2012

Identifying and mitigating the cognitive implications of semi-natural virtual locomotion techniques

William Marsh
Iowa State University

Identifying and mitigating the cognitive implications of semi-natural virtual locomotion techniques

by

William Eric Marsh

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Co-majors: Human Computer Interaction;
Computer Engineering

Program of Study Committee:
James H. Oliver, Major Professor
Veronica J. Dark
Julie Dickerson
Jonathan W. Kelly
Les Miller

Iowa State University
Ames, Iowa
2012

Table of Contents

List of Figures

List of Tables

Acknowledgements

It is a pleasure to express my gratitude to all of the people who have contributed to my success as a graduate student. First I must thank my advisor, Dr. Jim Oliver, and the rest of my committee for their hours of guidance with respect to research, writing, and university life in general. I could always count on plenty of useful feedback when analyzing data and writing, and the papers benefited greatly.

In addition to my committee, I would like to thank all of the people from around the VRAC and Iowa State University who have helped me out along the way. I must thank Kevin Godby for his help on software engineering and, of course, his LaTeX feedback in the preparation of this document. The fuzzy logic aspects of this dissertation have been significantly improved with feedback from Rafael Radkowski. I thank my friend Mike Oren for his guidance and conversation on university matters. I would also like to thank Jon Hobbs from the Department of Statistics for his help with R and statistics throughout all of my experiments. I have worked with many undergraduate students over the past two years and I appreciate their hard work that has greatly simplified my life. Of course none of this would be possible without the many people who help keep VRAC running, which includes maintaining the C6, making sure I get paid, and directing me to free food.

I would like to thank Dr. Forouzan Golshani, who has been a friend and a mentor in some capacity at each stage of my education. Without his guidance, I might not have gone to graduate school in the first place.

I thank my parents, Richard and Elizabeth Marsh, for their support throughout my academic endeavors.

Finally, I thank my girlfriend, Tiffany Tisler, for her endless support and her dedication to keeping me on task so I can graduate one day.

Abstract

Users of virtual reality systems often need to navigate to distant parts of the virtual environment in order to perform their desired tasks. Unfortunately, physical space restrictions as well as tracker range limitations preclude the use of fully natural techniques for navigation through an infinite virtual environment. This necessitates the use of a locomotion interface, and the closer that interface matches the analogous real world actions, the easier it will be for the user. Unnatural techniques require cognitive effort on the part of the users. Many authors have attempted to address this problem by creating locomotion interfaces and techniques that more closely approximate real world counterparts to the extent possible. In addition to requiring these unnatural movements, current virtual reality systems are incapable of providing the high-fidelity sensory feedback used to guide real-world movements. This may cause users to resort to more cognitively demanding strategies.

There is a large body of research in the psychology domain regarding the structure of cognitive resources. In particular, Baddeley's multi-component model of working memory describes a separation between the resources used for verbal and non-verbal storage and processing. It is likely that semi-natural locomotion techniques require some of these resources, which will then be unavailable for concurrent tasks. A pair of studies was conducted, investigating the cognitive resource requirements of several atomic locomotion movements by manipulating the user interface and field of view. The results indicate that semi-natural locomotion interfaces generally require a user's spatial cognitive resources. Based on the conclusions from the working memory studies, an adaptive system was designed that can learn how to adjust parameters of the locomotion technique according to a user's present cognitive task load.

Chapter 1. Introduction

Interaction with a virtual environment (VE) often involves compromises as system designers attempt to provide a "natural" interface for interaction with the virtual world, while confined by the physical constraints inherent in the hardware and software implementations of that interface. Natural interfaces are generally defined as those that use techniques that are similar to real world movement. This means that the same body segments should be used to actuate the interface and virtual control actions should have similar cause-effect relations as in the real world (Templeman & Sibert, 2006; Wells, Peterson, & Aten, 1996). A natural interface is more transparent to the user, enhancing the sense of immersion and potentially increasing the effectiveness of the VE.

The tasks that a user wishes to perform often require navigation to distant parts of the VE. Virtual navigation requires the use of atomic locomotion actions while building, maintaining, and using a spatial model. It is not possible to navigate an infinite VE from within a finite virtual reality (VR) system using only completely natural locomotion techniques. Unnatural locomotion mechanisms are often necessary but they can negatively impact user experience and success. An important concept is that navigation, and thus locomotion, is seldom the purpose of a VE. Users often wish to navigate to remote virtual locations while performing other primary tasks along the way (Bowman, Kruijff, LaViola, & Poupyrev, 2005; Darken, Cockayne, & Carmein, 1997). Figure 1.1 shows a hypothetical learning curve and illustrates how locomotion, navigation, and the primary task combine to form overall performance. Before a user can perform a task, basic navigation must be learned. Before navigating, the user must figure out how to perform basic locomotion actions using the provided interface.

Both physical and cognitive aspects must be considered when choosing a locomotion interface to be compatible with given primary tasks. For example, if locomotion requires the use of a handheld controller, the hands will be unavailable for other tasks, such as grabbing or gesturing (Wells et al., 1996). Working memory, mediated by conscious attention, is thought to be required when humans learn a novel skill. It seems that unnatural locomotion mechanisms may require working memory to maintain a model mapping possible actions to expected outcomes. Working memory resources used for locomotion cannot be directed to

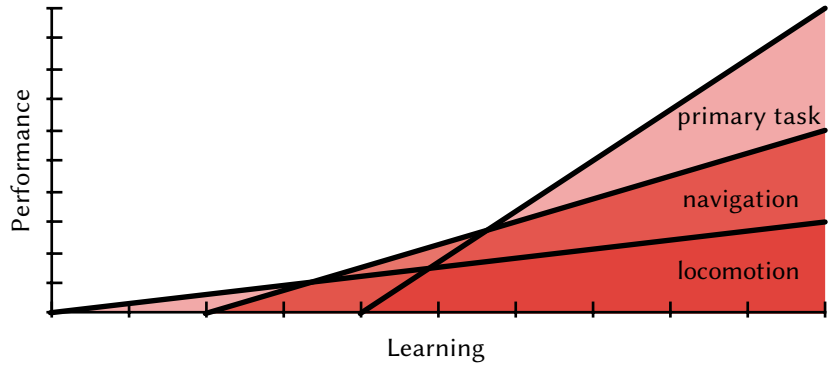

Figure 1.1 Hypothetical learning curve, showing how the primary task, navigation, and locomotion combine to form overall performance.

completing a simultaneous primary task. Similarly, working memory that is in use by primary tasks is generally unavailable for locomotion (Gopher & Donchin, 1986). Figure 1.2 depicts the flow of information and competition for resources when performing tasks in a VR system.

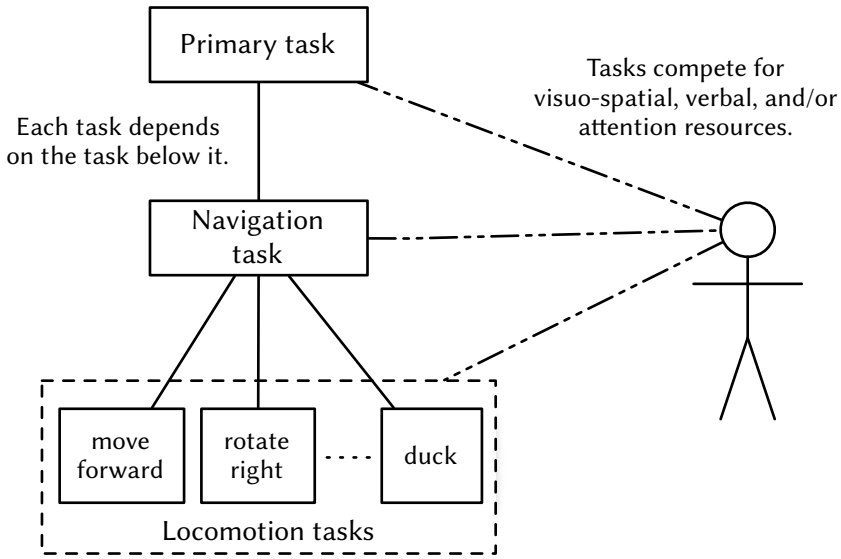

Figure 1.2 Flow of information and competition for resources when performing tasks in VR.

Clearly, aspects of the VR system may impact navigation through an environment in terms of spatial performance. For example, a system with a low-resolution display may not provide enough information for a user to adequately differentiate landmarks. Additionally, some system specifications may also impact basic locomotion abilities. For example, the field of view (FOV) afforded by the display may change the working memory requirements and, in

turn, affect performance. Humans use visual information, in conjunction with other sensory input, to judge distance traveled and relative orientation (Gibson, 1958). A limited FOV could limit the availability of these cues, possibly causing a user to resort to other, cognitively demanding, strategies to encode and manipulate such information.

Understanding the bi-directional impact of different types of cognitive tasks during a simultaneous locomotion task can inform the design of locomotion interfaces and VR systems. Additionally, this understanding will motivate the design of an adaptive locomotion interface that will use its knowledge of the user's current abilities, given a particular parallel task load, to maximize user performance. In doing so, this human-computer interaction research bridges the fields of psychology and computer engineering.

This was accomplished by first conducting studies related to basic psychology phenomena and using the results to inform the design of a software system that can adapt in response to a user's changing workload. The flow of the research described in this dissertation is shown in Figure 1.3. The results from each study informed the details of the next step. The primary components of this research are as follows.

1. Study 1. This study incorporated a dual-task paradigm to compare gamepad-based locomotion with a more natural body-based locomotion interface. Information was gained regarding which aspects of locomotion (rotation, forward, side-step, ducking, stopping, etc.) suffer from different types of simultaneous working memory load. The expected result was that a concurrent spatial task would cause greater detriment to users of the gamepad as compared to users of the body-based interface, because the gamepad is less natural. However, it was expected that certain aspects of the body-based interface would also be greatly impaired.

2. Study 2. This study investigated the possibility that locomotion in systems that provide a reduced FOV requires more working memory resources than locomotion using a similar system with a high FOV. In particular, it was hoped that this study would reveal which particular resource pools were used and to what extent aspects of locomotion suffered during concurrent cognitive tasks. In this study, the body-based locomotion interface was used by all participants.

3. Adaptive System Implementation. An adaptive system was designed to exploit the findings from the first two studies and knowledge of a user's current cognitive task load to adjust locomotion parameters in real-time. The system learns how to adjust the parameters of the body-based interface with the objective of maximizing user locomotion performance given current cognitive resource utilization while still allowing for infinite virtual locomotion from within a constrained physical space.

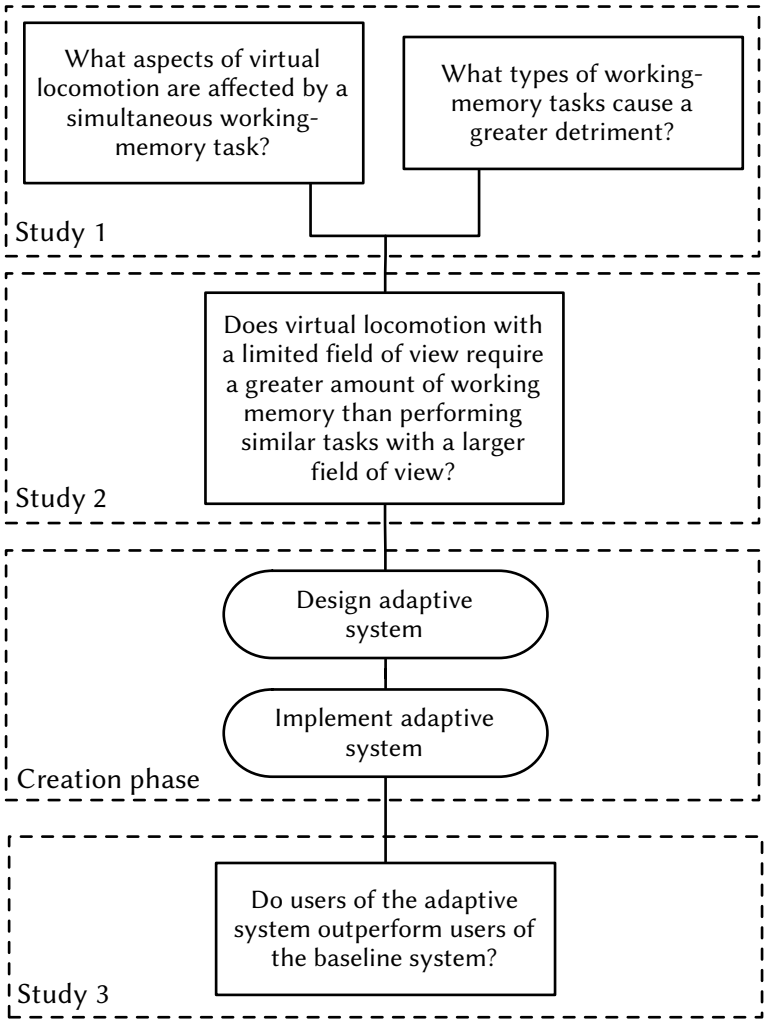

Figure 1.3 Flow of research described in this dissertation. Results from each study were used to inform the details of the next step.

4. Study 3. This basic study was intended to test performance when using the adaptive system described above. The new system was compared to the baseline, non-adaptive body-based interface, specifically checking if the adaptation helped users in the ways intended.

Working memory studies

First a pair of studies was conducted to explore the use of working memory in virtual locomotion. In all studies, the dual-task selective-interference paradigm was used to investigate the impact of simultaneous working-memory-intensive tasks on locomotion ability. The work-

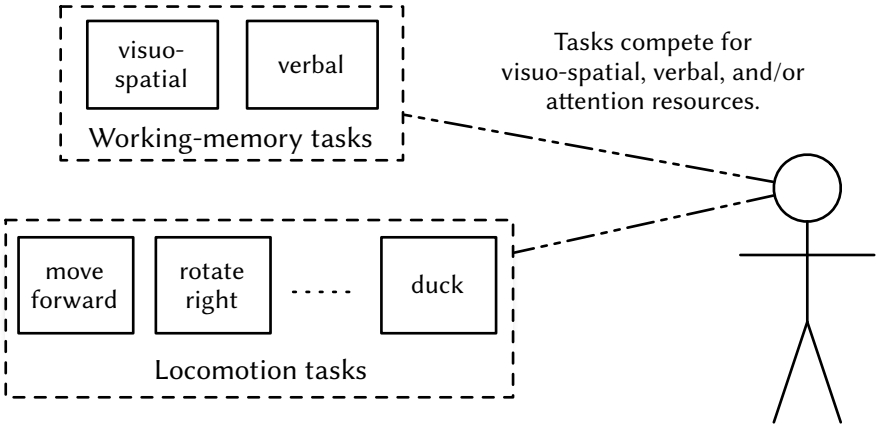

Figure 1.4 Competition for resources in working-memory studies. Large-scale navigation is not required.

ing memory tasks were chosen strategically to tax specific cognitive resources (spatial or verbal) and provide insight into which ones are required for locomotion (spatial, verbal, or more general attention resources). Figure 1.4 shows the competition for cognitive resources that was expected to occur in the working memory studies. Notice that the studies involve no large-scale navigation task, as would be typical in a practical VE application, because this research is focused on working memory use during basic locomotion actions. The working memory tasks were intended to simulate the existence of primary tasks, as might be performed in a "real-world" VE, with known resource demands. Participants were also given a perspective-taking test beforehand to check if individual differences impacted results. It was expected that, for example, a participant's abilities on this test may relate to the ability to predict the results of a locomotion movement, possibly affecting the choice of strategy and/or the cognitive demands of a given movement.

