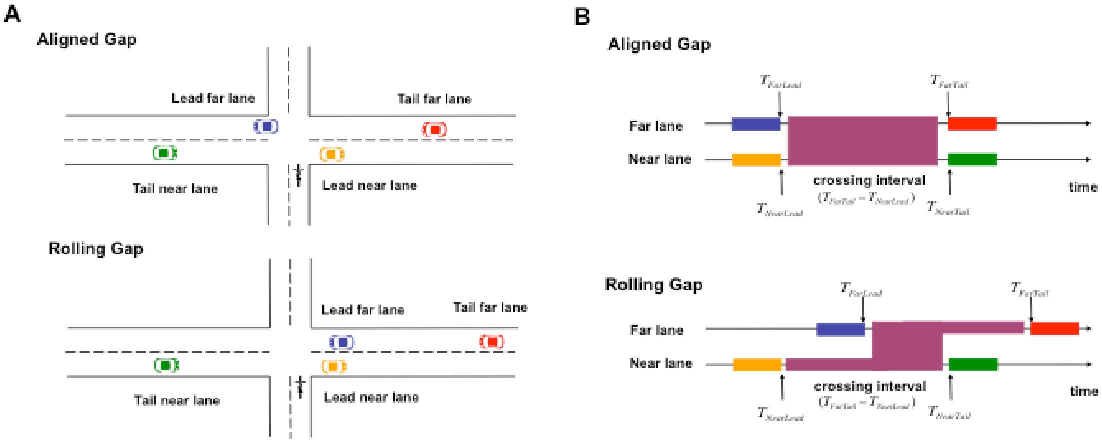

Figure 4.1: Aligned and rolling gap pair shown as both spatial and temporal arrangements from the cyclists perspective. The spatial presentation (A) shows vehicle positions at a snapshot in time. The temporal presentation (B) shows the times that the riders line of travel is obstructed by a vehicle. The interval when the gap is open is highlighted in purple. Note that the temporal presentation represents the progression of time for both lanes of traffic in a temporally aligned (left to right) coordinate system. In this idealized case, both types of gap pairs are composed of the same size near and far gaps.

gap (i.e., $T_{FarLead} T_{NearLead}$), then $d^* = d_{near}$. A cyclist can consider the available temporal crossing interval when both lanes are clear of vehicles as a singular gap (or an“aligned” gap pair) spanning both lanes of traffic (Figure 4.1A). This makes only relatively large far lane gaps crossable since the far gap must remain open long enough for the rider to clear both lanes. Conversely, in a“rolling” gap pair (named by Brewer, Fitzpatrick, Whitacre, & Lord [6]) the near lane gap opens before the far lane gap (Figure 4.1B). If the temporal offset is sufficiently large, the cyclist can enter the near lane before the far gap opens, bringing d^* close to zero. This potentially allows the rider to cross much smaller gaps in the far lane compared to an aligned gap pair.

4.2.2 Coordinating self and object movement

Once cyclists identify a gap (in case of a single lane crossing) or a gap pair (in case of crossing two lanes of traffic) that affords safe crossing, they must coordinate self and object movement to cross the intersection safely.

When crossing a single lane of traffic, cyclists should time their movement so that they enter the intersection soon after the lead vehicle in the gap clears their path. This strategy maximizes the temporal safety margin with the oncoming tail vehicle in the gap. Earlier work shows that child pedestrians and cyclists tend to delay initiation of road crossing relative to adults [54, 75, 76, 78, 99, 118]. Plumert and colleagues [76, 78] found that even 10- and 12-year-old cyclists delayed initiation of movement in a single lane crossing scenario, leading to reduced safety margins relative to adults. This delay is at least in part due to difficulties with coordinating self and object movement by children and early adolescents [14, 100].

Coordinating self and object movement is significantly more complicated when crossing multiple lanes of opposing traffic than when crossing a single lane of traffic. When crossing through two gaps, the second lead vehicle to pass the rider's line of travel acts as a gate to the gap pair. Once it passes the rider, both gaps are open and the rider can move from the near lane to the far lane. By synchronizing their movement to the arrival of the second lead vehicle, the riders can simplify the problem of crossing two one-lane gaps into the easier problem of crossing one two-lane gap.

Due to the differences in the configuration of rolling and aligned gap pairs, riders should use different approaches to coordinate their crossing of the two types of gap pairs. For aligned pairs, the second lead vehicle is in the near lane. Although the far gap is already open, the rider must wait for the near gap to open before entering the roadway. Once the near lane gap opens, however, the roadway in front of the cyclist is fully clear of traffic. This means that the rider should enter the near gap as soon as the lead vehicle in the near lane passes and then simply shoot through both gaps. For rolling pairs, the second lead vehicle is in the far lane. Although the gap in the near lane is already open, the rider must wait until the lead vehicle in the far lane passes before entering the far lane. To maximize the time to spare in both the near and far lanes, the rider should enter the near lane so as to cut in behind the lead vehicle in the far lane as soon as possible (once the near lane is open), and then ride across the remainder of the intersection.

Timing ones movement may be more difficult when crossing through a rolling gap pair than through an aligned gap pair for two reasons. The first is that the focus on the lead vehicle in the far lane (as opposed to a lead vehicle in the near lane) means that riders must coordinate their actions with an event that occurs further ahead in time and distance, making crossing through a rolling gap pair a more demanding perceptual-motor coordination problem. The second reason is that crossing through a rolling gap pair requires riders to split their attention between the lead cars in the near and far lanes. When the lead car in the near lane arrives shortly before the lead car in the far lane, the rider must be careful to not enter the near lane before the lead car in the near lane has completely passed. This is not the case with an aligned gap pair, where the lead car in the far lane can be dismissed because it will have passed before the rider begins to cross the near lane. The added demand of dividing attention between both lead vehicles may make rolling gap pairs more difficult to cross than aligned pairs.

4.2.3 The current study

In this study participants rode an instrumented bicycle across a series of intersections in an immersive, interactive simulator. Each intersection had two continuous streams of relatively dense cross traffic approaching from opposite directions in both the near and far lanes. Children and adults were asked to cross both lanes of traffic without colliding with any vehicles. Participants had to evaluate a continuous series of opportunities to cross that appeared in a realistic manner until they identified a pair of near and far gaps in the stream that they judged to be safe for crossing. The task was further complicated by the fact that two large overlapping gaps were rarely available, forcing participants to make difficult, sometimes suboptimal choices.

Overall, we expected to see a general preference for rolling gap pairs over similar-sized aligned gap pairs, especially for adults, who tend to be more skilled riders. Although aligned gap pairs present a simpler timing problem than do rolling gap pairs, aligned gap pairs provide less total time available for crossing than do rolling gap pairs for the same sized gaps in the near and far lanes. Due to the relative difficulty of crossing through a rolling gap pair, the choice of whether to cross an aligned vs. a rolling gap pair presents a tradeoff between simplicity and safety, especially for riders with less-developed perceptual-motor skills.

We also expected that differences between the two gap pair configurations would also translate into differences in crossing performance. For example, if the riders are in fact timing their motion relative to the second lead vehicle in the pair, cyclists should exhibit different timing of entry into the intersection when crossing aligned vs. rolling gap pairs. Based on previous work, we also expected to see that the children would time their entry into the intersection less precisely than adults, cutting in less closely behind the second lead vehicle when crossing either rolling or aligned gap pairs. Furthermore, when crossing through an aligned gap pair the simplest crossing strategy would be to ride through the intersection as fast as possible. This suggests that riders should achieve higher average speeds when crossing through aligned than rolling gap pairs.

4.3 Method

4.3.1 Participants

A total of 105 children and adults participated. There were thirty-eight 12-year-olds (M = 12.5 years; SD = .36, 17 females), thirty-one 14-year-olds (M = 14.33 years; SD = .16, 15 females), and thirty-six adults (M = 19.08 years, SD = 1.56, 20 females). All children knew how to ride a bike, with an average of 6.6 years of riding experience for 12-year-olds and 8.5 years of riding experience for 14-year-olds. All adults reported learning how to ride a bike as a child. The children were recruited from a child research participant database maintained by a psychology department at a Midwestern university. Parents received a letter describing the study followed by a telephone call inviting children to participate. Children were paid $10 for their participation. Adult participants were recruited from an introductory level

Figure 4.2: Photographs of the bicycling simulator. Note that the visual angles are correct from the viewpoint of the rider

psychology course at the university, and received course credit for their participation.

4.3.2 Apparatus and Materials

The study was conducted using the same large-screen immersive display setup as described in section 2.2.1.2. Four speakers and a subwoofer provided spatialized traffic sounds.

An instrumented bicycle was mounted on a stationary frame was positioned in the middle of three screens (Figure 4.2). The pedals, handlebars, and right hand brake were all functional. However, participants were not required to balance the bicycle because the bicycle mount was rigid. The steering angle and wheel speed were combined with virtual terrain information to render the graphics corresponding to the bicyclists real-time trajectory through the virtual environment. The rear wheel was mated to a friction-drive flywheel that was connected to a torque motor to generate an appropriate dynamic force, taking into account rider and bicycle mass and inertia, virtual terrain slope, ground friction, and wind resistance.

The computing platform for the simulation environment was a networked cluster of six PCs. The software system was a highly refined real-time ground vehicle simulator developed in-house. This system supported complex scenarios consisting of ambient and programmatically-controlled traffic [17, 113].

The virtual environment was populated with residential buildings, trees, and other roadside features typical of a small town. Participants rode through the town on a two-lane residential roadway intersected by cross streets with continuous traffic at 150 m intervals. There were stop signs at each intersection, indicating that the bicyclist should stop. All roadways were 12 m wide, and at a level grade. There was no ambient automobile traffic on the roadway with the participant.

While children were riding the bike, mothers were asked to complete a nine-item questionnaire regarding their childs bicycling history. The questionnaire was developed in-house to collect general information about children's bicycling skills and experience. Of particular interest in the current study was the mothers report of when their child starting riding a bike without training wheels and their rating of how skillful of a bicyclist you think your child is for his or her age (on a 5-point scale).

4.3.3 Design and procedure

The experiment began with a brief warm-up session designed to familiarize participants with the characteristics of the bicycle and the virtual environment. The experimenter informed participants that they would be riding through a virtual neighborhood, and instructed them to ride as though they were riding in a similar, real-world environment. Participants were asked to accelerate up to a comfortable speed, and to stay in the right lane of the roadway. During the familiarization session, participants were instructed to notify the experimenter if they experienced any simulator sickness. The warm-up session provided participants with the opportunity to learn how to steer, pedal, and stop the bicycle.

Following the warm-up session was a practice session in which participants crossed two intersections with a single lane of traffic. At the first practice intersection, the traffic approached from the left-hand side and was restricted to the near lane. At the second practice intersection, the traffic approached from the right-hand side and was restricted to the far lane.

Participants were instructed to stop at each intersection and to cross when they felt it was safe to do so. The practice session was used to familiarize participants with the basic road-crossing task and with traffic coming from each direction. After crossing the practice intersections, participants crossed 12 test intersections. Again, their task was to stop at each intersection and then cross when they felt it was safe to do so. This time, however, there was traffic in both near and far lanes, approaching from both directions. The cross traffic in each lane consisted of a series of cars traveling toward the intersection at 11.176 m/s (25 mi/h). The gaps in each lane were generated in seamlessly connected sets of six gap sizes ranging between 1.5-6.5 s at 1-second intervals. Each set was comprised of a randomly ordered permutation of the six gap sizes. The first vehicles in each near and far lane set arrived at the intersection simultaneously.

4.3.4 Coding and measures

The coordinates of the rider and the vehicles were recorded on every time step of the simulation. Three key events were automatically identified for each intersection: 1) the time when the rider arrived at the intersection, defined as the time the bicyclist was 10 m from the edge of the intersection; 2) the times when the rider entered the near and far lanes, defined as the time that the front wheel crossed the edge of the lane; and 3) the times when the rider cleared the path of the approaching cars in the near and far lanes, defined as the time when the rear wheel of the bike cleared the path of the approaching car.

If at a given intersection a participant was intercepted by one of the vehicles on the road in either lane, the trial was classified as a collision. Of the 1,260 total crossings in the experiment, participants collided with a vehicle only 39 times for an overall collision rate of 3.1%. The collision rate was similar across the three age groups (12-year-olds = 3.9%, 14-year-olds = 3.5%, and adults = 1.9%). Of the 39 collisions, 11 were close calls (the participant missed the gap by less than 1 s.) The other 28 attempted crossings were dropped from the analyses because it was impossible to tell which gap pair the participant was attempting to cross.

Gap pairs were classified as either aligned or rolling. To take full advantage of a rolling gap pair configuration, the cyclists must be able to cross the near lane before the far lane opens. Based on the average near lane crossing time of 1.5 s, we defined rolling gap pairs as those with a temporal offset ($T_{FarLead} - T_{NearLead}$) of at least 1.5 seconds. Temporal offsets of less than 1.5 s were classified as aligned gap pairs.

4.4 Results

The data analyses are organized into three major sections: gap selection, crossing performance, and safety margins. With respect to gap selection, we were particularly interested in whether participants' choices of gap pairs to cross were influenced by the age of the rider and the type of gap pair (i.e., aligned or rolling). With respect to crossing performance, we were interested how participants in the three age groups coordinated their movement through aligned and rolling gap pairs in terms of direction and speed of travel, and timing of entry into the near and far lanes. Finally, with respect to safety margins, we were interested in whether children had smaller safety margins than adults, and whether participants had greater safety margins when crossing rolling than aligned gap pairs.

4.4.1 Gap Selection

We constructed a series of mixed-effects logistic regression models to analyze participants choices of gap pairs to cross. These models attempted to predict the value of a response variable representing the decision to select or reject each pair of near and far gaps. The gap pairs were described using four independent variables (fixed-effects predictors): near gap size, far gap size, pair type (aligned or rolling), and age group. Participants observed continuous streams of gap pairs until they selected a gap pair for crossing. Therefore, each participant contributed approximately 12 positive responses (i.e., the gap pairs they chose to accept for crossing at the 12 intersections) and an a priori undefined number of negative responses (i.e., the gap

Table 4.1: Mixed-effects logistic regression model for likelihood of selecting a gap pair

Variables	**Estimate**	**SE**	**z-value**	**p-value**
Fixed Effects				
Intercept	-11.83	.45	-26.05	.0001
Age 12	1.35	.44	3.07	.002
Age 14	.44	.48	.92	.36
Rolling Gap Pair	3.06	.37	8.26	.0001
Near Gap Size	.75	.03	25.15	.0001
Far Gap Size	1.03	.07	15.21	.0001
Age 12 x Rolling Gap Pair	-.65	.17	-3.86	.0001
Age 14 x Rolling Gap Pair	-.64	.17	-3.68	.0001
Age 12 x Far Gap Size	-.21	.07	-2.96	.003
Age 14 x Far Gap Size	-.10	.08	-1.29	.198
Rolling Gap Pair x Far Gap Size	-.39	.06	-6.31	.0001
	Variance	**SD**		
Random Effects				
Subject	.18	.43		

pairs that they observed, but chose to reject). To account for individual differences between the participants we also clustered the responses from the same person together by including a random-effect variable for subject. Due to a limited number of observations for each participant, we decided to exclude the effect of intersection from the analysis (i.e., we collapsed observations across intersections).