In the dual-task selective-interference paradigm, participants are given a working memory task with known working memory requirements to perform alongside a second task of interest. If concurrent completion of the working memory task causes decreased performance at the task of interest, then it can be concluded that interference exists and the two tasks rely on the same cognitive resources. Specifically, concurrent tasks are known to tax either the spatial or verbal pools of working memory resources. For example, if performance on a task of interest declines when a participant is also performing a spatial working memory task but is not affected by a simultaneous verbal task, then it can be concluded that the task of interest also requires spatial resources. If either a spatial or a verbal secondary task causes an equivalent performance decrease, then it may be concluded that the task relies on general

attention resources. Alternately the task may have equal verbal and spatial working memory demands.

When using the dual-task selective-interference paradigm, it is important to verify that participant performance on the working memory task remains high. If performance drops on the working memory task then it can still be concluded that the task of interest requires the resource in question. It simply means that the user allocated resources to the task of interest as opposed to the working memory task.

Some models of working memory combine visual and spatial resources into a single pool, while many recent models draw a distinction. Because unnatural locomotion is essentially an unlearned skill, working memory should be used in its performance. More specifically, it intuitively seems that spatial working memory should be used for this purpose. Additionally, locomotion certainly relies on visual feedback and when movement is completely natural, position can likely be updated automatically through perceptual processes. However, when using an unnatural interface, visual working memory resources may be required to handle this feedback. Because the focus was on working memory use during skill acquisition, a spatial span task was used in this study. Before each locomotion scenario, a card was displayed, upon which a randomly ordered sequence of boxes was displayed at different spatial locations. The participant was instructed to remember the order in which the boxes were displayed. After the locomotion actions, letters were randomly displayed at all possible locations from the initial display. Because the participant will have moved since the initial presentation, these letters were presented relative to the user . The participant was prompted to state, in order, the sequence of letters corresponding to the locations of the boxes in the remembered sequence.

The verbal memory task was a span task similar to the spatial task described above except that the cues were verbal. Each participant saw a random sequence of numbers before each locomotion action. Afterwards, the participant was prompted to recite the numbers from memory in the order in which they were presented.

These tasks must be configured to ensure that they have a similar level of difficulty. Based on the pilot study conducted with help from Research Experience for Undergraduates (REU) students in the summer of 2010, a high degree of difference in abilities between participants was expected. Because both of the tasks required remembering a sequence, the length of that sequence should be customized to each participant to achieve a fixed, high level of performance in isolation of any other tasks. This was accomplished by adjusting the sequence length during a practice round to ensure that the user could successfully remember 100% of the items. If performance is lower, the task might be too hard, which might lead a participant not to try hard enough. If the task is too easy, then participants might not need to fully tax

the working memory resource in question.

Adaptive system

Based on the results from the working memory studies described above, an adaptive system was created. This system is intended to help users of the body-based interface stop quicker while also increasing the extent to which locomotion is natural.

Fuzzy system

The adaptive system used fuzzy logic to map the user's current spatial and verbal working memory loads (both represented as fuzzy sets) to parameter values for the body-based interface. Starting values for the fuzzy terms were set initially based on expert knowledge gained from interpreting the results from the first two studies.

Learning

These initial settings were intended to generally help users, but each user is different and has a different level of experience. To address this, the system tracked the following user performance metrics in order to improve itself by adjusting the fuzzy terms.

Collisions. When the user's virtual body comes into contact with the virtual walls.

Stop time. How long it takes the user to come to a complete stop.

Percent of interface utilization. How much of the physical space was being used.

Adaptive system study

The new adaptive system was tested in a formal user study. Participants were placed into one of two groups according to locomotion interface: normal body-based or adaptive body-based. Analysis focused on performance using each interface and also differences in participant learning in terms of the performance metrics. Two other measures of performance, physical distance traveled and virtual distance traveled, were used to assess the efficiency of participants' movements.

Chapter 2. Literature Review

Many authors have acknowledged and attempted to address the limitations posed by interfaces for virtual locomotion. Much of the previous research has attempted to address the limitations by creating more natural interfaces, in some cases exploiting limitations of human perception. However, it is generally accepted that none of the current interfaces are truly natural and achieving such an ideal may be impossible.

Virtual reality

Virtual reality (VR) usually refers to some combination of immersive stereoscopic graphics display, tracking, interaction devices, and sometimes aural (audio) and haptic (touch/force) interfaces intended to evoke a sense of presence in a virtual environment (VE). Tracking is a particularly important factor in producing an "egocentric" user experience because the graphics view frusta are continuously updated based on the user's head position. Virtual reality is used in many domains for a wide range of activities, spanning training, the natural sciences, and the humanities. Frequently, these activities involve using a navigation interface for traveling through a VE to perform tasks. Immersive graphics displays are typically implemented with either a head-mounted display (HMD), essentially small screens mounted relatively close to the eyes, or a CAVE (Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992), a room-sized display with walls and floor illuminated with stereoscopic graphics. CAVEs generally boast higher-fidelity graphics and larger field of view than HMDs but they are typically much more expensive. However, both types of displays offer limited physical mobility of users, due to tracking system constraints, and, in the case of CAVEs, the walls of the device itself.

Virtual navigation

It is useful to speak of interaction within an environment in terms of three scales of space. In fact, there is evidence that processing may even be scale dependent (Sholl, 1998). The following scales of space have been previously proposed (Montello, 1998).

Figural. Small relative to body.

Vista. Can mostly be seen from a given vantage point with a possibility of minor movements.

Environmental. Large relative to body.

It is generally not necessary to move about in the figural or vista space so navigation is not necessary. However, navigation is often required in the environmental space. In fact, environmental spaces cannot be seen from a single vantage point, thus requiring movement. Spatial learning therefore requires a person to mentally store environmental information (Ittelson, 1973).

Similarly, interaction generally involves some combination of the following spatial tasks (Kulik, 2009).

Exploration. Looking around, changing movement direction, and observing the environment.

Search. Tends to involve relatively straight trajectories while performing a systematic coverage of an area.

Maneuvering. Usually making fine movements while interacting with a single object or small area of interest within the environment (Darken et al., 1997).

Virtual locomotion

In this paper, the term "virtual locomotion" is used to refer to the atomic movements that a person makes when navigating through a VE. There are two basic competing objectives when designing an interface for virtual locomotion:

1. allow for navigation between any two points in the VE; and
2. maximize naturalness.

A locomotion technique should maximize the match between proprioceptive information corresponding to actions and sensory feedback generated by the VR system. A good match will allow the user to learn a predictive model of interaction within the environment (Slater, Usoh, & Steed, 1994). An untrained user already possesses natural perceptual-motor abilities and knowledge of interaction in the real world. Therefore, when designing an effective interface, it is beneficial for virtual techniques to replicate how humans usually move about in the real world (Kulik, 2009; Wickens & Baker, 1995). This includes use of the same body segments and a similar level of effort. These requirements will both benefit transfer

of training. This means that a virtual walking technique should ideally use the legs and be about as effortful as walking in the real world, meaning that performance will be limited by the strength and agility of the user (Templeman, Denbrook, & Sibert, 1999; Templeman & Sibert, 2006). There are some compelling reasons why natural locomotion techniques are usually preferred. The choice of movement technique has been shown to have a significant impact on cognition (Zanbaka, Lok, Babu, Ulinski, & Hodges, 2005) and presence, the subjective sense of being in a VE (Slater, Steed, McCarthy, & Maringelli, 1998; Slater et al., 1994; Usoh et al., 1999). There is also an impact on spatial abilities. Real walking seems to provide the translational and rotational information needed to accurately update position automatically with perceptual processes (Klatzky, Loomis, Beall, Chance, & Golledge, 1998). Finally, the use of natural locomotion techniques provides ecological validity for many of the types of studies and training exercises for which VR systems are used. Natural techniques may cause users to use movement strategies based on real-world experience, increasing transfer of skills learned in VR to the real world (Templeman et al., 1999). Templeman and Sibert (2006) used driving as an example:

> Learning just the rules and strategies of driving is not enough. A driving simulator needs a dashboard, steering wheel, shift controls, control knobs, pedals, a panoramic windshield, and rear-view mirrors because to learn to drive a car, you need to learn to coordinate actions as well. A driver must coordinate actions while applying the rules of the road.

Using wide-area trackers to allow for real walking is the ideal solution in terms of ease and naturalness (Usoh et al., 1999). However, it can be challenging to meet both of the stated objectives simultaneously as the confines of the physical environment often limit the use of natural locomotion techniques to navigate through an infinite VE. In some cases, actual system boundaries, such as walls, get in the way. In other systems, the range of user tracking hardware restricts allowable physical movement. Overcoming these constraints often involves scaling and automation (Kulik, 2009). Real world properties, such as visual-vestibular coupling, are often violated by the system (Wickens & Baker, 1995). A wide variety of locomotion mechanisms have been implemented and studied.

A basic feature of all locomotion techniques is control order. When using zero-order controls, an input by the user will produce a specific change in the user's location in the VE. Zero-order controls, such as walking, are good for precise positioning (Wickens & Baker, 1995) and such interfaces are often referred to as "position control." When using a first-order, or "rate-control," controller, a control input will produce a change in the user's virtual velocity. First-order controls are found in most hand-operated controllers and are useful for

traveling over long distances. They work best when visual feedback is present (Sibert et al., 2008). First-order interfaces have a characteristic called control-display gain, that refers to sensitivity in terms of how control input is mapped to output movement (MacKenzie, 1995).

There are several handheld locomotion devices currently in use. These involve hardware such as a wand, DataGlove, or joystick. Wands and DataGloves are often used with a push-button-fly metaphor where the user presses a button to travel in the direction in which the device is pointing (Wells et al., 1996). Joysticks are so standard that it makes sense to consider them as baselines for comparison to alternate techniques. These handheld devices are not usually considered natural because they use completely different muscle groups than real walking, yet they are often appropriate when coupled with a flying metaphor. Additionally, handheld devices do not facilitate spatial understanding as users are not good at judging distances while flying (Gibson, 1958).

Body-based locomotion techniques can often allow for much more natural movement because they incorporate the same muscle-groups used in the analogous physical-world locomotion activities. In recent years, much research has been devoted to the development of hardware-based solutions, such as treadmills (Christensen, Hollerbach, Xu, & Meek, 2000; Darken et al., 1997; Iwata, 1999; Wang, Bauernfeind, & Sugar, 2003), unicycles (Darken et al., 1997), and even large hamster balls (Medina, Fruland, & Weghorst, 2008) that allow for semi-natural movement within a constrained physical space. These solutions show promise but they are still expensive, inflexible, and they fail to produce 100% natural interaction. But it is also possible to simulate infinite walking using software-based solutions coupled with hardware to track the physical location and movements of the user. Walking in place is a semi-natural technique in which the system monitors head or body movements to detect steps (Templeman et al., 1999). Redirected walking (Razzaque, Kohn, & Whitton, 2001) and motion compression (Engel, Curio, Tcheang, Mohler, & Bulthoff, 2008) are both software-based approaches to exploiting limitations of the vestibular and proprioceptive systems. The brain combines information from multiple senses to form an estimate of reality. When the senses report conflicting information, vision is usually trusted because it has the highest acuity and it may have shorter latency (Mohler, Thompson, Creem-Regehr, Pick, & Warren, 2007). This allows for a user to be "redirected" along a curved physical path while following a straight path in the VE. Similarly, motion can be compressed so that steps in the real world produce slightly larger than normal movements in the VE. The result of each of these software-based techniques (or a combination of the two) is that a user can navigate a VE that is larger than the physical world. Redirected walking and motion compression are considered to be very natural because the user is actually walking and does not usually notice the illusion. However, it is less clear how the conflicting perceptual information may subconsciously

impact factors involving navigation because it has been shown that the vestibular sense of motion contributes to these activities (Chance, Gaunet, Beall, & Loomis, 1998). Walking-in-place is another common body-based technique that relies on both hardware and software. Hardware is often required to keep the user centered and software is used to convert head motion from walking-in-place into virtual movement. These leg-based techniques provide some kinesthetic feedback and minor cyclical vestibular feedback of real walking, but they still fail to provide any translational information (Sibert et al., 2008).

In addition to ease of use, another benefit of using semi-natural, body-based techniques is that they tend to help with path integration. Path integration allows a person to update current position and orientation based on velocity and acceleration information. This ability to use kinesthetic and vestibular input to augment vision is helpful in situations with low visibility (Sibert et al., 2008).

It is sometimes desirable to mix real-world locomotion techniques with less natural virtual techniques (Templeman et al., 1999). In this vein, hybrid rate/position-control systems have been created in which locomotion is natural to the extent that a given VR system allows. In the Virtual Motion Controller (Wells et al., 1996), for example, the user wears a belt that is suspended by elastic straps. When the user moves around the center of the physical environment all movements are natural as in the "real" world, allowing for precise position control. When the user steps past a certain threshold distance from the center, the interface becomes rate-controlled and the user's virtual velocity increases as a function of the distance from the physical center. As the user moves away from the center, resistance increases due to the elastic straps. The vector from the center to the user sets the virtual travel direction. This scheme allows for rapid movement over large distances as well as natural fine-position control. In addition to the fact that not all movement is completely natural, problems with this interface are largely due to the bulky hardware. In this case, the hardware configuration typically requires the use of a head-mounted display (HMD) because otherwise the user's view would be blocked. Another, similar system tracks the user position but does not have the force feedback hardware. In this system, a barrier tape metaphor is used to depict the outer boundaries of the rate-control threshold. If any part of the user's body crosses the graphical depiction of barrier tape, rate-control is used as in the Virtual Motion Controller (Cirio, Marchal, Regia-Corte, & Lécuyer, 2009).