Following a general model-building strategy suggested by Hosmer & Lemeshow ([41], pp. 91-116), we explored a series of models with univariate predictors, main effect combinations, as well as two-way and three-way interactions between fixed predictors. We first explored the significance of the fixed predictors by constructing four separate univariate models containing a single fixed predictor and the random effect. All four proved to be strong predictors of participants choices ($p = 0.01$). Second, we created a main effects model that simultaneously included all four fixed-effects and the random effect of subject. This model showed all four variables remaining significant predictors in combination with each other. As a third step we explored the inclusion of all possible 2-way interactions (one at a time) into the main effects model. We retained those interactions that were significant at the p = 0.1 level. There were no significant two-way interactions involving near gap size. Next we constructed a model that simultaneously included the main effects and all remaining two-way interactions from step three. All of the terms remained highly significant in combination. Finally, we attempted to add the only possible (due to the absence of two-way interactions with near gap size) three-way interaction between far gap size, age group and gap type into this model and found that it was not significant (p = 0.83). In the end, we obtained a robust model (Table 4.1) with a good fit to the data and excellent discrimination (concordance index c = 0.84). The model highlights three key aspects of participants gap selection:

1. preferences for choosing rolling vs. aligned gap pairs,
2. selectivity with respect to gap size, and
3. age differences in willingness to accept a given gap pair.

4.4.1.1 Preference for Rolling vs. Aligned Gap Pairs

Both children and adults exhibited a preference for rolling over aligned gap pairs. Due to the interaction between far lane gap size and pair type, the relative odds of taking a rolling vs. aligned pair depended on and exponentially declined with increases in the far gap size (Figure 4.3). Adults consistently preferred the rolling

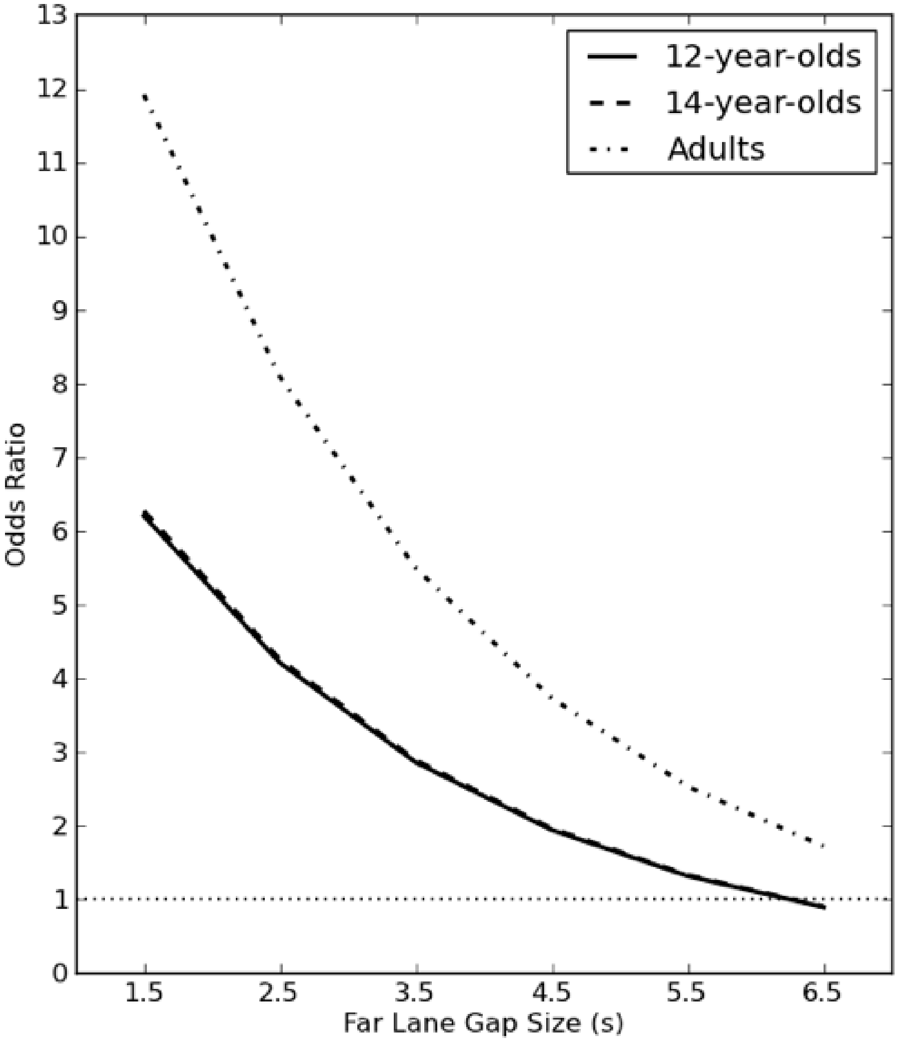

Figure 4.3: Estimated odds ratios for taking rolling over aligned gap pairs for each age group as function of the far lane gap size. An odds ratio of 1 corresponds to equal odds of accepting aligned and rolling gap pairs; an odds ratio grater than 1 corresponds to preference for rolling gap pairs

to the aligned pairs of similar size (as evidenced by the odds ratio greater than 1 over the entire observed range of far lane gap sizes). With the exception of the pairs containing large far lane gaps (6.5 s), this was also true for children. In addition, the odds of taking a rolling over an aligned gap pair were significantly higher for adults than children ($p < 0.001$ for both 12 year-olds and 14 year-olds). Odds ratios for 14-year-olds and 12-year-olds did not differ significantly ($p = 0.93$).

4.4.1.2 Sensitivity to Near and Far Gap Sizes

Overall, participants preferred larger gaps in both near and far lanes. The probability of selecting a gap pair increased with the size of the near gap, regardless of age group, gap pair type (aligned or rolling), or the size of the far lane gap. With every 1-second increase in near gap size, the odds of accepting a gap pair increased by 2.11 (95% confidence interval [1.99, 2.24]).

The probability of selecting a gap pair also increased with the size of the far gap. However, far gap size also interacted with pair type and age group. Participants were significantly more sensitive to changes in far gap size in aligned than in rolling gap pairs ($p = .0001$). Twelve-year-olds were significantly less sensitive to changes in far gap size compared to adults ($p = .003$), whereas 14-year-olds did not differ significantly from adults ($p = .198$) or 12-year-olds ($p = .145$). Figure 4.4 shows the corresponding probabilities of accepting aligned and rolling gap pairs as a function of far gap size for each of the three age groups. The corresponding odds ratios for a unit (1-second) change in far gap size for aligned and rolling pairs were $OR = 2.11$ (95% confidence interval [1.85, 2.41]) and $OR = 1.43$ [1.24, 1.66] for adults, $OR =$

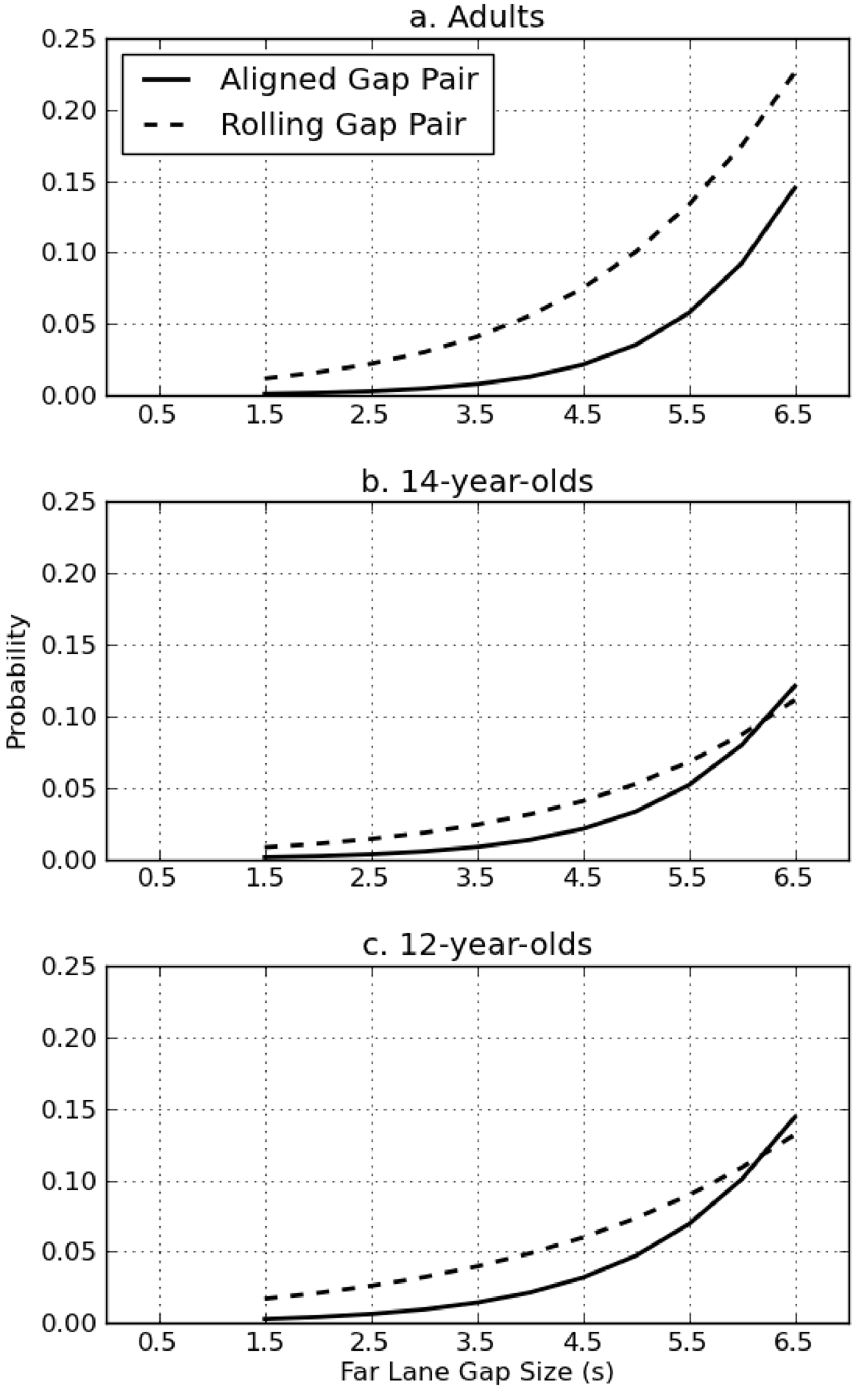

Figure 4.4: Estimated probability of accepting aligned and rolling gap pairs for adults (a), 14-year-olds (b) and 12-year-olds (c). The near gap size is fixed at 4.5 seconds.

1.91 [1.61, 2.27] and OR = 1.3 [1.08, 1.56] for 14-year-olds, and OR = 1.71 [1.46, 2.01] and OR = 1.16 [0.98, 1.39] for 12-year-olds, respectively. Here, the higher odds ratios correspond to steeper slopes of the probability functions.

These results suggest that participants generally selected larger gaps in both the near and the far lanes whenever traffic conditions presented such opportunity. In addition, steeper slopes of the probability functions for aligned gap pairs indicate that the riders were significantly more selective in choosing far lane gaps for the aligned pairs than for the rolling pairs. This confirms our theoretical prediction that riders could identify crossable rolling gap pairs that include smaller far lane gaps compared to the aligned gap pairs. Finally, the age-related differences in the slopes of the probability functions indicate that more mature riders also had more precise criteria for selecting crossable gap pairs.

4.4.1.3 Age Differences in Willingness to Accept Given Gap Pairs

Our model indicates that overall willingness to accept a gap pair depended on age group, pair type, and far gap size. For both aligned and rolling gap pairs, the odds ratio comparing the willingness of children relative to adults to accept a gap pair with a given far gap size exponentially decreased for larger far lane gaps (Figure 4.5). For rolling gap pairs, 14-year olds were consistently more conservative than adults in their gap pair choices (the corresponding odds ratio is consistently lower than 1 over the entire range of observed far lane gap sizes). In contrast, 12-year-olds were more willing than adults to select a rolling gap pair that included a tight gap in the far

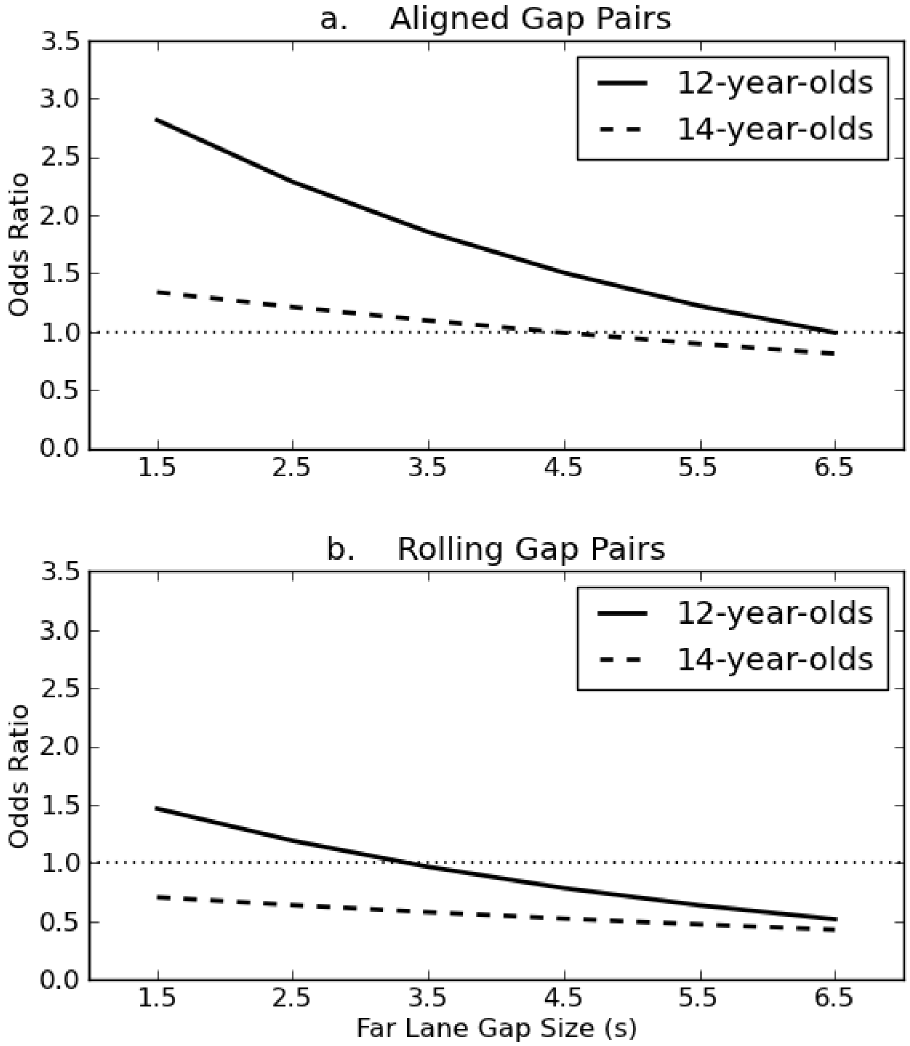

Figure 4.5: Estimated odds ratios for willingness of children relative to adults to accept an aligned (a) or a rolling (b) gap pair as a function of far lane gap size. An odds ratio of 1 corresponds to equal odds of accepting a gap pair for adults and children; an odds ratio greater than one shows that children are more likely to accept a given gap pair than are adults.

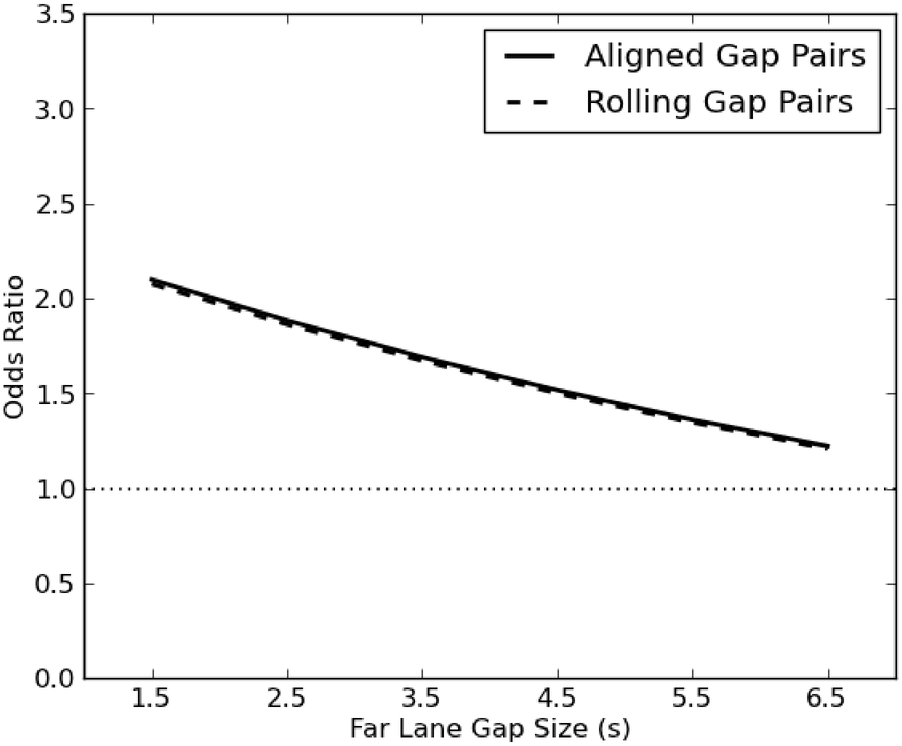

Figure 4.6: Estimated odds ratios for relative willingness of 12-year-olds compared to 14-year-olds to accept an alinged or a rolling gap pair. An odds ratio of 1 corresponds to equal odds of accepting a gap pair for 12-year-olds and 14-year-olds.

lane. For aligned gap pairs, 14-year-olds were more willing than adults to accept an aligned gap pair that included a relatively small far lane gap. Similarly, 12-year-olds were more willing than adults to accept aligned gap pairs with far gaps of any size, except for the very large (6.5s) gaps. Odds ratios indicate that 12-year-olds were also more willing than 14-year-olds to accept both aligned and rolling gap pairs (Figure 4.6).

4.4.2 Crossing Performance

Our second set of analyses focused on how children and adults negotiated crossing through selected gap pairs. Of particular interest was whether the choice to cross an aligned or rolling gap pair constrained how children and adults coordinated their movement through the gap pair. First we explored whether participants adjusted their direction and speed of travel when crossing aligned vs. rolling gap pairs. Then we looked at how children and adults timed their entry into the roadway when they crossed aligned vs. rolling pairs. All variables were analyzed in Age (12 years vs. 14 years vs. adults) x Pair Type (aligned vs. rolling gap pair) mixed model ANOVAs with the first factor as a between-subjects variable and the second as a within-subjects variable. All follow-up pair-wise comparisons were conducted using Fishers Protected Least Significant Difference (PLSD) test with an alpha level of .05.

4.4.2.1 Crossing Trajectories

Travel direction. The analysis of participants lateral motion from entering to exiting the roadway for aligned and rolling gap pairs revealed significant main effects for both age, F (2, 102) = 10.43, p <.001, η_p^2 = .51, and pair type, F (1, 102) = 78.55, p <.001, η_p^2 = .44. Follow-up tests showed that 12-year-olds overall travel direction (M = .013, SD = .035) was shifted leftwards while the travel direction of the 14-year-olds (M = -.009, SD = .036, p <.001), and adults (M = -.01, SD = .034, p ¡ .001), were shifted to the right. More importantly, as Figure 4.7 shows the participants were moving leftward when crossing an aligned gap pair (M = .015, SD = .025) and rightward when crossing a rolling gap pair (M = -.018, SD = .039). This difference between the two pair types suggests that when participants chose a rolling gap pair, they aimed for the opening of the gap in the far lane. In other words, because the gap in the far lane was not open when participants entered the near lane, they veered

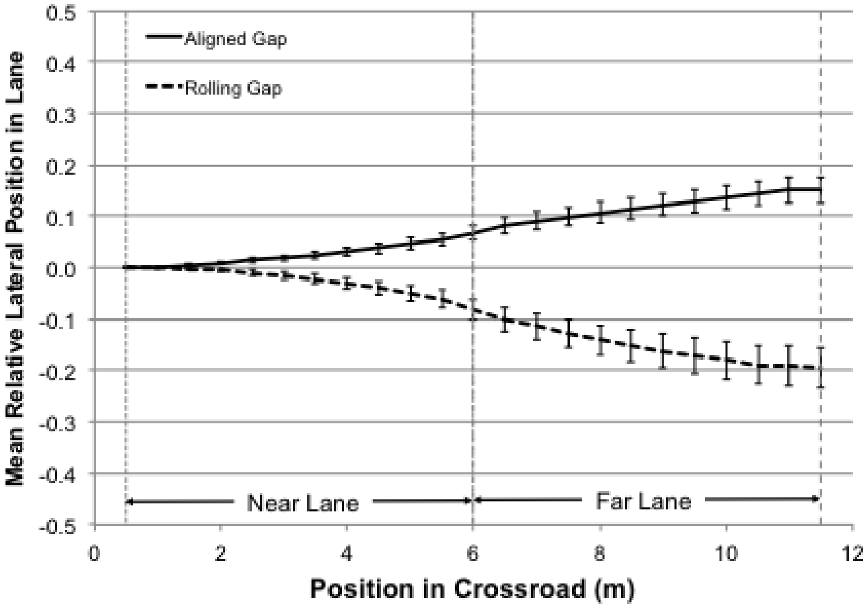

Figure 4.7: Mean y-coordinate trajectories associated with crossing either an aligned or rolling gap pair. The trajectories are normalized relative to the mean y-coordinate when the cyclists started crossing the intersection

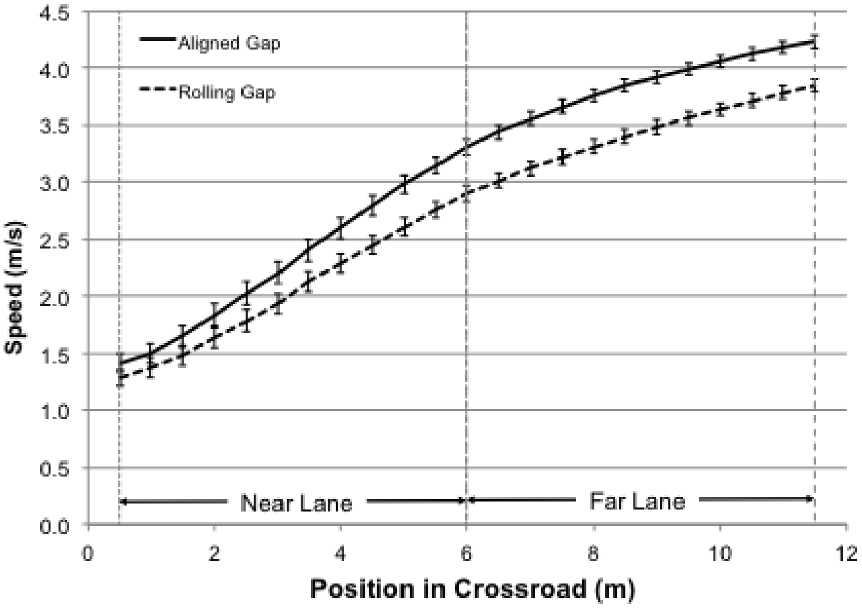

Figure 4.8: Mean speed profiles associated with crossing either an aligned or rolling gap pair.

slightly to the right as they aimed for the opening of that gap. In contrast, when participants chose an aligned gap pair, they veered to the left to avoid being hit by the closing of the far lane gap.

Crossing speed. Figure 4.8 shows the mean speed trajectories associated with crossing an aligned and a rolling gap pair. Participants crossing speed was higher when they crossed aligned gap pairs (M = 6.8 mi/h, SD = 1.5) than rolling gap pairs (M = 6.1 mi/h, SD = 1.5), F (1, 102) = 112.42, p <.001, $\eta_p^2 = 0.52$. This difference between the two pair types indicates that participants were able to pick up greater speed when crossing through an aligned gap pair.

4.4.2.2 Timing of Entry into Near and Far Lanes

A critical measure of how participants timed their movement is how closely they cut in behind the lead car in the gap (Figure 4.9ab). Small values indicate that the rider cut in closely behind the lead vehicle in the gap, while large values indicate that the rider hesitated longer. The key advantage of entering the lane shortly after the lead car clears the riders line of travel is the ability to fully utilize the time available within the gap and therefore maximize the safety margins with the tail car.

The analysis of the near lane revealed effects of age, F (2, 102) = 5.68, p <.01, $\eta_p^2 = .10$, and pair type, F (1, 102) = 288.61, p <.001, $\eta_p^2 = .74$. Overall, adults (M = 1.50 s, SD = .61) cut in closer behind the lead vehicle than 12-year-olds (M = 1.89 s, SD = .83, p = .001), and 14-year-olds (M = 1.75 s, SD = .65, p = .048). In addition, participants cut in substantially closer behind the lead vehicle in the near lane when crossing an aligned gap pair (M = 1.29 s, SD = .55) than when crossing a rolling pair (M = 2.14 s, SD = .62). This difference between the gap types suggests that participants movement timing was influenced by the type of gap pair they chose to cross. When crossing a rolling pair, entering the near lane too soon would risk either stalling in the middle of the lane or crashing into the lead vehicle in the far lane. These factors did not apply when crossing an aligned gap pair because both gaps were simultaneously open, allowing riders to cut in closer behind the lead vehicle in the near lane.

The analysis of the far lane also revealed effects of age, F (2, 102) = 13.95, p <.001, $\eta_p^2 = .22$, and pair type, F (1, 102) = 883.43, p <.001, $\eta_p^2 = .90$. Again, adults

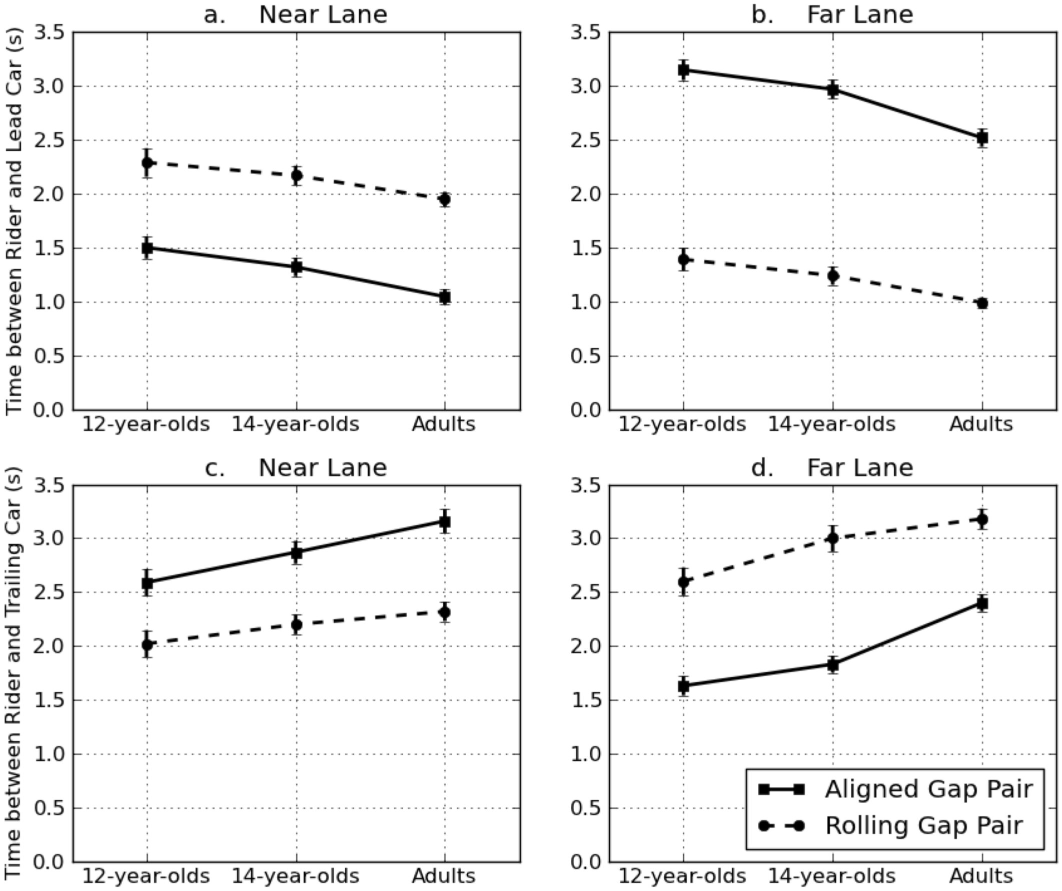

Figure 4.9: Timing of entry relative to the lead car in the near lane (a) and far lane (b), and time-to-spare relative to the trailing car in the near lane (c) and far lane (d) for 12-year-olds, 14-year-olds, and adults as function of selected gap pair type.

(M = 1.75 s, SD = .87) cut in closer behind the lead vehicle in the far lane than 12-year-olds (M = 2.27 s, SD = 1.08, p <.001), and 14-year-olds (M = 2.11 s, SD = 1.00, p = .001). Participants also cut in substantially closer behind the lead vehicle in the far lane when crossing a rolling gap pair (M = 1.2 s, SD = .52) than when crossing an aligned pair (M = 2.88 s, SD = .59). The difference between the pair types in timing relative to the lead vehicle in the far lane and in rightward veering while crossing the intersection suggests that participants were keying their movement timing off of the lead vehicle in the far lane when they crossed a rolling gap pair.

4.4.3 Safety Outcomes

The final issue of interest was the outcome of the crossing in terms of time to spare on clearing each lane of traffic and the overall minimum margin of safety. All performance measures were analyzed in Age (12 years vs. 14 years vs. adults) x Pair Type (aligned vs. rolling gap pair) mixed model ANOVAs with the first factor as a between-subjects variable and the second as a within-subjects variable. All follow-up pair-wise comparisons were conducted using Fishers Protected Least Significant Difference (PLSD) test with an alpha level of .05.

4.4.3.1 Time-to-spare When Exiting the Near and Far Lanes

We first examined how participants fared as they cleared each lane. A critical measure of safety is the time left to spare when participants clear the path of the oncoming car (Figure 4.9cd). Less time-to-spare indicates a reduced safety margin between the cyclist and the approaching vehicle.

The analysis of the near lane revealed an effect of pair type, F (1, 102) = 84.03, $p < .001$, $\eta_p^2 = .45$, indicating that as expected participants had significantly more time-to-spare when crossing an aligned pair (M = 2.87 s, SD = .72) than when crossing a rolling gap pair (M = 2.18 s, SD = .64). There was also a significant main effect of age, F (2, 102) = 6.28, $p < .01$, $\eta_p^2 = .11$. Although the Age x Pair Type interaction was not significant, F (2, 102) = 1.08, 12-year-olds had significantly less time-to-spare than adults for both aligned and rolling gap pairs, p = .001, whereas adults and 14-year-olds did not differ significantly for either aligned or rolling gap pairs.

The analysis of the far lane revealed an effect of pair type, F (1, 102) = 245.49, $p < .001$, $\eta_p^2 = .70$, and an Age x Pair Type interaction, F (2, 102) = 3.14, $p < .05$, $\eta_p^{=}.06$. In this case, participants had substantially more time-to-spare when they crossed a rolling gap pair (M = 2.91 s, SD = .73) than when they crossed an aligned pair (M = 1.95 s, SD = .59). In addition, 12- and 14-year-olds had significantly less time-to-spare than adults when they crossed an aligned gap pair, $p < .001$ and $p < .001$, respectively. When crossing a rolling pair, 14-year-olds and adults had significantly more time-to-spare than 12-year-olds (p = .019 and $p < .001$, respectively).