A similar body-based, position-to-velocity interface (P2V), depicted in Figure 2.1, has been implemented for use in the C6 CAVE at the Virtual Reality Applications Center (VRAC) at Iowa State University. The C6 surrounds the user with six rear-projected surfaces (four walls, ceiling, and floor), providing 3.048 m $\times$ 3.048 m of horizontal movement area. The P2V interface is particularly well-suited for use in a six-sided CAVE because all virtual rotations can be per-

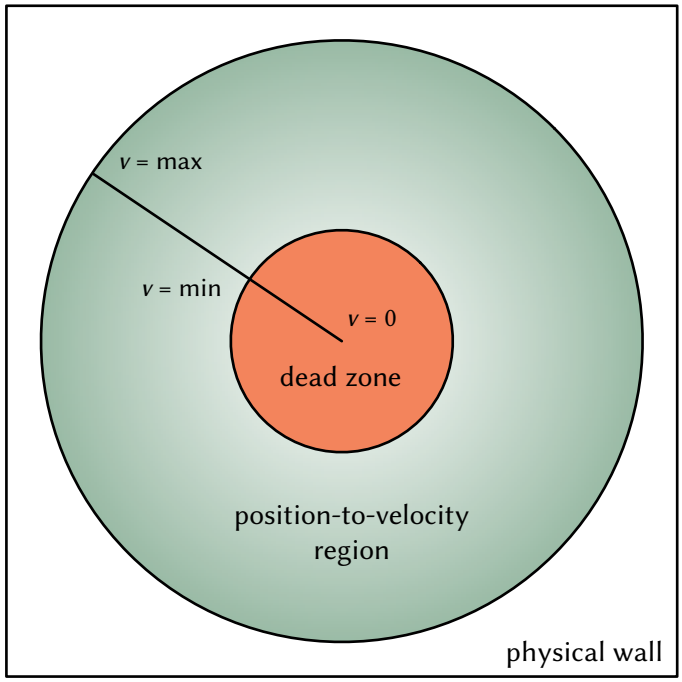

Figure 2.1 Top-down view of the body-based, position-to-velocity (P2V) interface.

formed using the completely natural movement of physically turning one's body, but without the need for restrictive headgear. Solutions such as treadmills are not appropriate for use in the C6 because the floor is a projection surface.

Motor control

Locomotion requires movements of the body so it is useful to look to the motor control literature when investigating unnatural locomotion. Body movements can be broken into multiple phases and understanding these phases may be helpful when attempting to quantify user corrections. Woodworth (1899) formulated the first two-part model describing the phases of movement. His model specifically applied to the speed and accuracy of upper limb movements in goal-directed aiming. The model was comprised of two stages: ballistic and correction. The ballistic phase was thought to contain the initial impulse while the correction phase involved perceptually guided precise control. More recently, Nieuwenhuizen, Martens, Liu, and van Liere (2009) expanded on this work, creating a five phase movement model:

1. latency,
2. initiation,

3. ballistic,
4. correction, and
5. verification.

In this model, additional phases have been added to the beginning and end of the movement. The latency phase encompasses the time before the actual movement starts and the initiation phase contains only small movements. After the actual movement stops is the verification phase. Some activities that require continuous corrections, such as steering, may not have an obvious ballistic phase (Nieuwenhuizen et al., 2009).

Movement intervals can be separated at pauses or times of minimal movement, allowing for phases to be parsed and analyzed. If significant progress toward the target is made during a given interval, it will be considered the ballistic phase (Nieuwenhuizen et al., 2009).

Motor control differences have been found when comparing virtual versus real environments. In particular, in virtual movements, speed tends to be lower in the ballistic phase and the correction phase tends to contain longer pauses, as compared to real world motions. Learning also affects phase characteristics. In particular, performance increases are visible in the ballistic phase more so than in the correction phase. Also the correction phase will tend to contain fewer sub-movements as a motor activity is practiced (Nieuwenhuizen et al., 2009).

Cognitive limitations

Human cognitive resources are limited and must be shared between simultaneous tasks. Different tasks have different processing demands and therefore impose different amounts of cognitive load (Gopher & Donchin, 1986). Current models of working memory vary in specifics, but they tend to distinguish between verbal and non-verbal storage. Baddeley and Hitch (1974) created the most widely accepted multi-component working memory model and it was expanded in 2002 (Baddeley, 2002). The original model separated working memory into two systems, the visuo-spatial sketchpad and the phonological store. In this model, the visuo-spatial sketchpad is used for maintaining visual and spatial information while the phonological store is primarily used for verbal information. According to the model, access to both of these is dependent on attention, a limited resource mediated by a third component, the central executive.

Spatial (but not exclusively visual) information is stored and manipulated in the visuo-spatial sketchpad (Baddeley, 1992). There is some evidence that within the visuo-spatial

sketchpad, there are two subsystems: one for visual appearance and one for location (Darling, Sala, & Logie, 2009). The visual appearance subsystem might be involved with aspects such as color, shape, or pattern while the location subsystem might be used for remembering locations themselves or movement between locations (Logie, 2003). This distinction is not universally agreed upon and thus it has been a popular subject of recent research (Vergauwe, Barrouillet, & Camos, 2009).

There are some tasks that have been used in the past that are known to tax visuo-spatial working memory. One task thought to be primarily visual in nature is the Visual Patterns Test (VPT), which involves remembering patterns depicted in a grid (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999). Tasks that are thought to be more spatial tend to involve remembering the locations or movements of cues through space. For example, past researchers have had participants remember spatial locations such as the movements of a ball (Vergauwe et al., 2009). Based on the Baddeley and Hitch model, this type of "span" task probably relies on the capacity of non-verbal working memory as well as the central executive, which may encode the input and reconstruct it upon recall (Baddeley, 1992). It is likely that visuo-spatial working memory is used when performing all but the most natural aspects of locomotion and specifically, if there is a dissociation between visual and spatial components, it is likely that locomotion primarily requires spatial resources.

The phonological store is used to maintain verbal information. Some standard tasks that are known to use the phonological loop are the verbal *n*-back task or remembering an ordered list of verbal items, such as digits. Similar to the visuo-spatial, this type of phonological span task should rely on the capacity of verbal working memory and the central executive (Baddeley, 1992). In some cases, phonological strategies, such as verbal encoding or counting, may be used to aid in tasks that are not phonological by nature.

Researchers in the Psychology field often use a dual-task selective interference paradigm to determine which working memory resource is used in a particular task. The idea is to load a participant's memory with a task that is known to tax a specific resource (visuo-spatial or phonological) while the user simultaneously completes a primary task. If performance at the primary task decreases while the user performs the VPT, for example, then we can conclude that the primary task also requires visuo-spatial working memory.

It has been shown that additional cognitive resources are required when using unnatural locomotion interfaces (Suma, Finkelstein, Clark, Goolkasian, & Hodges, 2010), but the specific conflicts have not been isolated. Research into skill acquisition has shown that declarative working memory is used when performing a novel task. This is termed "controlled processing." It is relatively slow and mentally demanding. If extensive training occurs and a consistent mapping exists between stimulus and response, the skill is learned. During the

learning process, a skill becomes proceduralized into long-term memory and performance will no longer require working memory (Anderson, 1982). This "automatic processing" is fast and occurs in parallel with minimal effort (Gopher & Donchin, 1986; Shiffrin & Schneider, 1977). If an expert is forced to devote attention to execution of a task, performance will actually suffer (Gray, 2004). Users learning an unnatural locomotion mechanism might follow this same learning progression as they internalize the model mapping control inputs to movement outputs. Body-based locomotion techniques are usually based on actions that have already been proceduralized (e.g., real walking), and thus should require minimal working memory resources. This is a possible benefit to choosing a body-based interface over a gamepad, for example. However, as discussed above, there is no interface that allows for infinite virtual locomotion in a constrained physical space in exactly the same manner as real locomotion. Some aspects of locomotion will be more natural and more proceduralized than others.

Individual differences also play a major role in skill acquisition. In particular, general intelligence and perceptual speed ability are likely to be involved in learning an unnatural interface. General intelligence involves differences of ability that may vary between individuals across many content domains. It is likely to affect initial performance when confronted with an unnatural locomotion interface. Perceptual speed is the speed with which simple cognitive items can be processed and it is likely to be involved as production systems are created (Ackerman, 1988).

Some tests have been devised to test some specific individual differences. One such test is the Perspective Taking and Spatial Orientation Test (PTSOT) (Hegarty & Waller, 2004). The PTSOT is a 12-question paper test with a series of top-down spatial layouts of objects in the world (stop sign, house, car, etc...). For each layout. the participant is asked to imagine standing at a given object, facing another object, and to imagine pointing at a third object. In the answer area, the participant must draw a line showing the direction to the third object. This ability to take an imagined perspective may relate to a user's ability to understand what the resulting sensory change due to locomotion will be. Different ability levels may cause users to employ different strategies, requiring different cognitive resources.

Field of view

Vision is essential for the effective control of locomotion. In some species when visual information is unavailable, locomotion stops (Gibson, 1958). Normal humans have a 200° horizontal and 135° vertical FOV (Wandell, 1995). Many VR systems make use of HMDs to display all visual information for an environment. These systems are popular because of their size,

flexibility, and relative cost. However, HMDs typically suffer from an extremely low FOV. For example, the NVIS nVisor SX HMD has a FOV of $47^\circ \times 38^\circ$ (Willemsen, Colton, Creem-Regehr, & Thompson, 2009). Modern CAVE-like systems are much more expensive and lack portability, but they often boast a vision-limited FOV. The use of stereoscopic shutter glasses limits this to about $90^\circ \times 140^\circ$, but such a system still typically provides a larger FOV than even very expensive HMDs.

It is known that users of VEs do not interpret spatial information such as distances as accurately as in real world scenarios and that this can have an impact on locomotion and navigation performance (Ruddle & Jones, 2001). However, it is uncertain what aspects of VEs lead to these discrepancies (Thompson et al., 2004). There is evidence that peripheral vision is important during locomotion and some studies have shown navigation and memory performance deficits associated with a limited FOV (Alfano & Michel, 1990; McCreary & Williges, 1998). Users of a system with a reduced FOV have been shown to perform worse at walking and search tasks than those with a higher FOV. However there are also findings that indicate that reduced FOV does not impair blind walking performance. One common finding is that users of VR systems tend to underestimate distances but, again, there is no consensus on the reasons for this. Some authors have shown that a limited FOV, combined with other HMD characteristics, does distort perceived distances within an environment (Kline & Witmer, 1996; Willemsen et al., 2009) while others have shown no impact of FOV (Péruch, May, & Wartenberg, 1997). Others have indicated that free head movement is more important than FOV, finding that FOV did not negatively affect distance estimations when full head movement was allowed (Creem-Regehr, Willemsen, Gooch, & Thompson, 2003), but that might be expected because allowing head movement increases the effective FOV. In any case, it is clear that FOV is relevant to the study of locomotion.

When humans interact with the world with unrestricted vision, they normally view the environment with multiple overlapping fixations, or saccades. As a person moves through the world or simply looks around, integration of information from one fixation to the next is generally unnecessary because most of the information is still available in the periphery (Dolezal, 1982). A reduced FOV may require storage and integration of information to be performed cognitively in visuo-spatial working memory.

Using patterns of visual stimulation, known as optic flow, humans can extract information about movements and displacements relative to their environment. Kinaesthesis is the sense of bodily motion and, thus, "visual kinaesthesis" can provide a good source of feedback for locomotion movements. Specifically, forward and backward movement cause expansion and contraction flows, respectively. The rate of optic flow corresponds to the velocity of movement. Its influence on human locomotion is profound, and optic flow rate has been shown

to impact the transition from walking to running and preferred walking speed (Mohler et al., 2007). Steering can be accomplished by moving such that the center of the flow pattern is in the desired travel direction (Gibson, 1958). Humans are capable of locomoting without adequate flow information (Macuga, Loomis, Beall, & Kelly, 2006; Warren, Kay, Zosh, Duchon, & Sahuc, 2001), but they may use alternate strategies to judge distance traveled and orientation. It is possible that such strategies require additional cognitive resources.

Measuring interface effectiveness

If the user's desired destination is known, there are a couple of ways to retrospectively track performance: route completion time and root-mean-squared (RMS) error along the path. However, neither of these methods makes sense for basic locomotion tasks because such tasks are comprised of atomic motions and not an entire path. Furthermore, these performance metrics are bad choices for real-time measurement.

Perhaps the easiest way to detect user locomotion errors is to record collisions with objects and walls in an environment. It can often be assumed that a user running into walls may be having trouble locomoting, regardless of his intended destination. Also it may be useful to refer back to the phases of movement describe above. Movements during the correction phase may reflect a user's problems with getting a locomotion interface to perform as intended. Delays in the latency or initiation phases may point to additional motor planning by the user.

Adaptive systems

When designing systems for heterogeneous users with ever-changing abilities, plans, and needs, it is sometimes useful for the system to learn to change according to a user's current state. Adaptive systems have been implemented to predict user needs in settings such as smart homes (Hagras et al., 2004; Vainio, Valtonen, & Vanhala, 2008) and, of particular relevance here, they have also been used to provide 3D navigation support according to a user's needs (Chittaro & Ranon, 2007) and to calibrate a locomotion system (Engel et al., 2008).

Fuzzy inference systems

In some cases when describing a continuous variable, such as cognitive resource utilization, it is not useful to define explicit bounds on set membership. In fuzzy logic, continuous numerical values are segmented into overlapping "fuzzy" sets. In this way, instead of describing

membership in the Boolean sense where states change abruptly, one can speak of degrees of membership, in that an input variable gradually loses membership in one set while gaining membership in another. A variable is then a member of several appropriate sets to varying degrees. The degree of membership in a given set is defined by a membership function for that set, commonly in the shape of a triangle or trapezoid. Fuzzy logic is complementary to probability. Probability deals with the likelihood of an event while fuzzy logic attempts to describe the degree to which it has happened (Kosko, 1994; Schwartz, 1992). Fuzzy logic is generally useful in situations where variables are continuous, a mathematical model does not exist, a large amount of noise is present, and a group of experts is able to specify rules that the system follows (Cox, 1992).

Implementing a simple fuzzy inference system is generally straightforward. It involves the following basic steps:

1. start with one or more continuous numeric input values;
2. using the set membership functions, determine the membership of the input variable in each particular set (known as "fuzzification");
3. using if-then production rules, map input set membership to appropriate output set membership; and
4. produce a single numerical value according to the output set ("defuzzification").

An inference engine decides which rules to "fire" according to a degree of truth determined by membership functions and current variable settings. In many cases, more than one rule may be selected. This set of rules produces multiple output sets according to the degree of membership of each premise. These output sets must be combined into a single set and commonly a logical OR composition is used for that purpose. For the output to be useful, the inference engine must typically produce a "defuzzified" final result in the form of a number (Schwartz, 1992). A frequently used method for arriving at this defuzzified value is by computing the center of gravity of the combined output set.

There are some freely available, open-source fuzzy inference system libraries. For the system described in Chapter 5, fuzzy-lite (Rada-Vilela, 2011) was selected because it is lightweight and has no dependencies aside from the Standard Template Library included in the C++ Standard Library, making it easy to use and high-performance.

Learning

Learning involves changing a system such that tasks can be done more effectively in the future (Simon, 1983). Through learning, as the system interacts with the environment it can

use past results to increase future success. A domain expert generally specifies initial fuzzy rules and set membership functions. Such a system can then attempt to minimize error over time by modifying rules, rule weights, or membership functions. Learning should be based on multiple error measures, as training data are often incomplete or noisy. Such a system is known as an adaptive fuzzy system (Hayashi, 1992; Kosko, 1994; Lin & Lee, 1992).

VirtuTrace

VirtuTrace is a VR experiment platform developed at the VRAC originally for use in firefighter decision-making studies. Because it is highly configurable, it is an ideal platform for testing and comparing diverse navigation/locomotion systems. VirtuTrace has recently been redesigned with the intent to simplify the introduction of new 3D scenes into a VR system, such as the C6, and run user studies. After creating some configuration files to specify scenes, navigation interfaces, and characteristics of the physics world, the application handles creation of the scene graph, creation of the physics world, switching between scenes, and switching between various types of navigation. If desired, multiple scenes can be sequenced in succession, each with a different locomotion interface and physics model. During an experiment, a participant's movements can be observed in real time, using the display attached to the cluster's head node, and logged to file.