The fact that participants had more time to spare in the near lane when crossing aligned pairs and more time to spare in the far lane when crossing rolling pairs is particularly interesting given the mean size of the gaps in the near and far lanes that participants crossed. Participants actually crossed smaller near lane gaps for aligned (M = 5.6 s, SD = .42) than for rolling gap pairs (M = 5.92 s, SD = .38), t (104) = -6.34, $p < .001$, and they crossed smaller far lane gaps for rolling (M = 5.34 s, SD = .65) than for aligned gap pairs (M = 5.98 s, SD = .33), t (104) = 10.13, $p < .001$. This suggests that by keying their movement off of the second lead vehicle, participants were able to achieve larger safety margins even when crossing smaller gaps.

4.4.3.2 Minimum Safety Margin

A final measure of the cyclists crossing performance is the closest call they encountered with one of the two trailing vehicles. For each intersection, the minimum safety margin was calculated by taking the smaller of the times-to-spare when clearing each lane of traffic. The analysis of the mean minimum safety margin revealed effects of age, F (2, 102) = 15.77, $p < .01$, $\eta_p^2 = .24$, and pair type, F (1, 102) = 7.75, $p < .01$, $\eta_p^2 = .07$, Adults (M = 2.10 s, SD = .45) had a larger minimum safety margin than 14-year-olds (M = 1.81 s, SD = 0.51, $p < .001$), who had a larger minimum safety margin than 12-year-olds (M = 1.51 s, SD = 0.64, p = .007). In addition, participants had a larger minimum safety margin when crossing a rolling gap pair (M = 1.87 s, SD = .63) than when crossing an aligned gap pair (M = 1.72 s, SD = .55). The fact that choosing a rolling gap pair resulted in a substantially larger minimum safety margin indicates that rolling gap pairs were actually safer than aligned gap pairs.

4.5 Discussion

This case study aimed to examine an example how perception of complex affordances can be studied in an immersive, interactive virtual reality simulator. To this end three groups of participants with different perceptual and motor skills (adults, 12-year-olds and 14-year-olds) were asked to cross two lanes of opposing traffic on a bike. We found that when facing the challenge of crossing two lanes of relatively dense traffic both children and adults preferred rolling over aligned gap pairs, but this preference was significantly stronger for adults. As argued earlier rolling gap pairs are safer, because they "stretch" total time available for crossing. Although crossing rolling gap pairs may be more difficult, the riders were able to take advantage of the additional available time. When crossing through rolling gap pairs bicyclists achieved an impressive 49% gain in time-to-spare in the far lane of traffic and a significant 8.7% gain in the mean minimal safety margin compared to crossings of aligned gap pairs.

As expected, participants appeared to key their movement off of the second lead vehicle to reach their line of travel. This meant that they cut in closely behind the lead vehicle in the near lane gap when crossing aligned gap pairs and they cut in closely behind the lead vehicle in the far lane gap when crossing rolling gap pairs. However, children significantly delayed their entry into the near lane compared to adults (by 26% for 12-year-olds and 17% for 14-year-olds). Given the similar crossing speeds this resulted in substantially smaller minimum margin of safety compared to adult riders (28% smaller for 12-year-olds and 14% smaller for 14-year-olds).

The complexity of the task played an important role in revealing the differences in perceiving affordances. In earlier studies, where participants had to cross a single lane of traffic [76, 78], children and adults chose the same size gaps for crossing. Similarly, in an observational study involving young and old adult pedestrians [72] the differences between participants were much more pronounced in a two-lane crossing scenario. This suggests that a more complex scenario may be better suited for revealing differences between experimental groups in studies of affordance perception.

The distinguishing feature in this particular task is that participants have to attend to multiple moving objects and choose a possibility to cross from a continuous stream of gaps. This makes the task considerably more complex, than a typical experimental task in a study of dynamic affordances, where participants respond to a single moving object [11, 12, 24, 28, 71]. Furthermore, the participants have to attend to objects that are not simultaneously observable and, therefore, integrate remembered and visual information.

The key take-away message for designing an experiment to compare different configurations of virtual reality systems is two-fold:

1. The study presented here provides an example of treating perception and action as an integrated, intertwined system. It identifies and defines quantitative measures to analyze two components of a perception-action system: choosing an opportunity to cross (perceptual component) and execution of action (action component). A final piece of analysis - the safety margin metric is used to evaluate how accurately participants were able to perceive the affordances.
2. The complex, realistic task selected here revealed between-group differences associated with both selection of a gap to cross (i. e. perceived affordances for action) and the timing of the actual crossing maneuvers. The findings presented here suggest that in order to reveal the differences between groups of participants, the task needs to be sufficiently complex. It is imperative to push the

participants towards the limits of their abilities by creating a situation that places participants outside of their comfort zone and requires making difficult choices.

CHAPTER 5
DEVELOPING METHODOLOGICAL APPROACH TO STUDY EFFECTS OF VR SETUP ON DYNAMIC INTERACTIONS

5.1 Introduction

Practical applications of virtual reality technology such as behavioral research, simulation training, and entertainment involve a wide range of user interactions with the virtual world. One particular type of interaction that often comes up in these scenarios requires user to coordinate self-motion with the motion of dynamically moving virtual objects. The user performance in these tasks depends to a large degree on the ability to accurately perceive the dynamically changing configuration of the the environment as well as accurately assess one's own capabilities for motion, which are determined by the type of the locomotion interface provided by the VR system and can have drastically different characteristics than the familiar real-world means of locomotion. In essence, the user performance depends on the ability to accurately perceive affordances in the virtual world.

The characteristics of the VR system can play a critical role in perception of affordances. After all, the user's perception-action loop with the virtual world is fully mediated by the VR hardware. From the human-computer interaction perspective the question of how the system configuration may impact performance for a given task is very important. However, relatively little has been done to study the effects of the system itself on perception-action and to compare the effects of different VR setups.

This chapter builds of theoretical approach proposed in Chapter 3 and the case study discussed in the previous chapter to propose a methodological approach to an experiment designed as a step toward systematic exploration of how the different types of VR system components can impact user performance (in terms of the percetion-action link) in a dynamic interactive task that requires coordination of self-motion with the motion of virtual objects. Specifically, we will compare VR systems using two common display types (large immersive screen setup and head-mounted display) and two locomotion modalities (physical walking and joystick-controlled motion).

5.2 Choice of experimental factors

The configuration of a VR system is defined by a multitude of components, corresponding to various aspects of multi-sensory feedback and participant's abilities to control her interaction with the system. However display type and the choice of interface to explore the virtual world are the most common choices facing the engineer of a virtual reality system. As such these aspects seem to be logical choices for experimental factors in an evaluation study. Moreover, these two components map in a straight-forward way to user's perception and action: display characteristics directly impact visual perception of the virtual world and locomotion interface defines action capabilities (at least for applications that require physical motion through the environment).

5.2.1 Locomotion interface

Due to our focus on full-body interaction, the choice of input interface narrows down to locomotion. Although a large number of locomotion interfaces have been proposed for VR systems, two types remain the most common: joystick-based interfaces and naturalistic walking in a tracked space. We propose to include these two interfaces as alternative locomotion modalities to be compared in an experiment.

A number of studies have shown cognitive benefits of naturalistic locomotion as compared to joystick [58, 86, 95] in exploration and navigation tasks. The benefits include faster exploration of the environment and improved mental maps of the virtual places visited during the task. Ruddle et al. [84] have also shown that realistic walking that supplies users with information about rotation and translation of the body provides superior results for navigation tasks compared to interfaces that support only rotations (orientation tracking + HMD) or no embodied motion at all (21 inch monitor non-stereo).

In addition, naturalistic walking seems to improve user's sense of presence in virtual environment [105]. Participants reported higher sense of presence in virtual reality application, when physically walking though the environment as compared to "flying" through the environment using a joystick.

Whitton et al. [112] investigated the impact of a locomotion interface on task performance, such as precision in reaching a target during the locomotion. While their results are somewhat inconclusive, they suggest that more "natural" interfaces (such as physical walking) enable better performance compared to joystick-based interface.

5.2.2 Display type

The modern display systems used in most of VR setups can be broadly grouped into two types: large-screen immersive displays displays and head-mounted displays.

In our study we propose to directly compare the effects of a CAVE-like[20] large-screen setup and an HMD.

The key characteristic of the large immersive screen setups that sets them apart from HMDs is the wide field-of-view (FOV). This is particularly true in the multi-screen settings, such as the CAVE-setup, composed of 3 to 6 screens arranged at right angles around the user. In these setups the field-of-view supported by the display approximates natural human FOV. Large screen setups vary in their ability to support motion parallax and stereoscopic imagery, which might also be important for accurately perceiving the location and motion of dynamic moving objects. The present investigation will focus on a CAVE-like setup enabled with tracking for the position of the user head to support motion parallax, but lacking stereoscopic capabilities. The key disadvantage of the large screens setup is that the stationary placement of the screens severely limits user ability to physically move in the environment. This limitation, however, can be at least partially negated by the use of the appropriate locomotion interface.

Direct comparisons between CAVE-like and HMD setups are relatively rare, because only a few research labs have both types of setups available. There is, however, some discussion in the literature concerning the benefits of using large-screen displays in general and relative benefits of large screen vs. HMD.

For example, Ruddle shows faster navigation time and better mental representation in a navigation task in an immersive HMD setup compared to a desktop display [85]. It is important to note, however, that a desktop display typically can not be immerse the user into the virtual world to a degree similar with an HMD or a CAVE setup. Tan et al. [98] on the other hand show that physically large displays enable faster navigation, improve mental maps and help users to perform better in mental rotation tasks.

5.3 Choice of experimental task

What kind of task would be appropriate to use for the planned experiment? The obvious requirement for the kind of the task to be selected is that it represents user interactions that we intend to study. As such, we are looking for a task that involves dynamic coordination of full-body self-motion with motion of objects in the virtual world. A number of other properties may be beneficial:

1. The participants should be able to walk in a relatively compact space, making it possible to conduct the experiment with a CAVE-like setup for large projection screens and the HMD with limited tracking area.
2. The task should be suitable for locomotion and can accommodate both physical walking and joystick-controlled movement.
3. For a strong candidate task, different aspects of the task design should alternatively play to the strengths and weaknesses the two display types that we plan to compare. For a non-stereoscopic larger-screen display the large field-of-view may help with need to track moving objects in the peripheral vision. On the other hand, the lack of stereoscopic clue may detract from accurate perception of motion in front of the user.

We propose to model the dynamic interaction task on the road-crossing scenario discussed in Chapter 4. In particular, we would like users to evaluate a continuous stream of action opportunities. Another useful aspect that we hope to imitate is the need to track multiple moving objects and make timing decisions based on estimation of their time of arrival to the user's line of travel. Finally, the split attention aspect of the task may be useful to create a sufficiently challenging problem for the users, which is necessary to reveal differences in preferences for choices of action opportunities.

One modification we propose to use involves the kind of anticipation that participants have to make. In the biking study they had to anticipate movement of the two streams of objects moving in opposite directions. In this task one of the streams of opportunities will be modeled as temporal anticipation for the stream of pass/no pass situations created by a gate that opens and closes periodically.

The full details of the proposed task are described in the methods section of the next chapter.

CHAPTER 6
EXPERIMENTAL STUDY: IMPACT OF DISPLAY TYPE AND LOCOMOTION MODALITY ON PERFORMANCE IN A DYNAMIC FULL-BODY INTERACTION TASK

6.1 Method

6.1.1 Task

The task immersed participants into a virtual environment representing a train station. At the beginning of each trial the participant was standing on a train station platform facing the railroad tracks. She was facing a transparent sliding gate located close to the edge of the platfom. There was also a freight train passing by the station with constant speed. The train was composed of two types of railroad cars: a flat platform car and a tall freight car. The participant was instructed to wait for a situation when there was a flat platform car directly in front her and she can pass through the open gate and then move forward to board the train (Figure 6.1).

The gate was located approximately 8 ft in front of the participant (this distance was selected to maximize the distance from the train and ensure that the gate was fully visible on the front screen of the CAVE setup). The gate opened and closed with periodic regular intervals. Each gate opening interval (from the moment the gate started to open to the moment when it was completely closed again) was 1.5 s long. There was a 3.0 s pause until next opening.

The train was moving with constant speed in such a way that each train car passed the participant in 3.0 s (both flat platform cars and tall freight cars had the same length of approximately 9 ft). The speed of the train was automatically

Figure 6.1: View of the virtual train and the gate in “Early” (left) and “Late” (right) configurations.

calculated based on the train car length and the required time interval for the car to pass the participant. The motion of the train and the gate opening cycle were synchronized so that they produced one of the following three configurations between the platform and the gate:

- *“Closed”* - the gate remained closed the entire time that the platform was passing in front of the participant.
- *“Early”* - the gate started to open when the front of the platform passed the participant and closed approximately half-way through the platform passing (1.5 s after platform arrival).
- *“Late”* - the gate started to open half-way through the platform passing interval (1.5 s after platform’s arrival) and closed when the back of the platform was about to pass the participant.

The algorithm for adding cars to the train was designed to maintain synchronization between the gate opening cycle and the passage of the train and to guarantee two important properties:

1. Each platform is followed by a tall freight car. This ensures that each opportunity to board the train is visually and logically distinct.
2. The three possible configurations between the platform and the gate opening listed above are uniformly distributed across the train.

These properties were achieved by synchronizing the arrival of the first platform with the gate opening in one of the desired configurations (randomly selected between gate closed, gate opens late, and gate opens early) and then uniformly randomly inserting one, two, three, or four tall train cars after each platform car. Regardless of the initial configuration this rule will maintain train-gate synchronization and produce uniformly random distribution of configurations.

To visually mark the end of each experimental trial the train froze as soon as the participant boarded. In this experiment participants were only provided with the visual stimulus, there was no sound.

The motion of the participants in the virtual environment was scaled relative to the real world. This was done to allow physical walking in the CAVE setup. If the motion of the participant is not scaled, she would have to walk through the front screen in order to reach the target location (the train) that was initially projected onto that screen. Scaling the motion in the virtual environment by a factor of 2 relative to the real-world motion guarantees that the participants will only have to walk part of the distance to the front screen in order to reach the train. When participants controlled their motion through the virtual environment using joystick, their speed was set to twice the average brisk walking speed of 1.5 m.s (or 3.35 mph) to simulate scaled motion.

Due to a programming error the viewpoint in virtual environment was set approximately 1 ft below the actual height of the participants. However, this error should not impact the results, because all comparisons are performed between conditions using the same environment. In addition the environment itself did not faithfully replicate real-world scale. Virtual world objects were modeled approximately and did not faithfully replicate real-wrold dimensions.

6.1.2 Experimental design

The experiment used a between-subject full-factorial two-by-two design with 20 replications for each participant. The two factors are the display type (either a large-screen projection display setup labeled as CAVE or a head-mounted display labeled as HMD) and locomotion modality (physical walking or joystick-controlled locomotion). In combination these factors produced the following four experimental conditions:

1. *CAVE+Walking* - participants performed the task with the CAVE display and physical walking scaled by a factor of 2
2. *CAVE+Joystick* - participants performed the task with the CAVE display and joystick-controlled locomotion
3. *HMD+Walking* - participants performed the task with the HMD display and physical walking scaled by a factor of 2

4. *HMD+Joystick* - participants performed the task with the HMD display and joystick-controlled locomotion

6.1.3 Apparatus and materials

For the HMD conditions we used an nVIS nVisor ST system. The HMD contains two small LCOS displays each with resolution of 1280 x 1024 pixels. Convergence distance for the HMD was set by the manufacturer to 10 meters. The field-of-view is 40.5 degrees vertical an 49.5 degrees horizontal.

For the CAVE conditions we used three 10-feet-wide x 8-feet-high vertical screens placed at right angle to one another, forming a three-walled room. Three projectiondesign F1+ projectors were used to project 1080 x 1024 pixel images onto the side screens, providing participants with approximately 224 degrees horizontal field-of-view of non-stereoscopic, immersive visual imagery. The floor screen was front-projected by a Sanyo PLC_WXE45 projector with the same resolution. The viewpoint was dynamically adjusted for each participant, enabling motion parallax.