VirtuTrace scenes

There are several existing scene classes, all inheriting from `Scene.h`. A basic scene class, SimpleScene, is provided for users who just want to quickly load a 3D model and navigate through the VE with an interface of their choice. Custom functionality is easy to implement by creating a new scene. Because the scenes are independent of the navigation and physics components, it is easy to mix and match scenes with various navigation interfaces and physical worlds. These choices can be made quickly using XML configuration files.

VirtuTrace is built on OpenSceneGraph (*OpenSceneGraph*, 2011), and thus supports any 3D model format for which an osgDB plugin exists. Typically, the scene models are created in 3D Studio Max 2012 64-bit or Google SketchUp and exported to Collada (`.dae`) files or one of the native OpenSceneGraph formats (`.osg` or `.ive`). Scenes can also be created from scratch in the code. In either case, a physics world can be automatically generated by the application, ensuring that users do not fall through the ground and allowing for tasks such as climbing stairs.

Figure 2.2 Logitech WingMan gamepad.

Navigation in VirtuTrace

The following navigation classes exist, all inheriting from `Navigation.h`. They can be selected and parameters can be adjusted using XML configuration files.

GamepadNavigation. Navigation using a gamepad with two joysticks (one for translation and one for rotation), such as the Logitech WingMan seen in Figure 2.2.

BodyNavigation. The P2V interface described above and depicted in Figure 2.1.

WiiSegwayNavigation. Navigation using a Wii Balance Board and Wii Remote in a manner similar to a Segway. Leaning on the board affects virtual translation, while turning the Wii Remote like handlebars affects virtual rotation.

RealWorldNavigation. Navigation using only real-world movements. There is no additional gain and navigation to distant areas of the VE is impossible.

The basic purpose of each navigation class is to convert user input from an input device or position tracker to a desired velocity in the physics world. The chosen physics class then determines what movement is allowable depending on obstacles and potentially other physical properties. Having a complete physics model for the VE also allows experimenters to observe and log when a participant collides with virtual objects, such as walls.

Chapter 3. Study 1: Working Memory Use During Semi-Natural Locomotion

The literature described in Chapter 2 suggests that virtual locomotion using a semi-natural interface requires cognitive resources. The study described in this chapter is intended to provide insight into the cognitive resource demands of three locomotion interfaces (gamepad, variant of P2V, and real world, in order from least to most natural). Participants were required to perform basic locomotion movements while simultaneously directing cognitive resources to either a spatial or a verbal memory task. The concurrent memory task was designed to simulate the existence of a cognitively demanding primary task in a real-world use case.

Research questions

The following research questions motivated this study.

1. What are the differences in performance under a concurrent working memory load of two locomotion interfaces (P2V and gamepad), and how do they compare to a "real-world" baseline?

2. How are verbal and spatial working memory resources used for the different isolated aspects of semi-natural locomotion?

Previous literature has not addressed the resource pools required for unnatural locomotion, but it is reasonable that spatial and/or general attentional resources would be required, because locomotion is an inherently spatial task. Additionally, it was expected that performance at the more unnatural aspects of the locomotion interface would suffer the most when competing with concurrent tasks for finite resources.

Pilot study

A team of three undergraduate students participating in the Summer Program for Interdisciplinary Research and Education - Emerging Interface Technologies (SPIRE-EIT) Research Experience for Undergraduates (REU) program in Human-Computer Interaction at Iowa State

University helped by conducting a pilot study intended to inform the design of the study described in this chapter. The pilot study compared the P2V locomotion interface to the Wii Segway interface while users, fellow students in the REU program, simultaneously performed spatial and verbal memory tasks.

Several aspects of the experiment and scenario design that came from the team meetings were used in the formal study. First, a task was designed to require users to travel from point to point in the virtual environment so their locomotion performance on translation tasks could be assessed. It was in these meetings that the team decided on having users travel to a virtual golden nugget, as in a video game. For modeling simplicity, this became a spherical "golden nugget." Second, a mechanism was needed for the system to present spatial memory sequences and allow for their recall. The chosen method should not require use of the hands, because participants in different groups (in both the pilot and formal study) would be holding different devices specific to the interface in use. The chosen format for the spatial memory task presentation phase was a random sequence of boxes highlighted in front of the user on a virtual "card." An example card sequence is depicted in Figure 3.1. After performing a series of locomotion tasks, the user was presented with random letters corresponding to the possible box positions, illustrated with an example in Figure 3.2. To recall the sequence, the participant was required to use the letters to recite, in order, the sequence of boxes. For the verbal tasks, a random sequence of number cards, such as that seen in Figure 3.3, was presented. Participants recited the numbers verbally after completing the movements.

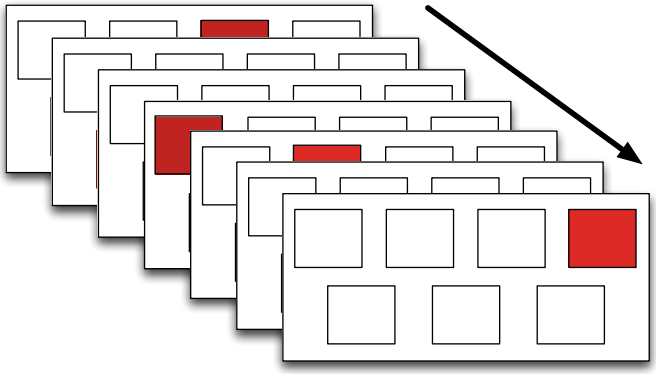

Figure 3.1 Example card sequence from the spatial memory task presentation phase in the pilot study. Boxes were highlighted in a random sequence.

The pilot study aided in identifying aspects that were likely to pose problems in the formal study. One issue was that users had trouble remembering the seven-item span tasks. It seemed that the novelty of being in the C6 interfered with participants' abilities to remember the sequences. This phenomenon was so prevalent that one participant reported that he

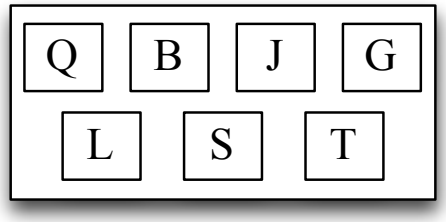

Figure 3.2 Example spatial recall card from the pilot study.

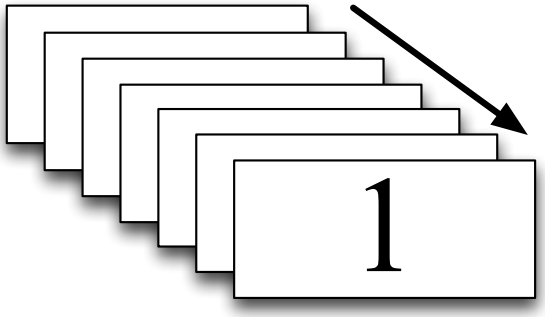

Figure 3.3 Random sequence of number cards from the verbal memory task presentation in the pilot study.

completely forgot to remember the items while doing the movements. Because undergraduate student participants who were also new to virtual reality would participate in the formal study, a solution to this problem was to reduce difficulty by assigning memory tasks with fewer items. Also, because there seemed to be extreme individual differences in participant memory abilities, a pre-test was added in an attempt to customize the difficulty levels. A second interesting finding was that, on several measures, the pilot results suggest that performance was reduced when participants had no memory task as compared to when they had either a spatial or a verbal task. These results suggest that there was increased motivation to complete the tasks in order to get back to remembering the sequence. These initial results will be supported by the findings from the studies described in detail in this dissertation.

The pilot study results also showed that participants were generally very quick at the locomotion tasks. The average time for all tasks to be completed was 22 s but it seemed that the memory sequence could be remembered for a longer period of time. For this reason, the type of task remained the same but the number of tasks in each block was increased for the formal study.

Overall, the pilot study results were in favor of the general flow and the tasks used, so while there were many changes to instructions and aesthetics, the basic experiment design

in the first study closely resembled that used in the pilot study. Before the formal study, all models were recreated to make participants feel more comfortable in the virtual environment and the VirtuTrace code underwent many revisions intended to increase application stability.

Experiment design

The first study incorporated a 3×3 design with three locomotion interfaces (gamepad, P2V, real world) and three levels of memory task (spatial, verbal, none). Locomotion interface was a between-subjects variable because it would have been logistically difficult for the participant to exit the virtual environment, train on a new interface, and then return to the environment. On the other hand, memory tasks were relatively quick and could be performed within the environment, so it made sense for that variable to be manipulated within subjects.

There were three between-subject groups, in order from least to most natural: gamepad (GP), P2V, and real world (RW). The GP group used a Logitech WingMan Cordless gamepad for all locomotion tasks. Participants were instructed to stand in one place, facing the front of the C6 CAVE. The gamepad button configuration was similar to many first-person video games, with one stick used to control the direction of "looking" (like "mouse look" on traditional PC first-person shooters) and the other to control the direction of movement with respect to that direction of looking. The P2V group used a modified version of the P2V interface described above. Because all of the required movements were axis-aligned, components of the velocity vector were computed separately based on the distance from each axis. Calculating the user's velocity in this way reduces the chance of drifting off course on axis-aligned tasks, which could make the movement easier. Movement was completely natural (as in the "real" world) until the user left the dead zone, which was configured, based on pilot testing, to have a radius of 42.67 cm (in CAVE-space) for this study. The outer extent of the P2V region was large enough that all movement outside the dead zone affected the participant's virtual velocity. Participants in the RW group moved about in the virtual world just as they would during locomotion in the physical world. No one interface should be superior overall, because it is assumed that each may be useful for different types of tasks under different concurrent task conditions. This is why having a detailed understanding of task interactions can be useful.

Methods

The study methods were closely modeled after those used in the pilot study described above. Aside from the cosmetic changes described above, the study had the same basic design.

Participants

Fifty-one undergraduate students (32 males) were recruited from the Iowa State University Department of Psychology research participant pool (SONA) and word of mouth. Participants came from multiple departments and majors across campus. All participants were required to have 20/20 (corrected) binocular vision and all played less than or equal to 3.5 hours of video games on average per week. Participants with too much gaming experience were not allowed because they were likely to be familiar with the gamepad. For this reason, users in the GP group were restricted to no more than than 1.5 hours of first-person video games (such as first-person shooters) per week. In this study, the gamepad was intended to be representative of a typical unnatural locomotion interface so it was important to ensure that it was, in fact, unnatural.

Procedures

The study took place at the VRAC. The tasks described below took less than an hour for each participant to complete.

First, participants were asked to complete a pre-questionnaire with topics involving demographic information and video game experience. This document is included in Appendix A. They also completed the PTSOT described in Chapter 2. The PTSOT was administered in order to explore the possibility that users with better spatial abilities may experience less competition for cognitive resources.

Next, participants entered the C6. In the C6, participants were given instructions and a demonstration of how to complete verbal memory tasks in the VE. For verbal tasks, a sequence of letters was presented, similar to that shown in the example in Figure 3.4. After a pause, when it was time for recall of a verbal sequence, the card shown in Figure 3.5 ("recite") was displayed, indicating that it was time to recall the the sequence. After the demonstration, participants were given a series of six verbal memory tasks to assess their individual verbal spans and allow them to practice so they would feel comfortable when doing the real tasks. The difficulty was increased from three items to five items, with two tasks at each difficulty level. Next, participants were trained on the spatial tasks. For these, a sequence of boxes was presented, similar to that shown in Figure 3.6. When it was time for a spatial sequence to be recalled, a card populated with random letters, similar to the example shown in Figure 3.7, was displayed and the participant was required to state the letters that corresponded to the order in which the boxes were presented. Participants were then given a series of six spatial memory tasks, increasing in difficulty from three to five items, to allow practice and assess individual abilities.

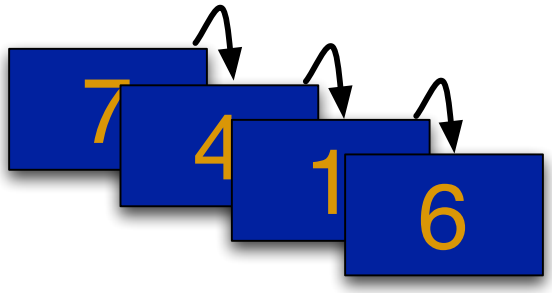

Figure 3.4 Study 1 sample verbal task presentation.

Figure 3.5 Study 1 verbal recall card.

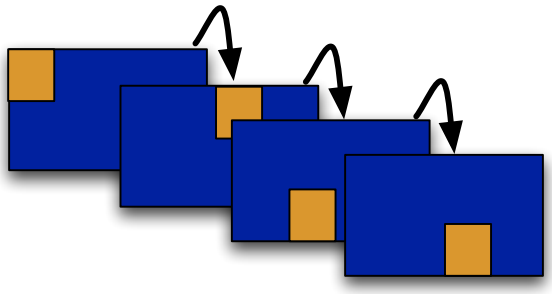

Figure 3.6 Study 1 sample spatial task presentation.

Figure 3.7 Study 1 sample spatial recall card, with random letters.

The practice tasks were designed for two purposes: 1) so the participant would be comfortable with the tasks, and 2) to assess individual abilities in order to customize the difficulty during the experimental phase. If a participant was unable to successfully complete the two tasks at the highest difficulty level (five), the span used during the experimental phase was dropped to four for that particular type (spatial or verbal) of task. This was done to ensure that the span used during the real locomotion tasks was sufficient to tax the cognitive resource in question but not so hard that the participant was incapable of recalling such a large sequence.

Before the experimental phase, each participant was given instructions and a detailed demonstration of the locomotion interface and all locomotion tasks. The VE was viewed using active-stereo shutter glasses. All tasks were performed in a virtual room with a grid texture, similar to the rendering of the room from Study 2 in Figure 3.8. The front wall of the room was purple and the other walls were black. The user was instructed to always face the purple wall and to stand in the center of the CAVE in between tasks. The participant was not allowed to practice the locomotion tasks, but there was a run-through in which the experimenter demonstrated what would be required to successfully complete all experimental tasks. The decision to not allow locomotion practice was made to prevent any learning from taking place before the actual experimental tasks. This would maintain the unnaturalness of the movements and probably the extent to which cognitive resources would be required. It was important for the user to feel comfortable and perform at a high level on the memory tasks and the pre-assessment tests were intended to provide practice.

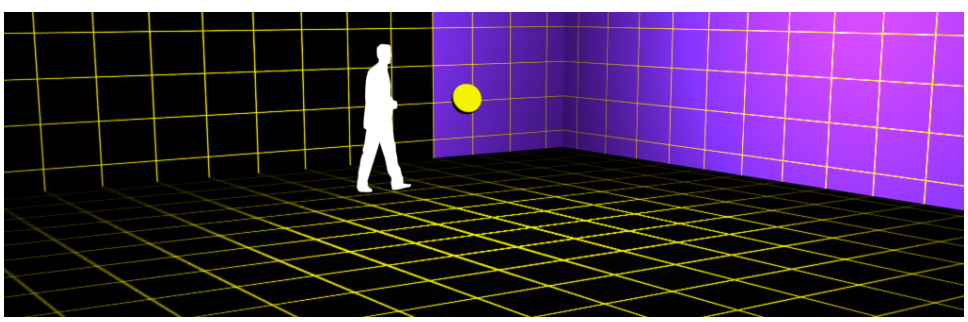

Figure 3.8 Rendered virtual room with grid texture from Study 2, similar to that used in Study 1.

The experimental phase consisted of six experimental blocks. A flow diagram of the experimental blocks is shown in Figure 3.9. Each was structured as a repeating series of locomotion tasks with a memory task presented beforehand and recalled by the participant afterward. Each block had a verbal, a spatial, or no memory task, assigned randomly such

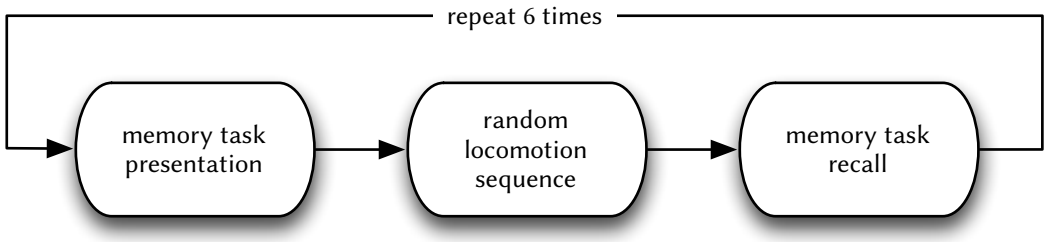

Figure 3.9 Flow of six experimental blocks in Study 1.

that each participant experienced two of each type over the course of all six blocks. The movement phase lasted at least 70.0 s to ensure that participants could not rush through the movements to get to the recall step quicker. The memory tasks were designed exactly as described above and depicted in Figures 3.4, 3.5, 3.6, and 3.7. Each sequence of locomotion tasks was also randomly ordered. The following locomotion tasks were each performed once during each block.

Translate left. The participant retrieved a nugget to the left. An arrow appeared in front of the user, indicating the location of the nugget.

Translate right. The participant retrieved a nugget to the right. An arrow appeared in front of the user, indicating the location of the nugget.

Translate forward. The participant retrieved a nugget to the front.

Rotate left. The environment rotated such that the purple wall was on the left side of the participant, requiring a 90° rotation to the left in order to continue facing the purple wall.

Rotate right. The environment rotated such that the purple wall was on the right side of the participant, requiring a 90° rotation to the right in order to continue facing the purple wall.

Duck. The participant had to duck to avoid being hit by a virtual I-beam flying overhead.

The nugget model used in the translation tasks had a 30.48 cm radius and it was centered 152.4 cm away from the center of the CAVE, 129.54 cm above the ground. Because the C6 has a horizontal movement area of 3.048 m × 3.048 m, this means that participants in the RW group were able to reach the nugget because they only had to come within 30.48 cm of the physical wall. The I-beam model used for the duck task flew 152.4 cm above the ground. Duck

failures were determined according to the height of the head tracking device mounted on top of the stereo shutter glasses worn by the participant. If the virtual I-beam clipped the head tracker, a failure was logged.

Between locomotion tasks, there was a 6.0 s pause, allowing participants to return to the center of the CAVE and await the next movement task. The experimental phase consisted of six such blocks of events described above, with two blocks of each memory task.

After completing all experimental blocks in the C6, participants were asked to complete a post-questionnaire, included in Appendix A, and answer questions in an unstructured interview. These interview questions were intended to uncover any strategies that participants may have used or any particular problems encountered, specifically involving competition for cognitive resources, and they were often tailored to specific problems that the experimenters observed during the locomotion and memory tasks.

Response variables

Recall that there were six movement tasks per block and six blocks per user so each user completed 36 tasks. Relevant metrics were calculated for each of these tasks. As described above, there were three basic types of movement tasks: translate, rotate, and duck. For rotate tasks, the only metric recorded was the elapsed time from task presentation (environment rotated by 90°) until the participant completed the required rotation. For duck tasks, only success or failure was recorded.

Four response variables were calculated for each translation task in Study 1 and Study 2, as shown in Figure 3.10. The elapsed time from when a nugget was presented until movement began is referred to here as *start time*. The elapsed time from when movement was detected until the participant reached the nugget's virtual location is referred to here as *movement time*. The elapsed time from when the nugget location was reached until the participant came to a stop is referred to as *stop time*. This measurement includes time to realize that the task has been completed (nugget reached) and time to manipulate the interface as required to come to a stop (GP: let go of stick; RW: physically stop; P2V: return to dead zone and physically stop). Additionally, one more variable will be used to refer to the total task time for a translation goal to be achieved:

$$\text{translation time} = \text{start time} + \text{movement time}$$

All participant responses on the spatial and verbal memory tasks were recorded and checked for correctness.

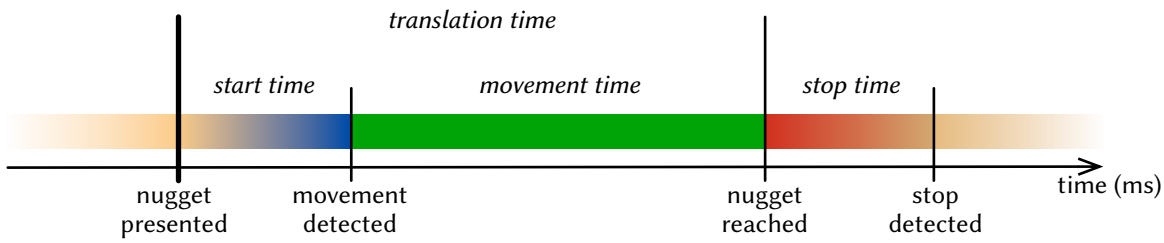

Figure 3.10 Timeline showing Study 1 and Study 2 translation task response variables.

Logging

During a trial, all scene data were logged to comma separated values (.csv) files. The files were then parsed with a Python parser to extract the variables described in the previous section. The following raw data were logged:

- participant head position in every third frame;
- timestamp when participant started movement;
- timestamp when participant stopped movement;
- timestamp when movement task was presented;
- timestamp when movement task was completed; and
- success or failure at end of ducking task.

The experimenters were able to watch a participant's movements on the head node, allowing for a subjective interpretation of the types of problems encountered by users. The participant's head position was logged every third frame using data from an InterSense IS-900 tracker. To determine start and stop time, the application calculated a moving average of the head positions. The length of the moving average window and the threshold for what would be considered movement were adjusted manually before the study began and, while these settings were not perfect for all participants, the same values were used throughout the study and overall the movement detection function was fairly accurate.

Results

The study results enabled analysis of the effects of the different interfaces and memory tasks on movement performance (start, move, stop, turn, and duck) and on memory task performance. In many cases, participants had problems that led to the movements not being atomic.

For example, if a participant sidestepped left but passed the nugget, a right sidestep movement must be performed to complete the task. This means that it would not have been appropriate to treat that movement simply as a left-sidestep action. For this reason, the left, right, and forward task data were combined for analysis. Stop time, duck failures, and memory items missed were most influenced by interface type and/or memory task, and so the following analysis focuses primarily on those measures.

Data cleanup

Some data points did not exist or had to be removed for experimental consistency reasons. Data were removed in the following instances.

- In many trials, the participant was not fully stopped before the next task was presented, so a stop time was not recorded. Likewise, a start time was not recorded if the participant was already moving when a task was presented. For consistency, the experimenters did not attempt to record times manually.
- Application crashes or other hardware and software problems led to incomplete data for some participants.
- In a few cases, participants missed the nugget but thought they had retrieved it. Because the objective was to measure the ability to successfully complete *intended* movements, head position data were manually inspected and discarded where it was clear that the user had passed the nugget and stopped, preparing for the next task, before realizing the mistake.
- Some participants reported using a verbal strategy for the spatial tasks (i.e., coding the locations as numbers). Because the spatial task was intended to tax spatial resources, the affected data were discarded any time a participant reported using such a strategy.

Across all analyses, the percentage of data missing or removed ranged from 7.3% to 24.7% with an average of 14.7%.

Stop time

Recall that the stop time was calculated after each translation task was completed (the nugget was reached). Mean stop time values for all translation tasks and interfaces are shown in Figure 3.11. A two-factor mixed-model analysis was performed with fixed effects for locomotion interface group and memory task combinations (9 means) and random effects for subject, as shown in Table 3.1. The analysis shows significant main effects of interface group

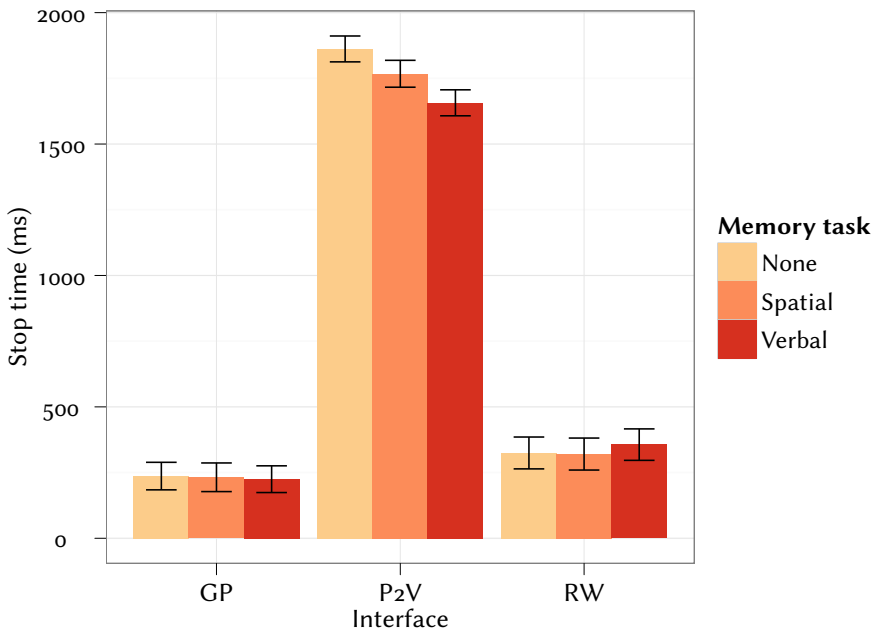

Figure 3.11 Study 1 mean stop time as a function of interface and memory task. Error bars show ±1 standard error of the mean.

[$F(2, 45) = 298.74$, $p < .001$] and memory task [$F(2, 638) = 4.22$, $p = .02$] as well as an interaction between locomotion interface and memory task [$F(4, 638) = 3.36$, $p = .01$]. A significant difference between locomotion interface groups was expected, because stopping with the gamepad (let go of stick) or real-world (stand still) locomotion was trivial, while stopping with the P2V interface required locating and returning to the center of the CAVE. This prediction is supported by the analysis. Also, because stop times were so low in the GP and RW groups, a difference should not exist between memory tasks in those groups. This is also supported by the analysis. A Markov chain Monte Carlo (MCMC) simulation from the posterior distribution for the model was used to obtain estimates and p-values for comparisons of interest. The most interesting results were found in the P2V interface group. Participants using this interface stopped significantly faster when performing a spatial memory task than when performing no task ($p = .04$), and significantly faster when given a verbal memory task as compared to a spatial memory task ($p = .02$).

An explanation for performance being slowest when there was no concurrent task is that participants were motivated to stop faster in order to end the competition between the locomotion task and the cognitive task for working memory resources. This conclusion is supported by the general trends found in the REU pilot study as well as visual inspection of no-task vs. with-task performance on other measures (for example, start time described

Table 3.1 ANOVA table for Study 1 stop time.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	2	298.74^{**}	$<.001$
Error	45		
Within subjects			
Memory task (M)	2	4.22^{*}	.02
M × I	4	3.36^{*}	.01
Error	638	(97613)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$, $^{**}p < .01$.

below) and participant feedback indicating a subjective sense that the tasks competed for resources.

The most intriguing result is the difference in stop time when performing spatial and verbal tasks. There are at least two possibilities that could lead to this difference. First, the participants could have been motivated to stop faster when given a verbal task than a spatial task if, for example, there was a subjective sense of competition for resources when performing a verbal task but not when performing a spatial task. Second, the participants could have been equally motivated during both types of memory tasks but they may have been incapable of stopping as fast during the spatial task, presumably due to competition for spatial resources. The second possibility is supported by a visual inspection of the start time results, which shows the same general trend but did not reach statistical significance, as explained in the following section. This result is intuitive, because it seems that returning to the CAVE center would be an inherently spatial activity. Self-reported feedback also supports the idea that spatial memory tasks interfered with movement performance to a greater degree than verbal tasks. Because stopping is trivial when using a gamepad (one must simply let go of the stick), the stop time data for the GP group do not show this trend.

Start time

A measure that did not reach significance, but seemed to exhibit relevant trends was the start time, with means plotted in Figure 3.12. A two-factor mixed-model analysis was performed with fixed effects for locomotion interface group and memory task combinations (9 means) and random effects for participants. The analysis, seen below in Table 3.2, shows a very significant effect of interface group [$F(2, 45) = 9.37$, $p < .001$], which is potentially interesting (though not surprising or directly relevant to the initial research questions) because it indi-

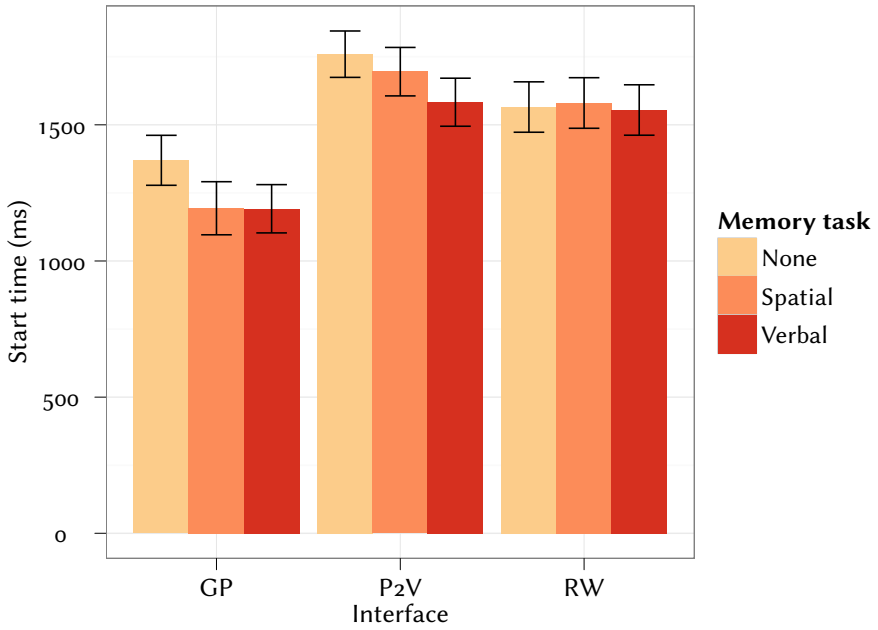

Figure 3.12 Study 1 mean start time as a function of interface and memory task. Error bars show ±1 standard error of the mean.

cates that participants take longer to plan and/or begin full-body movements than to plan and/or begin moving a finger to control a gamepad joystick. The memory task also had a marginally significant effect [$F(2, 728) = 2.34$, $p = .097$]. Looking at the plot, it seems that this was mostly driven by the low performance in the no-task case, which provides more evidence to support the conclusion about motivation described above because the difference is not exhibited in the fully-natural, RW group. The plot also shows a trend in the spatial versus verbal P2V performance that is not significant but is visually similar to that seen previously in the stop-time data.