A 6 DOF Ascension trakStar electro-magnetic tracker with extended range transmitter was used to determine position and orientation of participant's head in the environment.

Participants were asked to complete Virtual Reality Experience survey based on the Slater, Usoh, and Steed presence questionnaire [106]. The survey attempted to quantify the sense of presence in the virtual environment that participants had during the task. The survey also asked participants to evaluate their own performance in the task on the scale of 1 to 7.

6.1.4 Procedure

6.1.4.1 Briefing

Participants were briefed about the experiment and the task they were asked to accomplish. After providing the informed consent they were asked to complete the computer experience survey and the initial simulation sickness questionnaire to account for symptoms associated with the simulation sickness that were present prior to VR exposure. Participants were also shown the apparatus used for the appropriate experimental condition.

6.1.4.2 Initial calibration

In order to ensure accurate position and orientation measurements the tracker and displays had to be calibrated before the experiment.

For the CAVE conditions the axes of the tracker coordinate system were assumed to be aligned with the coordinate system of the virtual environment. Therefore, initial calibration required only a one-time alignment of tracker axes with the axes of the CAVE setup during the apparatus setup. There was no need for individual calibration prior to each experimental session.

For the HMD conditions each participant needed to adjust the head-mounted display to match the inter-pupular distance and to ensure that the screens were centered. The procedure used on this step was the same as the one described earlier in section 2.2.1.3. In addition, due to variation of HMD fit to the head of the participant the alignment of the tracker had to be calibrated individually for each participant. To measure the rotational offset between the coordinate systems of the tracker and the virtual environment the participants were asked to visually match two markers. The first marker (grey cross) was placed in a fixed position in the environment's coordinate systems. The second marker (orange cross) floated in front of the participant at a fixed position relative to the tracker coordinate system. When the two markers visually overlapped, the tracker's orientation measurements corresponded to the offsets between the two coordinate system. By pressing a button on the keyboard the experimenter recorded the offsets and initialized the correction of alignment.

6.1.4.3 Adaptation to virtual environment

To familiarize participants with the scaled motion through the virtual environment each experimental session included an adaptation phase. The procedure was designed to give participants some visual feedback to adapt to an unfamiliar rate of motion flow due to scaled walking. The participants were placed into an indoor environment (waiting hall inside the train station) with regularly spaced tables and chairs. They were asked to physically walk forward or move by using joystick until they were standing between the nearest pair of chairs. Then the screen was blanked and the participants were asked to return to their initial position. This procedure was repeated three times.

6.1.4.4 Practice trial

Next participants were placed on the platform of the virtual train station. They again received the instructions detailing the task. We encouraged the participants to visually explore the environment by looking around. We also asked them to observe the passing of several train platform cars to gain some understanding of interaction between the train and the gate. Finally, the participants completed a practice trial by attempting to board the train. The experiemters verbally commented on their performance to ensure that they fully understood the task.

6.1.4.5 Experimental trials

The participants were asked to complete 20 experimental trials. In the end of each trial we blacked the screens and asked the participants to return to their initial position before commencing the next trial.

6.1.4.6 Debriefing

After the VR exposure participants were asked to complete a second copy of the simulation sickness questionnaire to measure any symptoms that they might have developed as a result. The participants also completed the Virtual Reality Experinece Assesment. Finally, we explained to the participants the general goals of the study to complete the debriefing session.

6.1.5 Participants

A total of 63 undergraduate students taking an introductory psychology course at a Midwestern university were recruited to participate in this study. There were 16 participants in CAVE+Joystick condition (7 male, 9 female), 16 participants in CAVE+Walking condition (7 male, 9 female), 16 participants in HMD+Joystick condition (8 male, 8 female), and participants in 16 HMD+Walking condition(8 male, 8 female).

Each participant completed 20 trials for a total of 1280. Due to experimenter, participant, or data recording errors the data in 49 trials were excluded from the analysis (3.8% of the total). Therefore, the total number of trials in this analysis was 1231.

6.2 System latency and simulator sickness

Before we proceed to analyze the data a special consideration needs to be given to an issue of system latency. The responsiveness of the system plays an important role in maintaining the sense of realism of the virtual world and can also contribute to user performance outcomes.

When talking about latency VR engineers often consider only rendering delay, i. e. the delay between the appearance of the two display frames. Due to the nature of the task used in this experiment, however, we were concerned with the overall sensor-to-display latency.

To measure the delay we produced video recording at 60 fps that showed both sensor and the image on the display screens. By recording a series of movements we could than count the number of video frames between the initiation of sensor movement and the first change of the displayed image. Then these counts were converted into the estimated latency of the system for both the HMD ($M = 178ms, SD = 20ms$) and large-screens setup ($M = 160ms, SD = 18ms$).

These numbers indicate that that system lag was somewhat high for an interactive immersive experience, which may lead to increased simulator sickness. To control for increase in simulator sickness symptoms all participants were asked to complete a Kennedy-Lane Simulator Sickness questionnaire [48] before and after VR exposure. The survey asked the participants to rate themselves for the presence of

Table 6.1: Distributions of simulator sickness (SSQ) scores across experimental conditions.

Condition	**Mean**	**Median**	**Max**
CAVE + Joystick	88.704	35.57	330.744
CAVE + Walking	52.43	3.79	242.483
HMD + Joystick	67.357	3.79	427.026
HMD + Walking	174.329	77.74	1058.492

16 symptoms associated with the simulator sickenss.

We compared the SSQ scores for before- and after surveys and adjusted the former by subtracting the latter. This adjusted score shows the change in simulator sickness symptoms due to VR exposure. To ensure that simulator sickness did not affect conditions unevenly, we compared the adjusted SSQ scores across experimental conditions.

The adjusted scores had a skewed distribution (Table 6.1: most of the participants had low scores (i. e. did not experience simulator sickness), but a relatively small number showed very high scores. Therefore, the comparison between conditions required the use of a non-parametric statistical method, which does not rely on normality assumption. The Kruskal-Wallis rank sum test (analogous to one-way ANOVA for four conditions) revealed no significant differences between the scores in the experimental confitions ($\chi^2_3 = 3.49, p = 0.32$).

6.3 Results

6.3.1 Errors in judgement

In several trials participants made an error of judgement. One type of errors occurred when participants ignored the gate and attempted to board the going though closed gate. There were a total of 28 trials with this type of error in the sample. The errors of this type did not appear to be systematically related to condition $\chi^2_3 = 4.67, p = 0.198$.

The second type of error occurred when participants realized they made a bad choice and aborted the boarding after passing the gate. In these cases they moved back and attempted a second boarding within the same trial. There were 5 trials exhibiting this type of error. We did not perform a separate analyses of errors, because they were very infrequent and did not appear to be systematically related to either display or modality.

As mentioned earlier trials, where participants made errors, were excluded from the analyses that followed.

6.3.2 Choice of boarding opportunity

The choice of the opportunity to cross may be depend on display type, locomotion modality and the configuration of the gate opening relative to platform arrival. In this experiment participants selected opportunities to board the train from a uniform random distribution of choices relative to early and late arrivals. Therefore, the proportions of early and late arrivals selected by participants reflect their overall preferences within each condition.

Figure 6.2 shows choices of boarding opportunities made by participants. The choices in both displays conditions appear to be similar. Participants in walking conditions were significantly more likey to select an early configuration for crossing (as confirmed by likelihood ratio χ^2-test, $p < .001$) compared to participants in

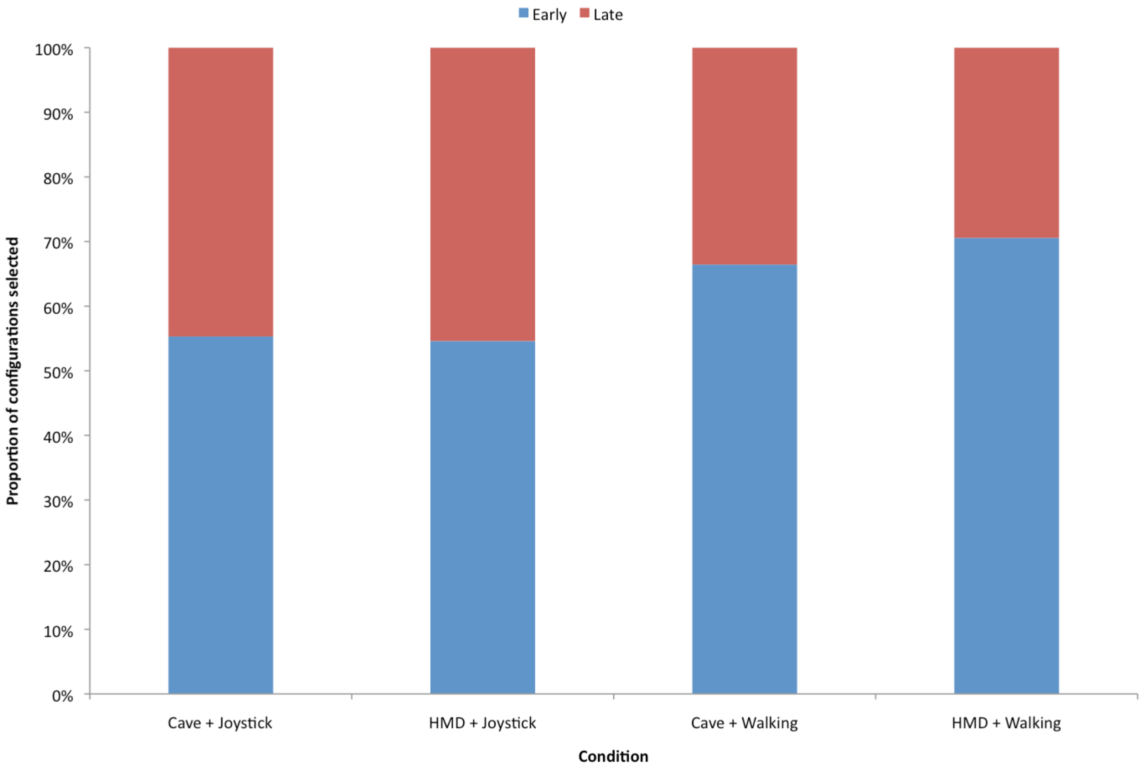

Figure 6.2: Proportion of ”Early” and ”Late” configurations selected for boarding in each experimental condition.

joystick conditions.

6.3.3 Boarding Performance

The analyses in this section consider how participants timed their boarding with respect to gate openings and platform arrival. In addition this section compares participants’ speeds across four experimental conditions. Since all participants started at the same distance from the gate, the differences in timing can be attributed to either choice of time to initiate boarding action or difference in speed. The choice of configuration between gate and platform may be associated to differences in timing and speed. So, two configurations are considered separately.

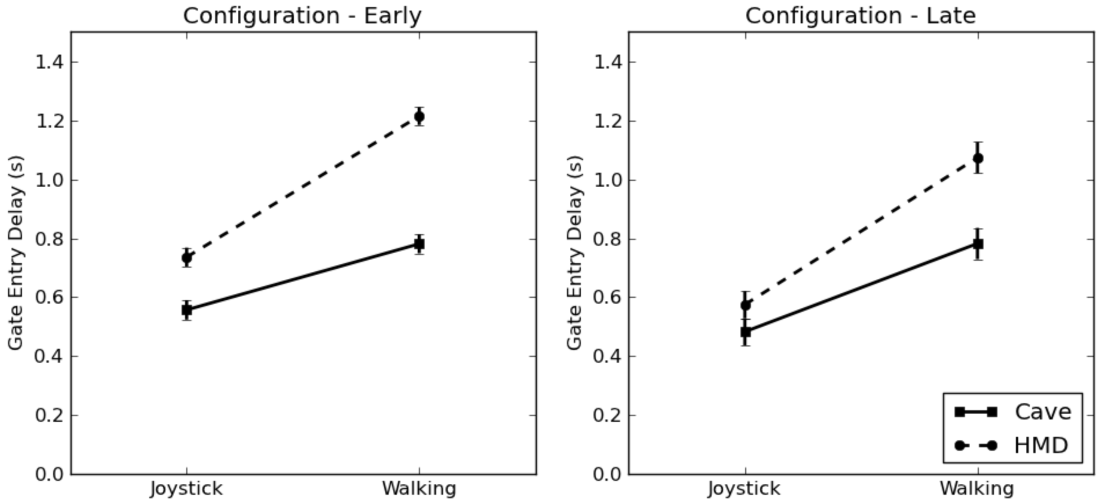

Figure 6.3: Gate entry delay separated for boarding "Early" (left) and "Late" (right) configurations between the gate and the platform. Error bars show standard error.

6.3.3.1 Gate entry delay

The first measure to consider is *Gate Entry Delay.* It is defined as the temporal delay between the moment the gate is fully open and the time the participant pass through the gate. Larger delays mean that participants waited longer to pass through the gate. Negative delays and delays greater 1.5 seconds indicate that participants passed through the gate either before it opened or after it has fully closed.

Figure 6.3 compares mean *Gate Entry Delay* across two display conditions and two locomotion modalities. The comparison is preformed separately for "Early" and "Late" openings of the gate relative to arrival of the platform that participants selected for boarding. Table 6.2 shows mean gate entry delays and corresponding standard errors for both configurations.

In order to compare results across participants we first found average dealy

Table 6.2: Mean gate entry delays separated for selection of early and late configuration.

Condition	**“Early”**		**“Late”**	
	Mean	**SE**	**Mean**	**SE**
CAVE + Joystick	0.572	0.043	0.468	0.059
CAVE + Walking	0.859	0.045	0.915	0.059
HMD + Joystick	0.735	0.043	0.634	0.057
HMD + Walking	1.236	0.043	1.270	0.061

values for each participant in each configuration and then analyzed the resulting values in a two-way ANOVA. Three participants did not contribute any trials in ”Late” configuration.

In the “Early” configuration a two-way ANOVA indicated significant main effects of display type $F(1, 59) = 37.91, p < .001$ and locomotion modality $F(1, 59) = 81.05, p < .001$. There was also significant two-way interaction $F(1, 59) = 5.99, p = .017$. Post-hoc comparisons using Tukey’s HSD procedure revealed significantly larger gate entry delays for walking compared to joystick for both CAVE ($t_{59} = -4.6, p < .001$) and HMD conditions ($t_{59} = -8.16, p < .001$). The difference between display conditions was also significant for both walking ($t_{59} = -6.03, p < .001$) and joystick ($t_{59} = -2.64, p < .05$) conditions.

In the “Late” configuration a two-way repeated measures ANOVA also showed significant main effects of display type $F(1, 56) = 19.62, p < .001$ and locomotion modality $F(1, 56) = 94.61, p < .001$. A two-way interaction was not significant $F(1, 56) = 2.57, p = .114$. Post-hoc comparisons using Tukey’s HSD procedure revealed significantly larger gate entry delays for walking compared to joystick for both CAVE ($t_{56} = -5.38, p < .001$) and HMD conditions ($t_{56} = -7.31, p < .001$). The difference between two display conditions using joystick was not significant ($t_{56} = -2.03, p = .189$). The was a significant difference between walking conditions using CAVE and HMD ($t_{56} = -4.20, p < .001$).

Overall these results show similar pattern in both configurations. First, participants using joystick had smaller delays in crossing the gate compared to walking participants. In addition, delays at the gate were also smaller for participants using larger-screen displays, particularly for walking conditions.

6.3.3.2 Platform entry delay

The timing of participant's boarding relative to platform's was measured using *Platform Entry Delay*, which was defined as the temporal difference between the platform's arrival and participant's crossing of the edge of the platform. Similar to the previous measure, larger delays indicate that participants waited longer to board the platform after its arrival. Positive delays indicate that participants missed the platform by attempting to board too early, delays larger than 3.0 seconds indicate that participants missed the platform by boarding too late.