Though there was no significant interaction of factors in the start time data, the start times in the P2V interface group seem to follow a pattern similar to that seen above in the stop time results. The evidence supporting the second interpretation of the stop data (that participants were incapable of stopping as fast during the spatial task) is in the GP group where we can see that the no-task performance was much slower and there was no real difference between the performances during a concurrent spatial or verbal task. If participants were more motivated to complete the movements quickly during a verbal task, then we would expect them to start faster as well. Therefore, this set of results supports the notion that participants are equally motivated when given spatial and verbal tasks.

Table 3.2 ANOVA table for Study 1 start time.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	2	9.37^{**}	$<.001$
Error	45		
Within subjects			
Memory task (M)	2	2.34	.097
M × I	4	0.67	.61
Error	728	(442034)	

Note. Value enclosed in parentheses represents mean square error. $^{**}p < .01$.

Duck failures

Mean failure rates at the duck task are plotted in Figure 3.13. Recall that the duck task required participants to avoid an overhead virtual I-beam. A failure at this task is defined as being hit by the beam. Because there were zero failures for some combinations of independent variables, two single-factor mixed-model analyses were performed, treating failures as binomial responses. These showed a significant effect of interface on chance of success [$F(2, 45) = 4.87$, $p = .01$] and a marginally significant effect of memory task condition on chance of success [$F(2, 214) = 2.75$, $p = .07$]. These results indicate that participants had trouble ducking when performing a spatial task, even when performing the action as they would in the physical world. Additionally, it seems that participants had particular problems using the gamepad to duck. Recall that participants using the in the GP group were allowed to duck as in real life, or using a button on the gamepad. Observation during the experiment revealed that many participants seemed to accidentally release the button prematurely, before the I-beam had passed.

Memory items missed

When a user simultaneously performs two tasks requiring common cognitive resources, this competition may cause a detriment on performance at either task, or both. For this analysis, the performance on each participant's two spatial tasks was combined into a single average. The number of missed items on the memory tasks is plotted in Figure 3.14. A two-factor mixed-model analysis (6 means) showed a significant main effect of memory task [$F(1, 41) = 23.34$, $p < .001$]. This significant effect could lead us to conclude that the spatial tasks were simply harder than the verbal tasks and so participants missed more items. While that is

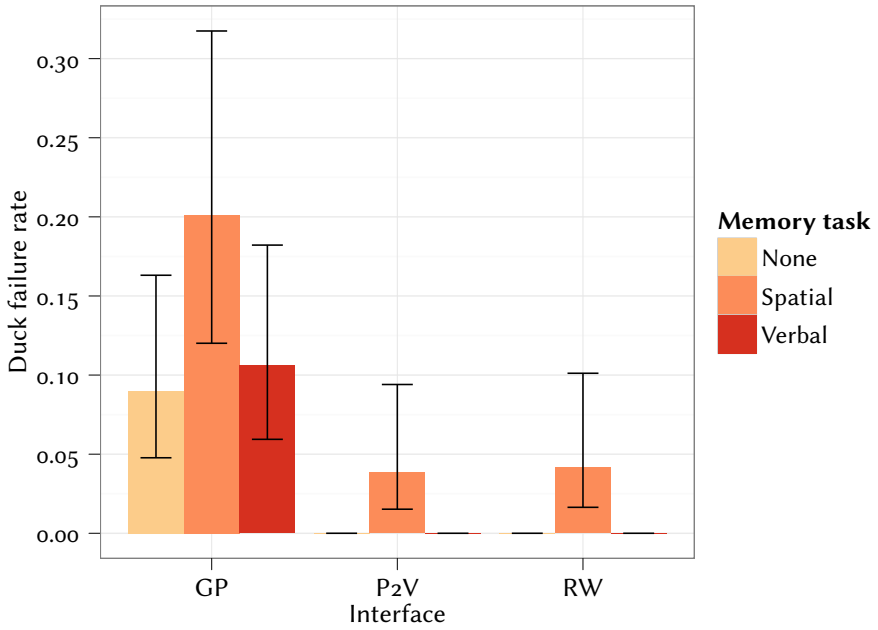

Figure 3.13 Study 1 mean duck failure rate as a function of interface and memory task. Error bars show ± 1 standard error of the mean.

a possibility, recall that participants performed at ceiling on both types of memory tasks with no concurrent task during the practice phase. Perhaps the extra time made the spatial task harder but, based on the expected results, overall patterns in the data, and self-reported participant feedback, it is likely that the difference is largely due to the concurrent locomotion movements. For this reason, analysis proceeded on the spatial results in isolation.

Recall that the gamepad was expected to be the least natural interface, P2V to be slightly more natural, and real world to be a completely natural baseline. To test this hypothesis in terms of the missed memory items, the pattern of performance across interfaces was tested using contrast weights $(1, 0, -1)$ determined by the hypothesis, corresponding to GP, P2V, and RW conditions, respectively. The predicted contrast significantly described the data $[F(1, 41) = 5.394, p = .03]$.

Other interesting findings

All analysis that was directly related to the initial research questions has been described above. Additional analyses were performed on other variables that had a potential to provide insight into virtual locomotion and the problems users have when using unnatural interfaces.

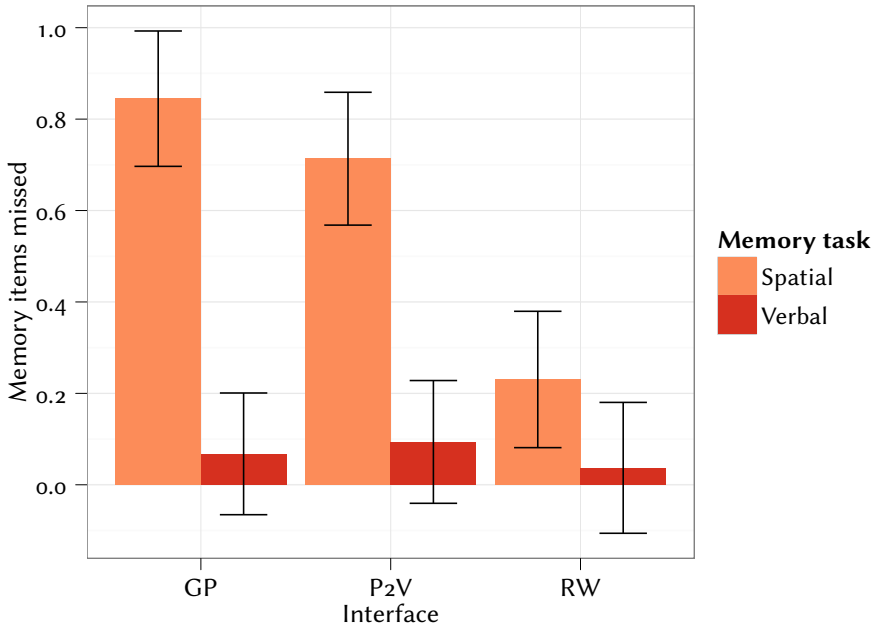

Figure 3.14 Study 1 mean number of memory items missed as a function of interface and memory task. Error bars show ±1 standard error of the mean.

Perspective taking and spatial orientation test (PTSOT)

The PTSOT responses were scored and analyzed. Recall that each participant was given five minutes to complete 12 paper-based tasks. Using a protractor, angles were measured between each participant response and the correct response for the given question. A participant's score was then the average deviation from the optimal responses on the attempted questions. Only 5.33% of the questions were unanswered in this study. The scores were used to divide participants into high- and low-ability groups, as described in Kozhevnikov, Motes, and Hegarty (2007). Participants were divided into high and low ability groups with the bottom quartile (8°–14.25°) in the "high" ability category (11 males, 2 females) and the upper quartile (32.92°–95°) in the "low" ability category (3 males, 10 females), discarding participants in the middle. For the following analyses, PTSOT ability was added as an independent variable to the mixed models used above, resulting in new three-factor mixed models (18 means). Recall that these models treated interface group and memory task as independent variables and also included a term for between-subject error. As in the analyses above, left, right, and forward movements were all treated as repetitions of the same translation task.

Because analysis of the stop time dependent variable produced interesting results, it was the first dependent variable that was examined for an effect of PTSOT ability. In addition to the

Table 3.3 ANOVA table for Study 1 stop time, including PTSOT ability and associated interactions.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	2	209.02**	<.001
PTSOT ability (P)	1	1.87	.19
I × P	2	4.63*	.02
Error	20		
Within subjects			
Memory task (M)	2	1.13	.32
M × I	4	2.26	.06
M × P	2	1.01	.37
M × P × I	4	0.90	.46
Error	331	(110796)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$, $^{**}p < .01$.

effects described above, a three-factor mixed-model analysis, shown in Table 3.3, revealed a marginally significant interaction between interface group and PTSOT ability [$F(2, 20) = 4.63$, $p = .02$]. The means are plotted in Figure 3.15. We can see that participants in the P2V group with a low perspective-taking ability took longer to stop than those with high ability. An MCMC simulation from the posterior distribution for the plotted model was used to obtain an estimate and p-value to confirm significance of this comparison and it was, in fact, significant ($p = .002$). This makes sense, as stopping requires locating and returning to the center of the CAVE, an inherently spatial task. It is not clear why the patterns exist in the GP and RW groups though stopping is trivial in both, resulting in very fast performance.

Another three-factor mixed-model analysis (18 means), shown in Table 3.4, was conducted, revealing a significant effect of PTSOT ability [$F(1, 20) = 5.08$, $p = .04$] on translation time, beginning at task presentation (includes start time). Looking at Figure 3.16, we can see that participants with low perspective-taking ability took longer to complete translation tasks than their counterparts with high ability. Most of this difference seems to be in the P2V group and an MCMC simulation confirms marginal significance ($p = .06$).

Finally, while a three-way mixed-model analysis revealed no significant effects involving PTSOT ability, the plot of rotation time shown in Figure 3.17 seems to indicate that participants in the GP group with a low perspective-taking ability may have taken longer to rotate than those with high abilities. Rotation in the other two groups is expected to be completely natural, so this pattern makes sense and might warrant further future investigation.

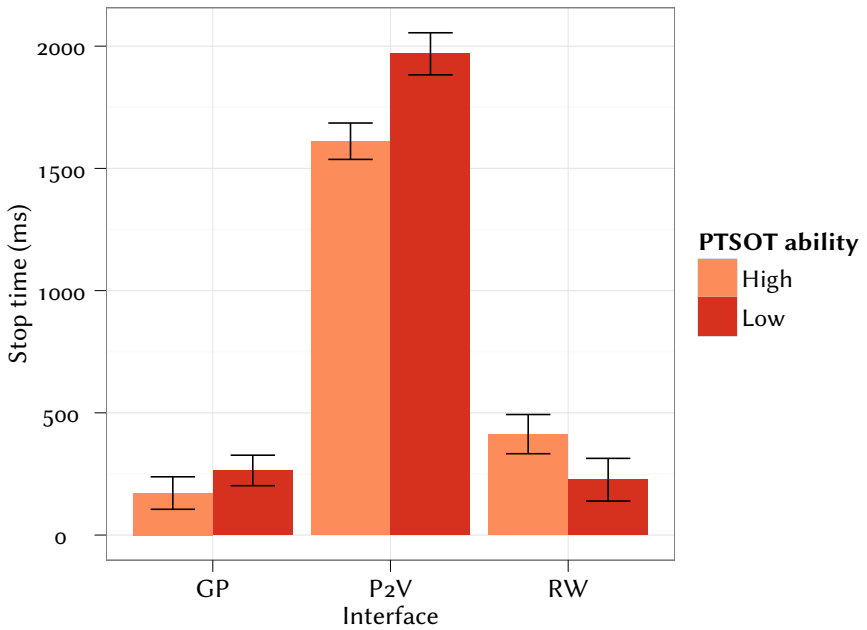

Figure 3.15 Study 1 mean stop time as a function of interface and PTSOT ability. Error bars show ±1 standard error of the mean.

Table 3.4 ANOVA table for Study 1 translation time, including PTSOT ability and associated interactions.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	2	7.65^{**}	.003
PTSOT ability (P)	1	5.08^{*}	.04
I × P	2	0.53	.60
Error	20		
Within subjects			
Memory task (M)	2	0.82	.44
M × I	4	0.71	.59
M × P	2	0.82	.44
M × P × I	4	1.37	.24
Error	364	(1.33×10^{7})	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$, $^{**}p < .01$.

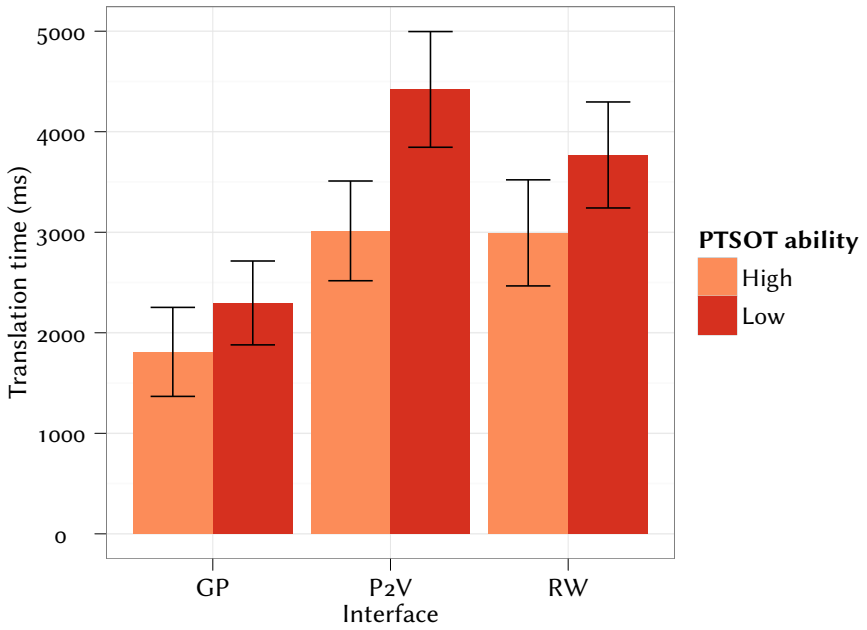

Figure 3.16 Study 1 mean translation time as a function of interface and PTSOT ability. Error bars show ±1 standard error of the mean.

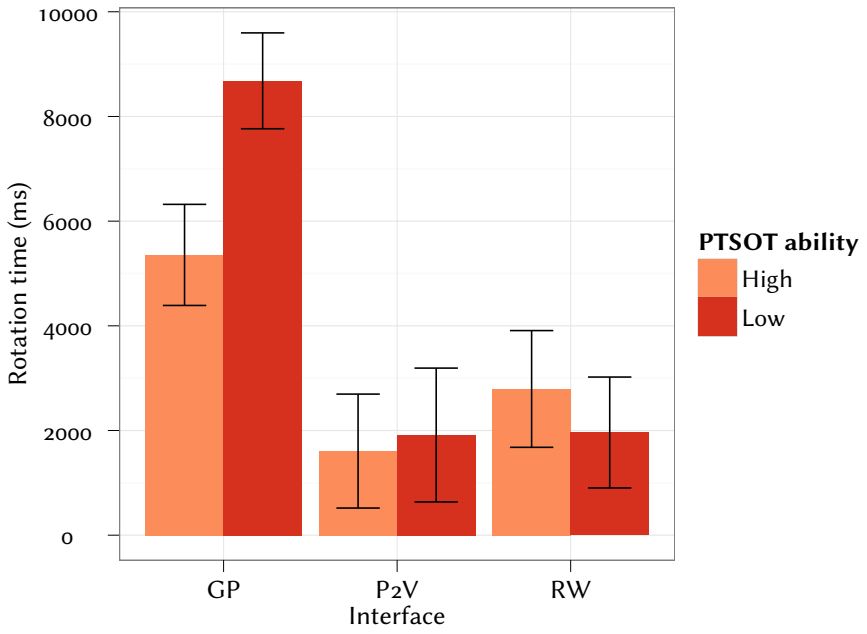

Figure 3.17 Study 1 mean rotation time as a function of interface and PTSOT ability. Error bars show ±1 standard error of the mean.

Table 3.5 ANOVA table for Study 1 translation time, including sex and associated interactions.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	2	7.73^{**}	.001
Sex (S)	1	2.83	.10
I × S	2	0.97	.39
Error	42		
Within subjects			
Memory task (M)	2	1.53	.22
M × I	4	0.75	.56
M × S	2	3.18^{*}	.04
M × S × I	4	0.85	.49
Error	707	(7219951)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$, $^{**}p < .01$.

Sex and locomotion performance

Because the assigned PTSOT categories were so closely aligned with participant sex, there was a concern that PTSOT ability may be a proxy for sex. Thus, more analyses were conducted in which sex was substituted for PTSOT ability in the previous three-factor mixed models. The following models all include interface group, memory task, and sex. The new models have more degrees of freedom than the PTSOT models because the middle two quartiles were removed in the latter. That was not necessary for this sex analysis, so all participants are included.

A three-factor mixed-model analysis of stop time (18 means) revealed no significance of sex or the associated interactions.

A three-factor mixed-model analysis (18 means), shown in Table 3.5, was conducted on translation time (from task presentation until the nugget was reached). The analysis revealed a significant interaction between memory task type and participant sex [$F(2, 707) = 3.18$, $p = .04$]. A plot of the interaction is shown in Figure 3.18. These results do not seem to mirror those found in the PTSOT analysis above, meaning that, with regards to translation time, PTSOT ability does not seem to be a proxy for sex.

Next a three-factor mixed-model analysis (18 means), shown in Table 3.6, was conducted on rotation time. The analysis revealed a marginally significant main effect of sex [$F(1, 44) = 3.47$, $p = .07$] and a significant interaction of sex and group [$F(2, 44) = 4.06$, $p = .02$]. This pattern of results does seem similar to the PTSOT analysis above.

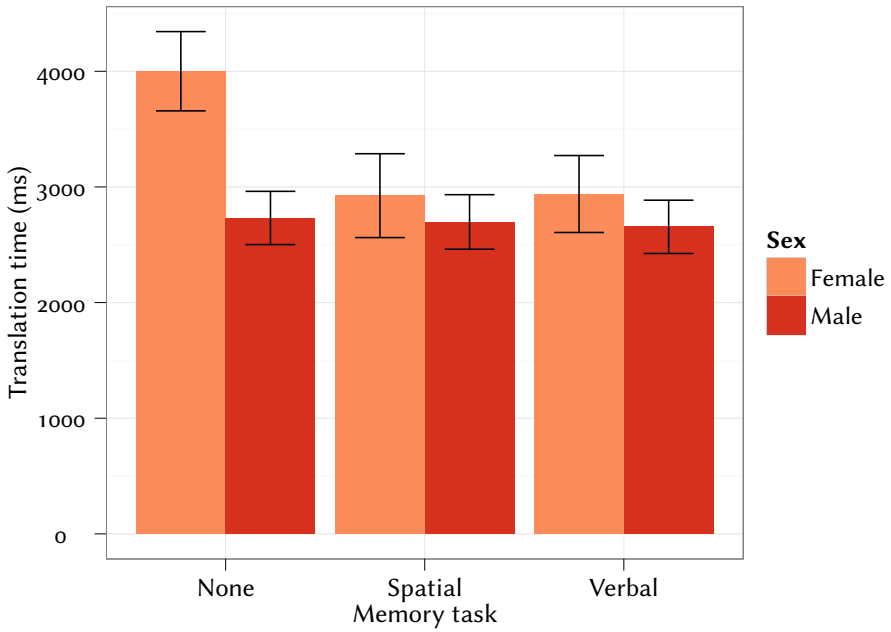

Figure 3.18 Study 1 mean translation time as a function of memory task and sex. Error bars show ±1 standard error of the mean.

Table 3.6 ANOVA table for Study 1 rotation time, including sex and associated interactions.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	2	60.87^{**}	<.001
Sex (S)	1	3.47	.07
I × S	2	4.06^{*}	.02
Error	44		
Within subjects			
Memory task (M)	2	0.30	.74
M × I	4	0.37	.83
M × S	2	0.31	.73
M × S × I	4	0.39	.82
Error	492	(19569679)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$, $^{**}p < .01$.

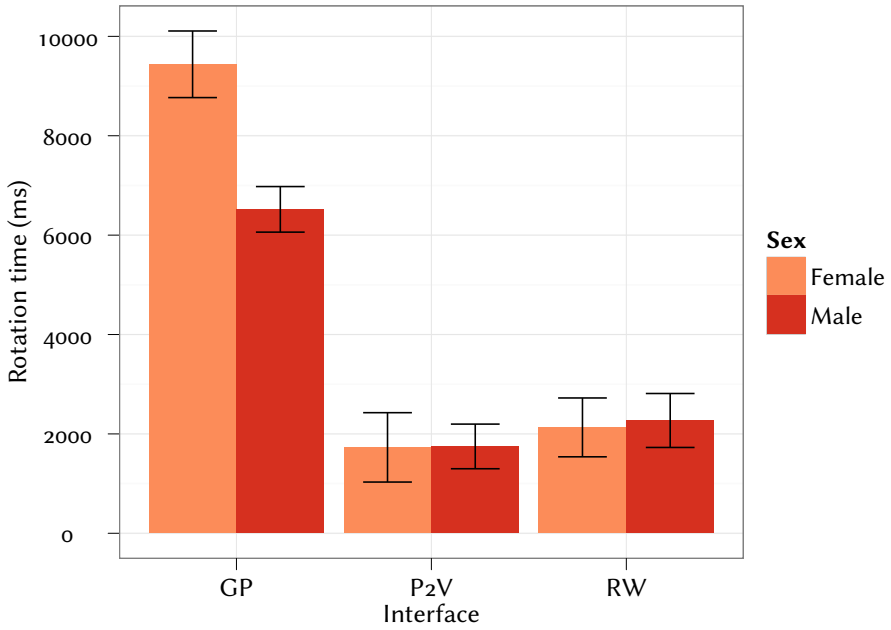

Figure 3.19 Study 1 mean rotation time as a function of interface and sex. Error bars show ±1 standard error of the mean.

Questionnaire responses

Additional analyses were performed on the self-reported post-questionnaire responses. ANOVAs on self-reported performance and immersion revealed no significant effects of interface group or participant sex. An ANOVA, shown in Table 3.7, on self-reported adaptation to the environment indicates a significant effect of sex [$F(1, 45) = 5.88$, $p = .02$]. These means are shown in Figure 3.20. The plot shows greater scores for males than females when using either the P2V or RW interface, but no difference is seen in the GP group. These patterns motivated an additional ANOVA with the two body-based interface groups (RW and P2V) condensed. The resulting graph is shown in Figure 3.21. In this analysis, shown in Table 3.8, the interaction between group and sex is marginally significant [$F(1, 47) = 3.42$, $p = .07$]. Though this does not directly relate to the primary questions posed in the study, it provides potentially valuable insight into body-based interfaces. It seems that females may be less confident in their ability to adapt to body-based systems as compared to a gamepad while males may show the opposite pattern.

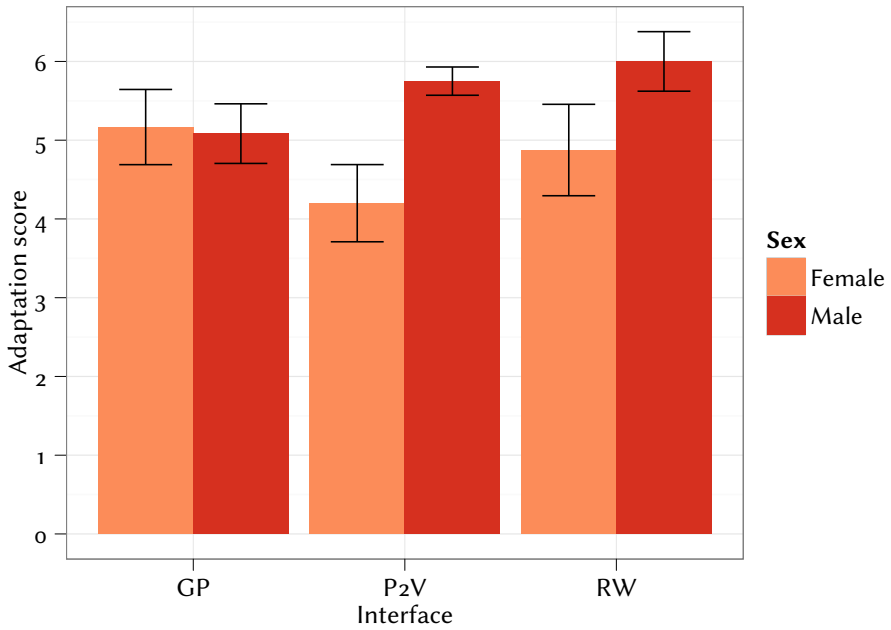

Figure 3.20 Study 1 mean self-reported adaptation scores as a function of interface and sex. Error bars show ±1 standard error of the mean.

Table 3.7 ANOVA table for Study 1 self-reported adaptation.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	2	0.33	.72
Sex (S)	1	5.88*	.02
I × S	2	2.01	.15
Error	45	(1.3706)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$.

Table 3.8 ANOVA table for Study 1 self-reported adaptation with P2V and RW groups combined.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	1	0.55	.46
Sex (S)	1	5.51*	.02
I × S	1	3.42	.07
Error	47	(1.3484)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$.

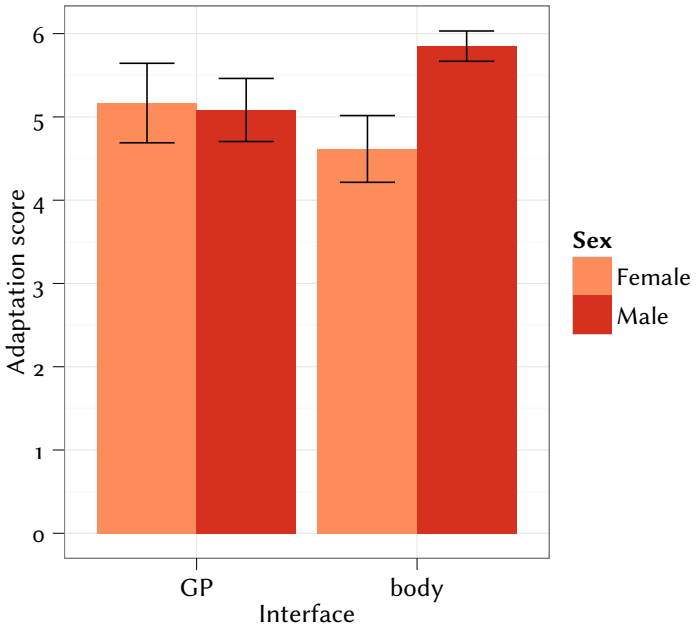

Figure 3.21 Study 1 mean self-reported adaptation scores as a function of interface and sex, with P2V and RW groups combined. Error bars show ±1 standard error of the mean. body: P2V and RW.

Conclusions

The findings from the study described in this chapter can be used to inform the design of future VR systems, particularly with respect to the choice of locomotion interfaces. It is interesting that users in the study tended to let cognitive task performance suffer in order to allocate additional resources to locomotion activities. This alone indicates that this is a worthy area of inquiry because in real-life use cases there would be no contrived memory task and instead performance on a critical primary task (such as battlefield operations) might suffer.

The stop-time results confirmed expectations because stopping with the P2V interface requires returning to the center of the CAVE, an inherently spatial task. This study showed that users of that interface are slower at stopping when given a concurrent spatial task as compared to a verbal task. In Chapter 5 this knowledge was used to motivate the design of an adaptive system that can adjust dead-zone size according to the user's concurrent task load. Specifically, knowledge of the impact of a concurrent spatial task guided the initial definition of rules for the fuzzy system.

Another interesting finding was that participants had problems ducking during a concur-

rent spatial task, regardless of how natural the interface was. It might be interesting to study this phenomenon in the physical world.

The individual differences analysis highlights the importance of an adaptive system's ability to learn about a user's needs, as opposed to being a one-size-fits-all solution. It is unclear to what extent the individual differences findings were due to PTSOT ability or sex. The results show different translation-time patterns for sex and PTSOT ability, indicating different effects are at play in each analysis. However, the rotation results seem to be very similar for PTSOT and sex, indicating that one may be a proxy for the other. In any case, it is clear that individual differences play a role and should be accounted for in future studies.

The sex differences in self-reported adaptation indicate that, while self-reported performance was statistically the same, females are less confident in their ability to adapt to body-based locomotion interfaces. These results can potentially inform the design of systems. Specifically it may help when choosing a locomotion interface to be used predominately by people of a given sex.

This study has also revealed some trends that will help in the design and analysis of future studies and, indeed, even the one described in Chapter 4. First, several participants reported employing a verbal strategy to remember spatial memory sequences. This phenomenon is hard to avoid (Brandimonte, Hitch, & Bishop, 1992), but the memory card design was modified in Study 2 in an attempt to reduce the participant's temptation to try this. Second, because participants seemed to sacrifice performance at the memory tasks in favor of maintaining high performance on the locomotion tasks, the training phase in study two was modified in an attempt to further emphasize the relative importance of the concurrent tasks. Finally, there is evidence that participants may be more motivated to complete movements quickly when performing concurrent memory tasks. There probably is not much that can be adjusted to directly change this in the future but as the tasks become less contrived, and more like what a user would see in the real world, this effect is likely to diminish. Generally, the dual-task selective-interference paradigm worked well for answering the questions posed.

Chapter 4. Study 2: Working Memory Use During Locomotion With a Constrained Field of View

Humans use sensory feedback, particularly visual, to guide movements through the world. Unfortunately, VR systems fail to provide the visual fidelity available during locomotion in the physical world. This likely leads to lower locomotion performance, but it may also cause users to resort to more cognitively demanding strategies when traveling through a VE. These strategies may compete with other ongoing tasks for finite cognitive resources.