For this measure there is a built-in difference between early and late configurations, because participants cannot board the platform before the gate is open. Therefore this measure cannot be used to compare across configurations. Figure 6.4 compares mean *Platform Entry Delay* across two display conditions and two locomotion modalities. The comparison is preformed separately for ”Early” and ”Late” openings of the gate relative to arrival of the platform that participants selected for boarding.

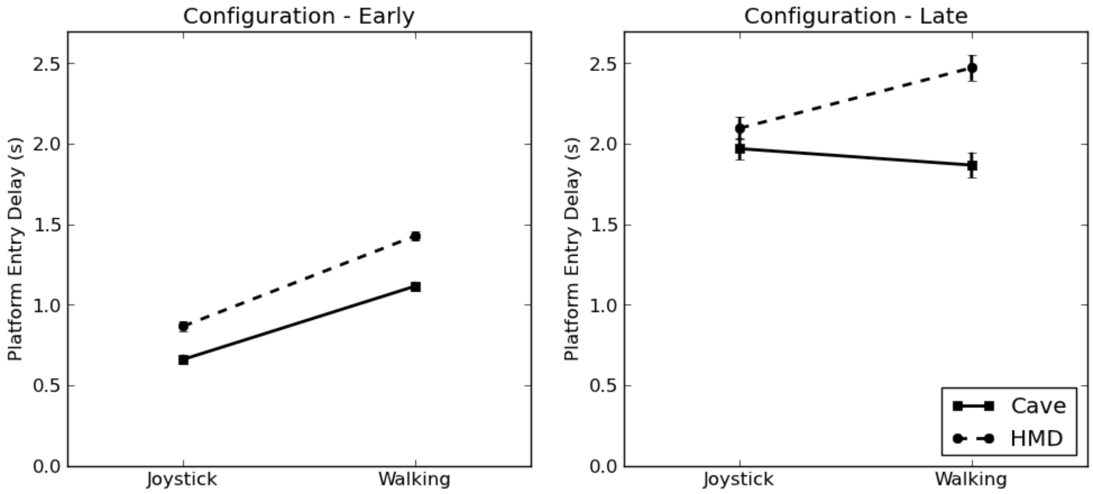

Figure 6.4: Platform entry delay separated for boarding "Early" (left) and "Late" (right) configurations between the gate and the platform. Error bars show standard error.

Table 6.3: Mean platform entry delays separated for selection of early and late configuration.

Condition	**"Early"** **Mean**	**SE**	**"Late"** **Mean**	**SE**
CAVE + Joystick	0.69	0.046	1.971	0.114
CAVE + Walking	1.224	0.048	1.846	0.114
HMD + Joystick	0.865	0.046	2.202	0.111
HMD + Walking	1.467	0.046	2.595	0.118

Table 6.3 shows mean platform entry delays and corresponding standard errors for both configurations. Here we again averaged delays for each participant (separately for each configuration) and then analyzed the data using a two-way ANOVA.

In the "Early" configuration a two-way repeated measures ANOVA revealed significant main effects of display type $F(1,59) = 20.13, p < .001$ and locomotion modality $F(1,59) = 149.04, p < .001$. The two-way interaction was not significant

$F(1,59) = 0.53, p = 0.47$. Post-hoc comparisons using Tukey's HSD procedure indicated significantly larger platform entry delays for walking compared to joystick for both CAVE ($t_{59} = -8.05, p < .001$) and HMD conditions ($t_{59} = -9.22, p < .001$). The difference between display conditions was also significant for both walking ($t_{59} = -5.42, p = .003$) and joystick ($t_{59} = -2.68, p < .05$) conditions.

In the "Late" configuration a two-way repeated measures ANOVA also showed significant two-way interaction between display type and locomotion modality$F(1,406) = 10.37, p = .001$. Post-hoc comparisons using Tukey's HSD procedure revealed no significant difference between joystick and walking for CAVE conditions ($t_{56} = 0.0.78, p = 0.87$), while the difference between walking and joystick in HMD conditions trended towards significance ($t_{56} = -3.57, p = .084$). The difference between display types for joystick conditions was not significant ($t_{56} = -1.45, p = .474$). There was a significant difference between walking conditions using CAVE and HMD displays ($t_{56} = -4.55, p < .001$).

6.3.3.3 Estimated maximum speed

The differences in timing discussed above may be a result of differences in speed, initiation of movement relative to the platform arrival and gate opening, or a combination of the two. This section explores differences in speed.

Although the momental speed produced by the joystick (set at 3 m/s) was selected to represent normal walking speed scaled by a factor of 2, it may be faster than the actual walking speed selected by walking participants. In addition, there might be some variation of the speed within two joystick conditions, because participants

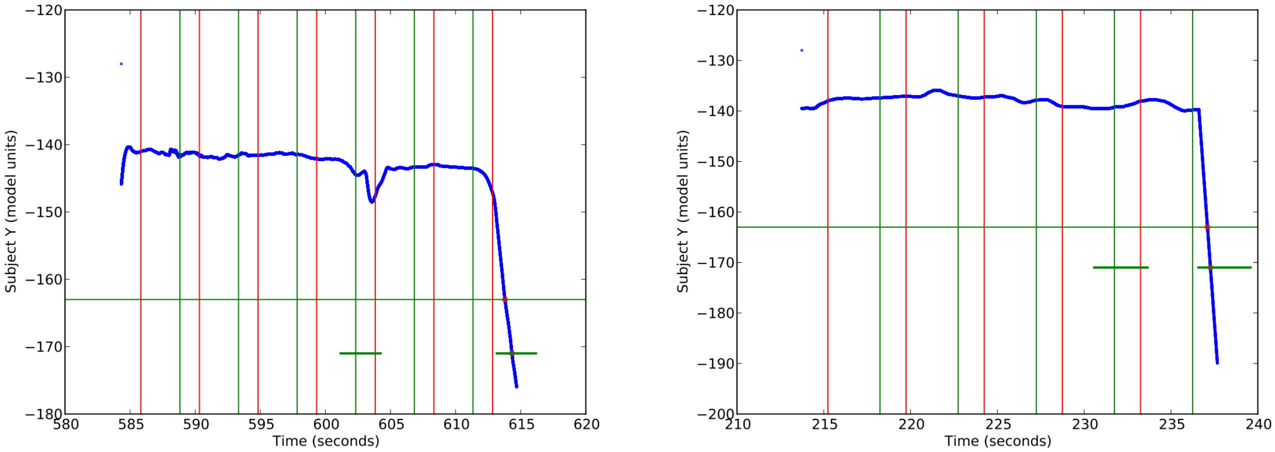

Figure 6.5: Participant's trajectory during a trial in a walking condition (left) and a joystick condition (right). Virtical lines indicate gate openings and closings. Horizontal lines show position of the gate and the edge of the train.

can control average speed by pushing and releasing the joystick. Figure 6.5 shows two representative trajectories illustrating these differences. To compare speed across conditions we used an estimate of participants maximum speed. It was obtained by computing the ration of distance between the gate and the edge of the platform and the time it took the participant to cover this distance.

Similar to timing delays, the speed may differ according to selected configuration of the gate and the platform. So, the comparisons in this sections looked at the speed for "Early" and "Late" configurations separately. Figure 6.6 compares mean *Speed* across two display conditions and two locomotion modalities. Table 6.4 shows mean estimated speed and corresponding standard errors for both configurations.

There is a virtually identical pattern for both configurations. In the "Early" configuration a two-way repeated measures ANOVA indicated significant main ef-

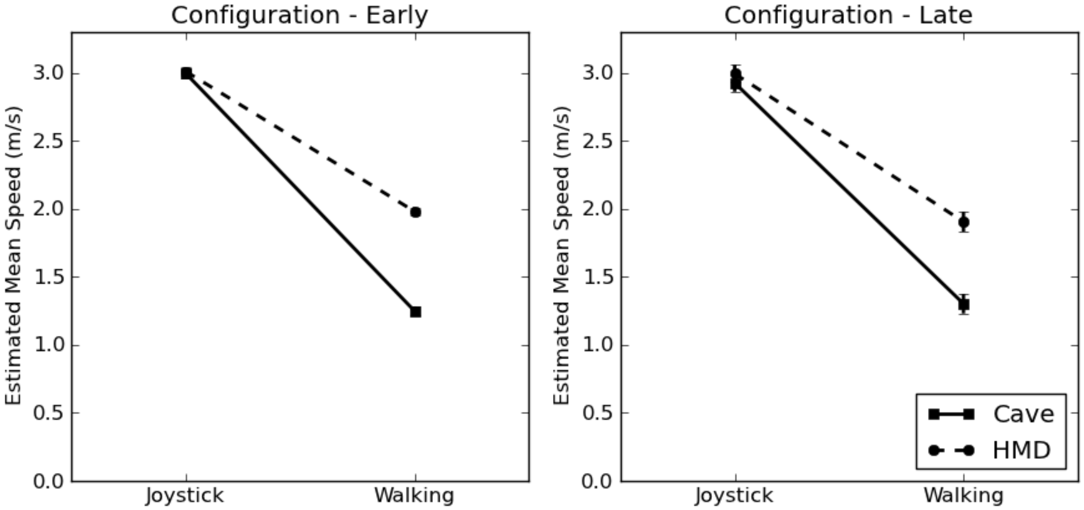

Figure 6.6: Mean estimated speed boarding ”Early” (left) and ”Late” (right) configurations between the gate and the platform. Error bars show standard error.

Table 6.4: Mean estimated speed separated for selection of “Early” and “Late” configuration.

Condition	**“Early”**		**“Late”**	
	Mean	**SE**	**Mean**	**SE**
CAVE + Joystick	3.016	0.056	2.973	0.068
CAVE + Walking	1.29	0.058	1.32	0.068
HMD + Joystick	3.014	0.056	3.017	0.067
HMD + Walking	1.996	0.056	2.027	0.070

fects of display type $F(1,59) = 39.26, p < .001$ and locomotion modality $F(1,59) = 594.92, p < .001$ and a significant two-way interaction $F(1,59) = 39.65, p < 0.001$. The same is true in the "Late" configuration with significant main effects $F(1,56) = 30.41, p < .001$ and$F(1,56) = 376.4, p < .001$ respectively and a significant interaction $F(1,56) = 23.66, p < .001$.

Post-hoc comparisons using Tukey's HSD shows significantly larger speeds for joystick compared to walking in both CAVE ($t_{59} = 21.52, p < .001$ and $t_{56} = 17.18, p < 0.001$ respectively) and HMD ($t_{59} = 12.9, p < .001$ and $t_{56} = 10.27, p < .001$ respectively) display conditions. There was no significant difference between CAVE and HMD display in joystick conditions ($t_{59} = 0.02, p = 1.0$ for the "Early" configuration) and ($t_{56} = -0.47, p = 0.97$ for the "Late"). The means were close to 3 m/s as can be expected. Walking speed in both configurations in HMD condition was faster than in CAVE condition ($t_{59} = -8.81, p < .001$ and $t_{56} = -7.22, p < .001$ respectively for "Early" and "Late" configurations).

These results indicate that participants were walking slower than moving using joystick. In addition there is evidence that walking in HMD was faster than walking in large-screen setup. Finally, participants did not seem to change their speed depending on configuration of selected pairing between the gate and the platform.

6.3.3.4 Door-to-Platform delay

A final measure of participants performance was a *"Door-to-Platform Delay"*. The idea behind this measure is to provide a overall metric of participants timing performance in the boarding task that does not depend on the choice of the configuration

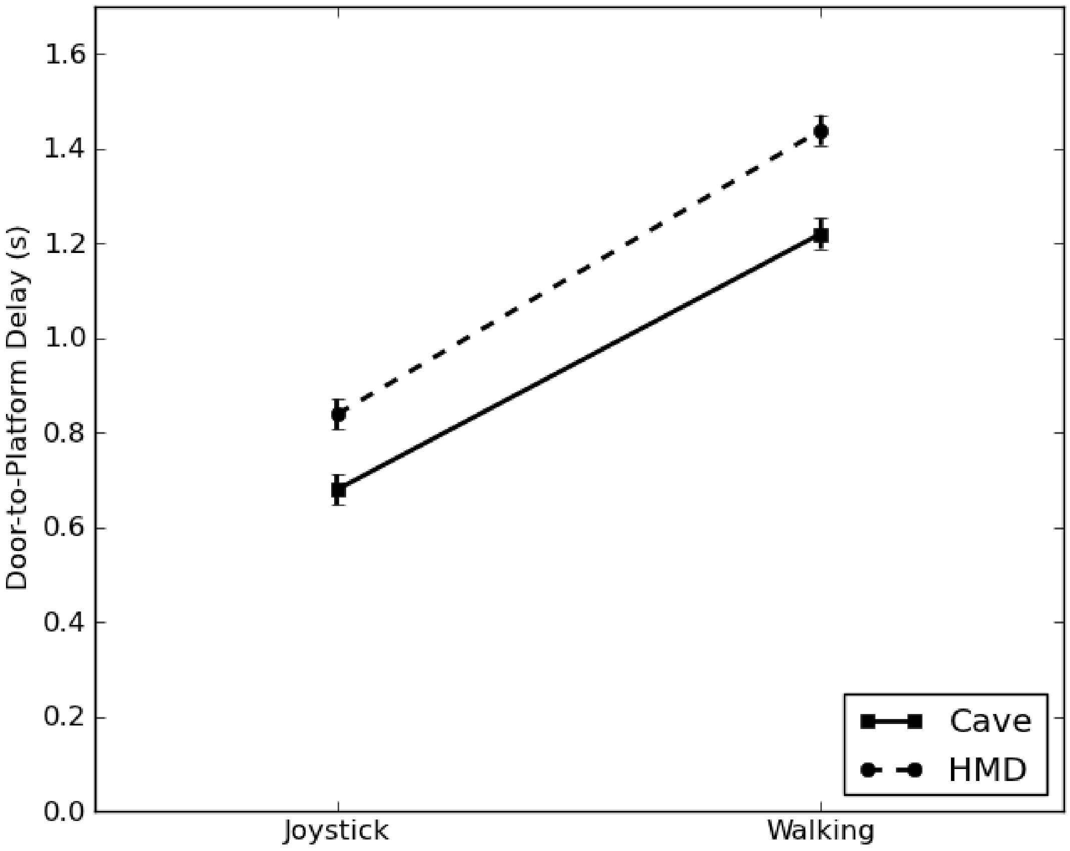

Figure 6.7: Mean door-to-platform delays for each experimental condition. Error bars show standard error.

for boarding. This measure was define as a temporal delay between the moment the gate if fully open and participant's crossing the edge of the platform.

Figure 6.7 compares mean door-to-platform delays for all four experimental conditions. The mean values are summarised in Table 6.5.

The comparison between experimental conditions was done using a two-way repeated-measures ANOVA with door-to-platform delay as dependent measure and

Table 6.5: Mean door-to-plaform delay for all conditions.

Condition	**Mean**	**SE**
CAVE + Joystick	0.679	0.032
CAVE + Walking	1.219	0.033
HMD + Joystick	0.838	0.032
HMD + Walking	1.439	0.032

Display and Modality as independent factors. The ANOVA analysis revealed significant main effects of both Display ($F(1, 1167) = 34.79, p < .001$) and Modality ($F(1, 1167) = 315.35, p < .001$). The interaction was not significant. The follow-up comparisons using Tukey HSD procedure showed that participants in walking conditions significantly delayed their boarding compared to joystick conditions ($t_{1167} = -5.9, p < .001$). In addition the delay in CAVE conditions was smaller than in HMD conditions ($t_{1167} = -17.76, p < .001$).

6.3.4 Sence of presence

To compare participants, reports on their sense of presence during the VR experience, we computed mean score across six questions in the presence questionnaire. Figure 6.8 show mean presence scores for all four condiitons. These means were than analyzed using a two-way ANOVA with mean presence score as a dependent variable and Display and Modality as independent factors. The analysis revealed significant interaction between the effects ($F(1, 59) = 5.29, p = 0.025$), although individual main effects were not significant. The follow up comparison using Tukey HSD procedure revealed that CAVE+Joystick condition had the overall lowest score. This score was

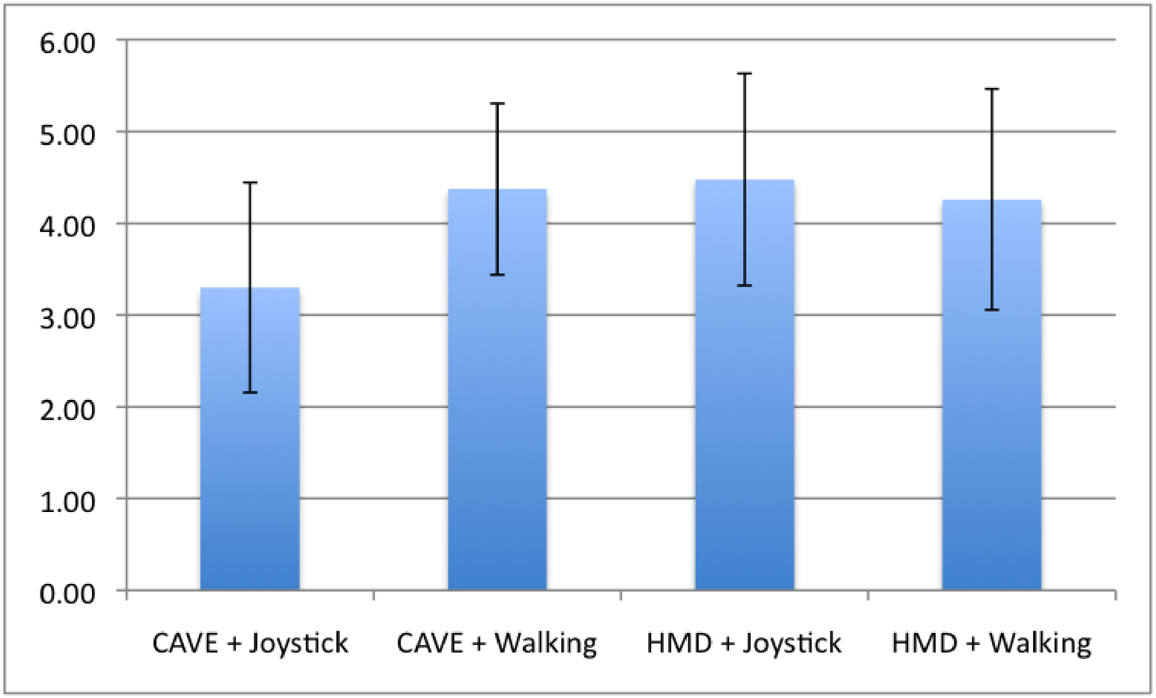

Figure 6.8: Mean scores for sense of presence during the virtual reality experience. Error bars show standard deviation.

significantly lower compared to HMD+Joystick ($p = 0.023$) and CAVE+Walking ($p = 0.045$) conditions. The difference between CAVE+Joystick and HMD+Walking approached significance ($p = 0.088$). The other three conditions (CAVE + Walking, HMD + Walking, and HMD + Joystick) had similar mean scores.

6.4 Discussion

The goal of this experiment was to examine the effects of locomotion modality and display type on participants' perception of affordances in an immersive interactive task that required coordination of full-body self-motion with the motion of objects in virtual environment. In particular, we were interested in how participants timed their movement though the gate while boarding a virtual train and how they selected an opportunity for boarding.

The participants were presented with two nearly identical types of opportunities for boarding the train. The “Early” and “Late” configurations are very similar in terms of affordances. In both cases the gate remained open for exactly the same amount of time and the train platform car was available in front of the participant for the whole duration of that temporal interval. If the participant’s motion capabilities were sufficient to go through the gate while it was open, she was also able to board the train in either case. The ”Early” configuration affords a marginal advantage, because travel between the gate and the train is not instantaneous and the platform car was available for a longer period of time after the gate closed.

The “Early” configuration might appear as a better option. First, the coinciding arrival of the front of the platform and the opening of the gate represent a powerful visual clue that can be used to trigger the motion. Second, the slow moving and less confident participants might be attracted by the longer availability of the platform in the early opening configuration.

6.4.1 Effects of locomotion modality

The effects of locomotion modality appear to be clear. Participants using joystick were able to achieve higher speeds compared to walking participants. Predictably this translated into better timing while boarding the train as evidenced by lower delays (in all three delay measures).

Furthermore, even though joystick action abilities may have been less familiar to participants than naturalistic walking, they were able to perceive own abilities to move faster. This is supported by the fact that the proportion of “Early” and “Late” configurations selected for boarding by participants using joystick was close to 50%. Walking participants were less likely to choose a “Late” configuration, confirming our hypothesis that “Late” configurations may appear difficult to slower moving participants.

Overall these results suggest that participants were able to perceive their action abilities regardless of the type of locomotion interface used in the VR system and are sensitive to changes they experience due to locomotion interface.

6.4.2 Effects of display type

The effects of the display type are somewhat surprising. There seems to be some evidence that participants in CAVE conditions were able to better time the initiation of their movement, when boarding the train. All three delay measures show that lower delays for CAVE conditions compared to HDM conditions. The outcomes are particularly notable for platform entry delay. The participants in CAVE+Walking condition were able to achieve delays similar to those in CAVE+Joystick condition and significantly better than participants in the other two groups. This is particularly noteworthy given the fact that they also had the lowest speed among all four groups. At the same time the type of display had no effect on participants’ choices of boarding opportunities.

Overall, these results suggest that large-screen displays provided a clear performance advantage, when participants need to visually track moving objects. However, participants were not able to accurately detect the effects of changes in their perceptual abilities and showed no evidence of adjusting their choices to capitalize on advantages provided by display.

6.4.3 Limitations

Two minor features of our method may require follow-up examination to broaden the application of our results:

1. The study of locomotion interface we presented here relied on scaled motion. A follow-up study comparing effects of scaled and non-scaled motion may prove to be a useful addition.
2. The software in this study scaled the eye-height of the user incorrectly. Although we have argued that this should not impact the results of the study, an experimental evaluation of this effect would provide a more compelling argument.

6.4.4 Implication for design of VR systems

The experimental results presented here have two implications for the design of VR systems:

1. Users seemed to be able to adapt easily to changes in their action abilities associated with the locomotion interfaces. This implies that the choice of the interface should be based primarily on matching the locomotion abilities to task requirements.
2. The choice of display had an substantial effect on user performance, but users did not show evidence of adapting their choices to changes in perceptual abilities. This suggests that the choice of display should rely on matching the

perceptual abilities to those available in the real world. In that sense the large-screen displays that resemble human's natural vision system more closely seems like a preferable choice.

CHAPTER 7
CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

7.1 Research contributions

In this thesis we attempted to develop a method for conducting user evaluation studies for virtual reality systems that focuses on VR support for immersive, dynamic, full-body interaction experience.

We began with one of traditional evaluation method - perceptual evaluation of distance estimation in real and virtual environments. Chapter 1 presented an experiment demonstrating that the choice of a response action can have a significant impact on the outcome of such perceptual evaluation. We argued for a need to approach perception and action holistically as an intertwined, integrated system.

Chapter two proposed a theoretical framework that uses the notion of affordances to connect perception and action and proposes the perception of affordances as a criterion for evaluation VR systems.

To illustrate the methods of exploring dynamic affordances in VR we described an experiment that studies complex affordances in crossing two lanes of opposing traffic in a bicycle simulator. This study was unique in both the level of realism of the modeled scenario and the complexity of the action-scaled affordance that it explores.

Finally, we formulated some methodological principles and conducted a study that explored the effects of display type and locomotion modality on user performance in a dynamic interaction task through the prism of affordances. The experimental results have two implications for the design of VR systems. First, the users seemed to be able to adapt easily to changes in their action abilities associated with the locomotion interfaces. This implies that the choice of the interface should be based primarily on matching the locomotion abilities to task requirements. Second, the choice of display had an substantial effect on user performance, but users did not show evidence of adapting their choices to changes in perceptual abilities. This suggests that the choice of display should rely on matching the perceptual abilities to those available in the real world. In that sense the large-screen displays that resemble human's natural vision system more closely seems like a preferable choice.

7.2 Limitations and future directions

The experimental study described in Chapter 6 of this thesis represents only a first step of using perception of affordances for evaluation dynamic interactions support in VR systems. Future research should expand it to other types of action-scaled affordance.

Importantly, in this thesis we did not attempt to explore the limits of action-scaled affordances. The task used in chapter 6 relates the temporal interval when the platform is available (and the gate is open) to maximum speed and timing abilities of the participant. By pushing the participant to the limits of her abilities, we may gain better understanding of how well the users can perceive the dynamic affordance in each of the experimental conditions we examined. This understanding may provide further insights into comparative advantages of locomotion interfaces and display technologies. A study of this type would require a slightly different approach to generating action opportunities present to the user. It would likely benefit from an adaptive strategy, that chooses the stimulus presented at each trial based on the previous responses.

In addition, it would be interesting to explore effect of additional sensory feedback such as sound on participants' perception of affordance. For example sounds of train or gate movement may provide useful feedback for participants and enable them to perceive the affordances for action in this task more accurately.

REFERENCES

[1] G. D. Abowd and E. D. Mynatt. Charting past, present, and future research in ubiquitous computing. In J. M. Caroll, editor, *Human-computer interaction in the new millenium*, pages 513–535. ACM Press, New York, NY, 2002.

[2] K. E. Adolph. The growing body in action. In R. L. Klatzky, B. MacWhinney, and M. Behrmann, editors, *Embodiment, Ego-Space, and Action*, pages 275–321. Psychology Press, Taylor & Francis Group, LLC, New York, NY, 2008.

[3] I. V. Alexandrova, P. T. Teneva, S. de la Rosa, U. Kloos, H. H. Bülthoff, and B. J. Mohler. Egocentric distance judgments in a large screen display immersive virtual environment. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization, APGV'10*, pages 57–60, 2010.

[4] M. l. Bhalla and D. R. Proffitt. Visual–motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4):1076–1096, 1999.

[5] B. Bodenheimer, J. Meng, H. Wu, G. Narasimham, B. Rump, T. P. McNamara, T. H. Carr, and J. J. Rieser. Distance estimation in virtual and real environments using bisection. In *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization, APGV '07*, pages 35–40, 2007.

[6] M. Brewer, K. Fitzpatrick, J. Whitacre, and D. Lord. Exploration of pedestrian gap-acceptance behavior at selected locations. *Transportation Research Record: Journal of the Transportation Research Board*, 1982(1):132–140, 2006.

[7] F. P. Brooks. What's real about virtual reality? *IEEE Computer Graphics and Applications*, 19(6):16–27, 1999.

[8] G. Bruder, F. Steinicke, and K.H. Hinrichs. Arch-Explore: A natural user interface for immersive architectural walkthroughs. In *Proceedings of IEEE Symposium on 3D User Interfaces. 3DUI 2009*, pages 75 –82, 2009.

[9] G. Bruder, F. Steinicke, D. Valkov, and K. Hinrichs. Augmented virtual studio for architechtural exploration. In *Proceedings of Virtual Reality International Conference, VRIC 2010*, pages 1–8, 2010.

[10] G. Burdea and P. Coiffet. *Virtual reality technology.* John Wiley & Sons, Inc., Hoboken, NJ, 2nd edition, 2003.

[11] A. Chardenon, G. Montagne, M. J. Buekers, and M. Laurent. The visual control of ball interception during human locomotion. *Neuroscience Letters*, 334(1):13–16, 2002.

[12] A Chardenon, G. Montagne, M Laurent, and R J Bootsma. A robust solution for dealing with environmental changes in intercepting moving balls. *Journal of Motor Behavior*, 37(1):52–62, 2005.

[13] A. Chemero. An outline of a theory of affordances. *Ecological Psychology*, 15(2):181–195, 2003.

[14] B. J. Chihak, J. M. Plumert, C. J. Ziemer, S. Babu, T. Grechkin, J. F. Cremer, and J. K. Kearney. Synchronizing self and object movement: How child and adult cyclists intercept moving gaps in a virtual environment. *Journal of Experimental Psychology: Human Perception and Performance*, 36(6):1535–1552, 2010.

[15] A. B. Craig, W. R. Sherman, and J. D. Will. *Developing Virtual Reality Applications: Foundations of Effective Design*. Morgan Kaufmann, San Francisco, CA, 2009.

[16] S. H. Creem-Regehr, P. Willemsen, A. A. Gooch, and W. B. Thompson. The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. *Perception*, 34(2):191–204, 2003.

[17] J. Cremer, J. Kearney, and P. Willemsen. Directable behavior models for virtual driving scenarios. *Transactions of the Society for Computer Simulation*, 14:87–96, 1997.

[18] J. F. Cremer, J. Severson, S. Gleo, J. K. Kearney, and M. McDermott. “This old digital city” one year later: Experience gained, lessons learned, and future plans. In *Proceedings of the Seventh International Conference on Virtual Systems and Multimedia, VSMM'01*, pages 1–8, 2001.

[19] J. F. Cremer, J. Severson, S. Gleo, J. K. Kearney, M. McDermott, and R. Riccio. “This old digital city”: Virtual historical Cedar Rapids, Iowa circa 1900. In *Proceedings of the Sixth International Conference on Virtual Systems and Multimedia, VSMM'2000*, pages 1–8, 2000.

[20] C. Cruz-Neira, D. Sandin, and T. A. DeFanti. Surround-screen projection based virtual reality:the design and implementation of the CAVE. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '93*, pages 135–142. ACM, 1993.

[21] P. Dourish. *Where the Action Is: The Foundations of Embodied Interaction.* MIT Press, Cambridge, MA, 2001.

[22] A. Drejer and A. Udmundsson. 3D technology in design and production. In P. Anderson and L. Qvortrup, editors, *Virtual Applications: Applications with Virtual Inhabited 3D Worlds*, pages 117–135. Springer-Verlag London Limited, London, UK, 2004.

[23] Klein. E., J. E. Swan II, G. S. Schmidt, M. A. Livingston, and O. G. Staadt. Measurement protocols for medium-field distance perception in large-screen immersive displays. *Proceedings of the IEEE Virtual Reality Conference, VR'09*, pages 107–113, 2009.

[24] B. Fajen and W. Warren. Behavioral dynamics of intercepting a moving target. *Experimental Brain Research*, 180(2):303–319, 2007.

[25] B. R. Fajen. Affordance-based control of visually guided action. *Ecological Psychology*, 19(4):383–410, 2007.

[26] B. R. Fajen. Rapid recalibration based on optic flow in visually guided action. *Experimental Brain Research*, 183(1):61–74, 2007.

[27] B. R. Fajen, M. A. Riley, and M. T. Turvey. Information, affordances, and the control of action in sport. *International Journal of Sport Psychology*, 40(1):79–107, 2009.

[28] B R Fajen and W H Warren. Visual guidance of intercepting a moving target on foot. *Perception*, 33(6):689–715, 2004.

[29] K. P. Fishkin, A. Gujar, B. L. Harrison, T. P. Moran, and R. Want. Embodied user interfaces for really direct manipulation. *Communications of the ACM*, 43(9):74–80, 2000.

[30] W. W. Gaver. Technology affordances. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, CHI '91, pages 79–84, 1991.

[31] J. J. Gibson. *The Ecological Approach to Visual Perception*, chapter The Theory of Affordances. Houghton Mifflin Company, Boston, MA, 1979.

[32] T. Grechkin, T. D. Nguyen, J. Plumert, J. Cremer, and J. Kearney. How does presentation method and measurment protocol affect distance estimation in real and virtual environments? *ACM Transactions on Applied Perception*, 7(4):Article 26, 2010.

[33] T. Y. Grechkin, B. J. Chihak, J. F. Cremer, J. M. Plumert, and J. K. Kearney. Perceiving and acting on complex affordances: How children and adults bicycle across two lanes of opposing traffic. *Journal of Experimental Psychology: Humam Perception & Performance*, (in press), 2012.

[34] D. Guth, D. Ashmead, R. Long, R. Wall, and P. Ponchillia. Blind and sighted pedestrians' judgments of gaps in traffic at roundabouts. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 47(2):314–331, Summer 2005.

[35] D. A. Guttentag. Virtual reality: Applications and implications for tourism. *Tourism Management*, 31(5):637 – 651, 2010.

[36] J. Haase, P. Musaeus, and E. Boisen. Virtual reality and habitats for learning microsurgical skills. In P. Anderson and L. Qvortrup, editors, *Virtual Applications: Applications with Virtual Inhabited 3D Worlds*, pages 29–48. Springer-Verlag London Limited, London, UK, 2004.

[37] R. Hartson. Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology*, 22(5):315–338, 2003.

[38] M. L. Heilig. Stereoscopic-television apparatus for individual use. U. S. Patent 2,955,156, October 4th 1960.

[39] M. L. Heilig. Sensorama simulator. U. S. Patent 3,050,870, August 28 1962.

[40] L. F. Hodges, R. Kooper, T. C. Meyer, J. J. De Graaf, B. O. Rothbaum, D. Opdyke, J. S. Williford, and M. M. North. Presence as the defining factor in a VR application: Virtual reality graded exposure in the treatment of acrophobia. Technical Report GIT-GVU-94-06, Georgia Institute of Technology, 1994.

[41] D. W. J. Hosmer and S. Lemeshow. *Applied Logistic Regression.* John Wiley & Sons, Inc., Hoboken, NJ, 2000.

[42] V. Interrante, B. Ries, J. Lindquist, M.l Kaeding, and L. Anderson. Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments. *Presence: Teleoperators & Virtual Environments*, 17(2):176–198, 2008.

[43] H. Ishii. Tangible bits: beyond pixels. In *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction*, TEI '08, pages xv–xxv, 2008.

[44] R. J. K. Jacob, A. Girouard, L. M. Hirshfield, M. S. Horn, O. Shaer, E. T. Solovey, and J. Zigelbaum. Reality-based interaction: a framework for post-

WIMP interfaces. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 201–210, 2008.

[45] J. A. Jones, J. E. Swan II, G. Singh, E. Kolstad, and S. R. Ellis. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization, APGV '08*, pages 9–14, 2008.

[46] K. S. Jones. What is an affordance? *Ecological Psychology*, 15(2):107–114, 2003.

[47] S. Kenderdine. Speaking in Rama: Panoramic vision in cultural heritage visualization. In F. Cameron and S. Kenderdine, editors, *Theorizing Digital Cultural Heritage: A Critical Discourse*, Media in Transition, pages 301–331. MIT Press, Cambridge, MA, 2007.

[48] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology*, 3(3):203–220, 1992.

[49] A. Kirlik. On Stoffregen's definition of affordances. *Ecological Psychology*, 16(1):73–77, 2004.

[50] E. Kjems. VR for decision support in urban planning. In P. Anderson and L. Qvortrup, editors, *Virtual Applications: Applications with Virtual Inhabited 3D Worlds*, pages 13–28. Springer-Verlag London Limited, London, UK, 2004.

[51] B. R. Kunz, S. H. Creem-Regehr, and Wi. B. Thompson. Evidence for motor simulation in imagined locomotion. *Journal of Experimental Psychology: Human Perception and Performance*, 35(5):1458–1471, 2009.

[52] J. S. Lappin, A. L. Shelton, and J. J. Reiser. Environmental context influences visually perceived distance. *Perception & Psychophysics*, 68(4):571–581, 2006.

[53] A.T. Lee. *Flight Simulation: Virtual Environments In Aviation.* Ashgate, Burlington, VT, 2005.

[54] D. N. Lee, D. S. Young, and C. M. McLaughlin. A roadside simulation of road crossing for children. *Ergonomics*, 27(12):1271–1281, 1984.

[55] N. Lehmann and J. Szatkowski. Virtual narrative, autonomous agents and interactivity in a dramaturgical perspective. In P. Anderson and L. Qvortrup, editors, *Virtual Applications: Applications with Virtual Inhabited 3D Worlds*, pages 49–88. Springer-Verlag London Limited, London, UK, 2004.

[56] J. Loomis, J. Blascovich, and A. Beall. Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods*, 31(4):557–564, 1999.

[57] J. M. Loomis and J. M. Knapp. Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger and M. W. Haas, editors, *Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues*, pages 21–46. Lawrence Erlbaum Associates, Hoboken, NJ, 2003.

[58] W. E. Marsh, M. Putnam, J. W. Kelly, V. J. Dark, and J. H. Oliver. The cognitive implications of semi-natural virtual locomotion. In *Proceedings of IEEE Virtual Reality Conference, VR'2012*, pages 47–50, 2012.

[59] J. McGrenere and W. Ho. Affordances: Calrifying and evolving a convept. In *Proceeding of Graphics Interface 2000*, 2000.

[60] R. Messing and F. H. Durgin. Distance perception and the visual horizon in head-mounted displays. *ACM Transactions of Applied Perception*, 2(3):234–250, 2005.

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

[62] B. J. Mohler, S. H. Creem-Regehr, W. B. Thompson, and H. H. Bülthoff. The effect of viewing a self-avatar on distance judgments in an HMD-based virtual environment. *Presence: Teleoperators & Virtual Environments*, 19(3):230–242, 2010.

[63] P. E. Napieralski, B. M. Altenhoff, J. W. Bertrand, L. O. Long, S. V. Babu, C. C. Pagano, J. Kern, and T. A. Davis. Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Transactions on Applied Perception*, 8(3):18:1–18:19, 2011.

[64] T. D. Nguyen, C. J. Ziemer, J. M. Plumert, J. F. Cremer, and J. K. Kearney. Effects of scale change on distance perception in virtual environments. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, APGV '09, pages 27–34, 2009.

[65] Tien Dat Nguyen. *Estimating Distances and Traveled Distances in Virtual and Real Environments.* PhD thesis, University of Iowa, 2011.

[66] D. A. Norman. *The Psychology of Everyday Things.* Basic Books, New York, NY, 1988.

[67] D. A. Norman. Affordance, conventions, and design. *Interactions*, 6(3):38–43, 1999.

[68] D. A. Norman. *The Design of Everyday Things.* Basic Books, New York, NY, 2002.

[69] M. Oliver. The problem with affordance. *E-Learning*, 2(4):402–413, 2005.

[70] S. O'Neill. *Interactive Media: The Semiotics of Embodied Interaction*, chapter Affordance: A Case of Confusion, pages 49–66. Springer-Verlag London Limited, London, UK, 2008.

[71] R. R. D. Oudejans, C. F. Michaels, F. C. Bakker, and M. A. Dolné. The relevance of action in perceiving affordances: Perception of catchableness of fly balls. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4):879–891, 1996.

[72] J. Oxley, B. Fildes, E. Ihsen, J. Charlton, and R. Day. Differences in traffic judgements between young and old adult pedestrians. *Accident Analysis & Prevention*, 29(6):839–847, 1997.

[73] R. Pausch, J. Snoddy, R. Taylor, S. Watson, and E. Haseltine. Disney's Aladdin: first steps toward storytelling in virtual reality. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, SIGGRAPH'96*, pages 193–203, 1996.

[74] J. W. Philbeck and J. M. Loomis. Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1):72–85, 1997.

[75] T. K. Pitcairn and T. Edlmann. Individual differences in road crossing ability in young children and adults. *British Journal of Psychology*, 91(3):391–410, 2000.

[76] J. M. Plumert, J. K. Kearney, and J. F. Cremer. Children's perception of gap affordances: Bicycling across traffic-filled intersections in an immersive virtual environment. *Child Development*, 75(4):1243–1253, 2004.

[77] J. M. Plumert, J. K. Kearney, J. F. Cremer, and K. Recker. Distance perception in real and virtual environments. *ACM Transactions on Applied Perception*, 2(3):216–233, 2005.

[78] J. M. Plumert, J. K. Kearney, J. F. Cremer, K. M. Recker, and J. Strutt. Changes in children's perception-action tuning over short time scales: Bicy-

cling across traffic-filled intersections in a virtual environment. *Journal of Experimental Child Psychology*, 108(2):322–337, 2 2011.

[79] F. Remondino, S. El-Hakim, S. Girardi, A. Rizzi, S. Benedetti, and L Gonzo. 3D virtual reconstruction and visualization of complex architectures - the "3D-ARCH" project. In *Proceedings of the ISPRS Working Group V/4 Workshop 3D-ARCH "Virtual Reconstruction and Visualization of Complex Architectures"*, pages 1–9, 2009.

[80] G. Reymond and A. Kemeny. Motion cueing in the Renault driving simulator. *Vehicle System Dynamics*, 34(4):249–259, 2000.

[81] A. R. Richardson and D. Waller. Interaction with an immersive virtual environment corrects users' distance estimates. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(3):507–517, 2007.

[82] M. J. Richardson, K. L. Marsh, and R. M. Baron. Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 33(4):845–859, 2007.

[83] F. P. Rivara and M. Aitken. Prevention of injuries to children and adolescents. *Advances in pediatrics*, 45:37–72, 1998.

[84] R. A. Ruddle and S. Lessels. The benefits of using a walking interface to navigate virtual environments. *ACM Transactions on Computer-Human Interaction*, 16(1):5:1–5:18, 2009.

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

[86] R. A. Ruddle, E. Volkova, and H. H. Bülthoff. Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Transactions on Computer-Human Interaction*, 18(2):10:1–10:20, 2011.

[87] B. Sanstrom and P. Longorio. Innovative 3D visualization tool promotes development-drilling efficiency. *Oil & Gas Journal*, 100(8):79–83, 2002.

[88] R. Satava and S. Jones. Medical applications of virtual reality. In L. J. Hettinger and M. W. Haas, editors, *Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues*, pages 325–343. Lawrence Erlbaum Associates, Inc. Publishers, Mahwah, NJ, 2003.

[89] O. Shaer and E. Hornecker. Tangible user interfaces: Past, present, and future directions. *Foundations and Trends in Human-Computer Interaction*, 3(1-2):1–137, 2009.

[90] W. R. Sherman and A. B. Craig. *Understanding virtual reality : interface, application, and design.* Morgan Kaufmann, San Francisco, CA, 2003.

[91] T. A. Stoffregen. Affordances as properties of the animal-environment system. *Ecological Psychology*, 15(2):115–134, 2003.

[92] T. A. Stoffregen. Breadth and limits of the affordance concept. *Ecological Psychology*, 16(1):79–85, 2004.

[93] T. A. Stoffregen, K. M. Gorday, Y. Sheng, and S. B. Flynn. Perceiving affordances for another person's actions. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1):120–136, 1999.

[94] D. Strickland. A virtual reality application for autistic children. *Presence: Teleoperators & Virtual Environments*, 5(3):319–329, 1996.

[95] E.A. Suma, S.L. Finkelstein, S. Clark, P. Goolkasian, and L.F. Hodges. Effects of travel technique and gender on a divided attention task in a virtual environment. In *Proceedings of IEEE Symposium on 3D User Interfaces, 3DUI'2010*, pages 27–34, 2010.

[96] I. Sutherland. The Ultimate display. In *Proceedings of 1965 IFIP Congress*, pages 506–508, 1965.

[97] I. Sutherland. A head-mounted three-dimensional display. In *Proceedings of the AFPIS Fall Joint Computer Conference*, volume 33, pages 757–764, 1968.

[98] D. S. Tan, D. Gergle, P. Scupelli, and R. Pausch. Physically large displays improve performance on spatial tasks. *ACM Transactions on Computer-Human Interactions.*, 13(1):71–99, 2006.

[99] A. F. te Velde, J. van der Kamp, J. A. Barela, and G. J. P. Savelsbergh. Visual timing and adaptive behavior in a road-crossing simulation study. *Accident Analysis & Prevention*, 37(3):399–406, 5 2005.

[100] A. F. te Velde, J. van der Kamp, and G. J. P. Savelsbergh. Five- to twelve-year-olds' control of movement velocity in a dynamic collision avoidance task. *British Journal of Developmental Psychology*, 26(1):33–50, 2008.

[101] W. B. Thompson, P. Willemsen, A. A. Gooch, S. H. Creem-Regehr, J. M. Loomis, and A. C. Beall. Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators & Virtual Environments*, 13(5):560–571, 2004.

[102] G. Torenvliet. We can't afford it!: the devaluation of a usability term. *Interactions*, 10(4):12–17, 2003.

[103] M. T. Turvey. Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, 4(3):173–187, 1992.

[104] B. Ullmer and H. Ishii. Emerging frameworks for tanglible user interfaces. In J. M. Caroll, editor, *Human-computer interaction in the new millenium*, pages 579–601. ACM Press, New York, NY, 2002.

[105] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks Jr. Walking >walking-in-place >flying, in virtual environments. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '99*, pages 359–364, 1999.

[106] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionnaires in reality. *Presence: Teleoperators & Virtual Environments*, 9(5):497–503, 2000.

[107] F. A. Voorhorst, C. J. Overbeeke, and G. J. F. Smets. Implementing perception-action coupling for laparoscopy. In L. J. Hettinger and M. W. Haas, editors, *Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues*, pages 391–411. Lawrence Erlbaum Associates, Inc. Publishers, Mahwah, NJ, 2003.

[108] W. H. Warren. Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, 10(5):683–703, 1984.

[109] J. A. Weast, K. Shockley, and M. A. Riley. The influence of athletic experience and kinematic information on skill-relevant affordance perception. *The Quarterly Journal of Experimental Psychology*, 64(4):689–706, 2011.

[110] M. Weiser. The computer of the 21st century. *Scientific American*, 265(3):66–75, 1991.

[111] S. L. Whitney, P. J. Sparto, L. F. Hodges, S. V. Babu, J. M. Furman, and M. S. Redfern. Responses to a virtual reality grocery store in persons with and without vestibular dysfunction. *CyberPsychology & Behavior*, 9(2):152 – 156, 2006.

[112] M.C. Whitton, J.V. Cohn, J. Feasel, P. Zimmons, S. Razzaque, S.J. Poulton, B. McLeod, and Jr. Brooks, F.P. Comparing ve locomotion interfaces. In *VR '05: Proceedings of the IEEE conference on Virtual Reality*, pages 123–130, March 2005.

[113] P. Willemsen, J.K. Kearney, and H. Wang. Ribbon networks for modeling navigable paths of autonomous agents in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 12(3):331–342, 2006.

[114] Peter Willemsen, M. B. Colton, S. H. Creem-Regehr, , and W. B. Thompson. The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments. *ACM Transactions on Applied Perception*, 6(2):1–14, February 2009.

[115] B. G. Witmer and W. J. Sadowski. Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Human factors*, 40(33):478–488, 1998.

[116] B. Wu, T. Ooi, and Z. He. Perceiving distance accurately by a directional process of integrating ground information. *Nature*, 428:73–77, 2004.

[117] T L Yang, M W Dixon, and D R Proffitt. Seeing big things: overestimation of heights is greater for real objects than for objects in pictures. *Perception*, 28(4):445–467, 1999.

[118] D. S. Young and D. N. Lee. Training children in road crossing skills using a roadside simulation. *Accident Analysis & Prevention*, 19(5):327–341, 10 1987.

[119] C. J. Ziemer, M. J. Branson, J. K. Kearney, J. F. Cremer, and J. M. Plumert. Recalibrating perception and action: Examining the role of the recalibration experinece, distance estimation task, and the testing environment. *(Manuscript in preparation)*, 2012.

[120] C. J. Ziemer, J. M. Plumert, J. F. Cremer, and J. K. Kearney. Estimating distance in real and virtual environments: Does order make a difference? *Attention, Perception, & Psychophysics*, 71(5):1095–1106, 2009.