A second study was conducted to investigate the connection between the ideas explored in the first study and FOV limitations. It seems reasonable that interfaces with a reduced FOV might decrease a user's locomotion performance and increase working memory load because reduced environmental movement cues, such as optic flow, may cause the user to resort to verbal strategies, such as counting. If so, these verbal strategies would require verbal working memory resources. Additionally, users provided with a limited FOV may be forced to store and manipulate perspective information in working memory rather than in the world, causing increased use of spatial working memory resources during locomotion activities. Finally, it is possible that the alternate strategies employed may require additional general attention resources.

The VirtuTrace codebase underwent many changes between Study 1 and Study 2. Additionally, the experiment scenes were completely rewritten to improve stability and user experience. The models were also recreated to improve aesthetics. Even so, the general study flow and look and feel were very similar to that experienced in Study 1.

Research questions

The following research questions motivated this study.

1. Does limited FOV cause users to resort to verbal locomotion strategies?

2. Does limited FOV during locomotion force users to store and manipulate spatial information in working memory that may have otherwise been available in the world?

There are two basic reasons why virtual locomotion may require working memory resources. The first, the use of semi-natural interfaces, was explored in the previous chapter. The second reason involves the fidelity of sensory feedback provided by VR systems. Previous research, described in Chapter 2, has shown that sensory feedback is used to guide locomotion and that humans, when using interfaces that involve real walking, are able to use automatic processes to update their position in the world. Therefore, when using a semi-natural locomotion interface, the fidelity of sensory feedback should be even more important. In the absence of real walking and adequate sensory feedback, users may be forced to resort to strategies that compete with ongoing tasks for finite cognitive resources.

Experiment design

A 2×3 (FOV, working memory load) design was used for this study. Working memory load was a within-subjects variable while FOV was manipulated between subjects. Participants in both groups viewed the VE through CrystalEyes shutter glasses, which provided stereoscopic vision. Vision for participants in the high-FOV group was restricted only by these glasses ($140^\circ \times 90^\circ$). Participants in the low-FOV group wore the same type of shutter glasses, but with cardboard pieces attached in front of the lenses, limiting FOV to approximately $60^\circ \times 45^\circ$. Field of view was a between-subjects variable because participants in each group wore a different pair of shutter glasses. This decision eliminated the need to enter and exit the CAVE between tasks to switch glasses. As in the study described in Chapter 3, working memory load was a within-subjects variable, because it was easy to present both types of tasks (verbal and spatial) quickly without requiring the participant to exit the CAVE.

All participants used the P2V interface described in Chapter 2. The dead zone was set to a radius of 30.48 cm and the outer extent radius was set to 152.4 cm.

Methods

This study had a very similar design to that described in Chapter 3 for Study 1. In this study, instead of manipulating the locomotion interface, FOV was changed between subjects. The P2V interface was used throughout in this study. According to the lessons learned when conducting the first study and analyzing the results, the following changes were made to the tasks that participants were asked to perform in the environment.

- Because several participants reported using verbal-coding strategies for the spatial working-memory tasks in Study 1, the graphical presentation of the spatial sequence

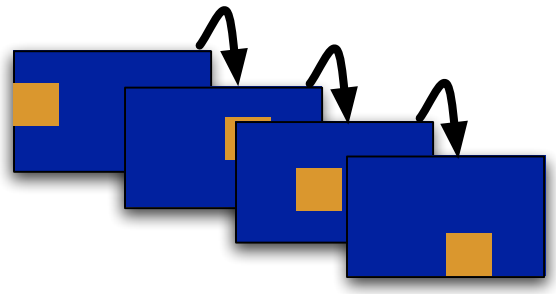

Figure 4.1 Sample spatial task from Figure 3.6, staggered for Study 2.

Figure 4.2 Sample spatial recall card from Figure 3.7, staggered for Study 2.

was changed such that each column was staggered, as seen in Figures 4.1 and 4.2. In this study the boxes were not arranged in a neat grid-like pattern, aiming to reduce the usefulness of a verbal strategy.

- Because participants generally seemed to perform well on the working memory tasks and most were pre-tested to have a span of five items (the maximum allowed) in Study 1, participants in this study were asked to remember between four and six items.
- Because no participants used real-world locomotion in this study, there was no need for task performance to be possible within the 3.048 m × 3.048 m area of the C6. For this reason, the distance to the center of each nugget was increased to 213.36 cm.
- Because participants seemed to have relatively little trouble with the duck task when using the P2V interface in Study 1, the height of the I-beam was lowered to 144.78 cm.

Participants

Thirty-one undergraduate students (20 males) were recruited from the Iowa State University Department of Psychology research participant pool (SONA), through word of mouth, and an announcement in an undergraduate course. In this study, participants were required to have 20/20 (corrected) visual acuity. In contrast to Study 1 recruiting, there was no restriction on allowable video game experience, because the gamepad was not used in this study.

Procedures

The procedures were very similar to those in Study 1. First, participants were asked to complete a pre-questionnaire with topics covering demographic information and video game experience. Also they completed the Perspective Taking and Spatial Orientation Test (PTSOT). Then they entered the C6 and were given instructions and a demonstration of how to complete working memory tasks in the VE. They were then presented with a series of six verbal working memory tasks, similar to the one depicted in Figure 3.4 to assess their individual verbal spans and allow them to practice so they would feel comfortable when doing the real tasks. The difficulty was increased from four items to six items, with two tasks at each difficulty level. Next, participants were trained on the spatial tasks and given a series of six spatial practice tasks, similar to the one depicted in Figure 4.1, again increasing in difficulty from four to six items. If a participant was unable to successfully complete the two tasks at the highest difficulty level (six), the span for the real tasks was dropped to five for that particular type (spatial or verbal) of task. If a participant was unable to successfully remember five items, then the span was dropped to four. This was done to ensure that the span used during the real locomotion tasks was sufficient to tax the cognitive resource in question but not so hard that the participant was incapable of recalling such a large span.

Before the experimental tasks, the participant was given instructions and a detailed demonstration of the locomotion interface and all locomotion tasks. All tasks were performed in a virtual room with a grid texture, rendered in Figure 3.8 and similar but not identical to the one used in Study 1. The room model was recreated for this study in an attempt to improve the quality of the visual feedback. The models were intended to be more aesthetically pleasing but also the room was smaller, intended to give the user more visual feedback to guide the locomotion tasks. The front wall of the room was blue and the other walls were black. The participant was instructed to always face the blue wall and to stand in the center of the CAVE in between tasks. The participant was not allowed to practice the locomotion tasks, but there was a run-through in which the experimenter demonstrated what the participant would be required to do. As in Study 1, the decision to not allow locomotion practice was made to

prevent any learning from taking place before the actual experimental tasks. This would maintain the unnaturalness of the movements and probably the extent to which cognitive resources would be required. The intent was to make the participant feel comfortable and perform at a high level on the working memory tasks and the pre-assessment tests provided practice.

As in Study 1 (Figure 3.9), the experiment was structured as a repeating series of locomotion tasks with working memory tasks interleaved. In each of six blocks, the participant was presented with a working memory span sequence, then a sequence of movement tasks, and then asked to recite the working memory sequence. Each block had a verbal, a spatial, or no working memory task, assigned randomly (two of each) and the movement phase lasted at least 85.0 s to ensure that participants could not rush through the movements to get to the recall step quicker. Each sequence of locomotion tasks was also randomly ordered. In each block, participants completed two of each translation task, one of each rotation task, and a single duck task.

Response variables and logging

The response variables were defined and calculated the same as they were in Study 1. The logging format was also nearly identical to that described for Study 1, with the minor exception that participant head positions in Study 2 were logged every frame instead of every third frame. In Study 2, each participant completed 6 blocks of 9 tasks each, for a total of 54 tasks.

Results

Clearly, if all aspects of Study 1 and Study 2 were identical, one would expect to see similar findings in the high-FOV group, considering that the interface was similar and the FOV was the same. But, because of the improvements described above based on problems encountered in Study 1, it was not appropriate to compare the results directly in this way. However one pattern in particular from Study 1 did provide insight when interpreting results from Study 2. Participants still tended to perform worse when they were given no concurrent working memory task, presumably due to some motivation effect. This similarity to the previous study helped when interpreting the results.

Data cleanup

As described above, numerous data points were collected for each user. In some cases, due to hardware problems, software problems, or participant confusion, affected data points were

discarded. Across all analyses, the percentage of data removed ranged from 3% to 28%. Data were discarded or non-existent for the following reasons.

- In many trials, the participant was not fully stopped before the next task was presented, so a stop time was not recorded. Likewise, a start time was not recorded if the participant was already moving when a task was presented. For consistency, the experimenters did not attempt to record times manually.
- Head tracker malfunctions and graphical anomalies affected a small subset of trials. Data points were removed where it seemed likely that task performance was impeded.
- A bug prevented failures on the duck task from being be properly logged for three participants.
- Several participants reported using a verbal strategy (i.e., assigning a number to each position) to remember the spatial tasks. Because the intent was to load a given cognitive resource, data from affected trials were discarded.
- Some participants missed the nugget and thought they got it on a subset of tasks. This led to the participant standing still, waiting for the next task. Because the intent was to measure intended movements, these affected data points were discarded.
- Some participants got close enough to the virtual walls that nuggets were displayed on the other side so, due to confusion, some data points were discarded.
- One participant reported disobeying directions and playing around. All data were discarded for this individual.

Across all analyses, the percentage of data missing or removed ranged from 2.9% to 28.2% with an average of 13.0%. If either a start time or translation time was missing, no movement time was calculated. For this reason, a higher percentage of movement times were missing from the analysis. Without considering the movement times, the percentage of data missing or removed ranged from 2.9% to 23.5% with an average of 10.6%.

Memory items missed

The primary observation from an initial analysis of the data was that participants seemed to be sacrificing performance on the memory task instead of the locomotion movements. As in Study 1, incorrect answers were scored by counting the minimum number of replacements or swaps required to convert the participant's answer to the correct answer. Because the

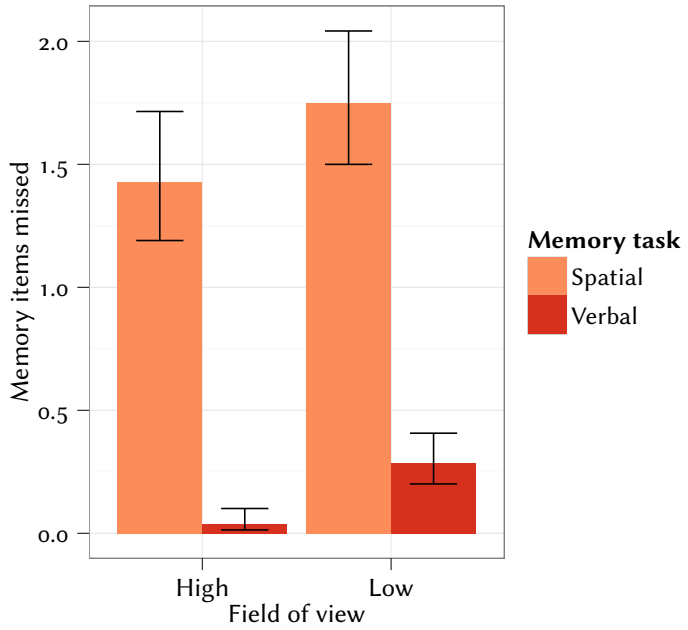

Figure 4.3 Study 2 mean number of memory items missed as a function of field of view and memory task. Error bars show ±1 standard error of the mean.

difficulty in this study was customized to between four and six items depending on individual abilities, allowing for a wider range (as compared to Study 1) of items possible to miss between participants, there was concern that those who had higher abilities would be penalized with the chance to miss more items. However, inspection of the data revealed only one trial in which a participant missed five items and no trials in which six were missed. The following analysis was re-run with that trial omitted but the conclusions were the same.

The pattern of memory items missed (shown in Figure 4.3) in this study is very similar to that in Study 1. Incorrect answers on the verbal tasks were very rare and on the spatial tasks they were much more common. The number of items missed on each memory sequence was treated as a Poisson distribution and a two-factor mixed-model analysis showed significant effects of FOV group [$F(1, 27) = 4.27$, $p = .049$] and memory task [$F(1, 69) = 25.26$, $p < .001$], with a marginally significant interaction [$F(1, 69) = 2.86$, $p = .095$]. It seems that restricting the FOV led to a decrease in both verbal and spatial performance, though note that the verbal performance dropped by a much larger percentage.

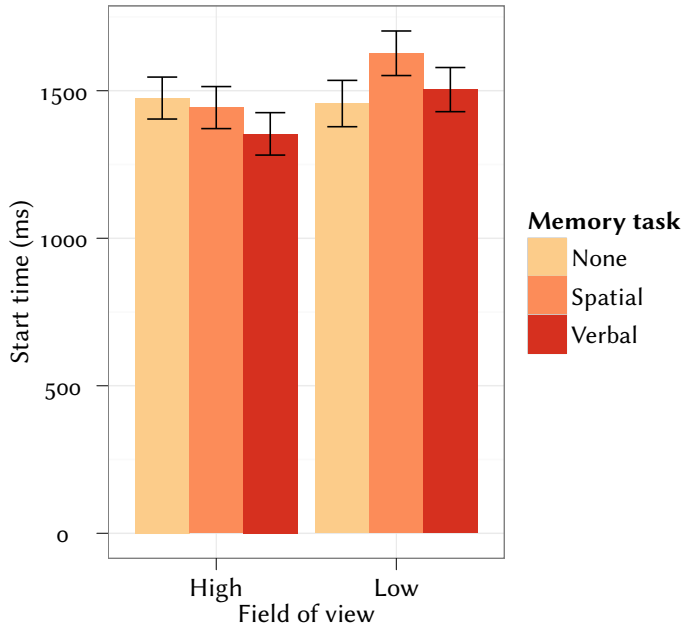

Figure 4.4 Study 2 mean start time for left and right translation tasks as a function of field of view and memory task. Error bars show ±1 standard error of the mean.

Start time

Individual aspects of movements were also analyzed independently. The mean start times for sidestepping left or right are plotted in Figure 4.4. Recall that start times were measured from translation task presentation until participant movement was detected. This means that start time reflects the time required to identify the task to be performed as well as motor planning. A two-factor mixed-model analysis, seen in Figure 4.1, shows memory task [$F(2, 525) = 2.37$, $p = 0.095$] and the interaction between memory task and FOV group to be marginally significant [$F(2, 525) = 2.38$, $p = .09$]. An MCMC simulation from the posterior distribution for the plotted model was used to obtain estimates and p-values for the comparisons between verbal and spatial for each FOV group. This comparison revealed a marginally significant difference between verbal and spatial in the low-FOV group ($p = .08$). Finally, an MCMC simulation was also used to obtain an estimate and p-value for the difference between the average of verbal and spatial with low FOV and that with high FOV. The analysis revealed marginal significance of the comparison ($p = .07$).

Table 4.1 ANOVA table for Study 2 start time on left and right translation tasks.

Source	*df*	*F*	*p*
Between subjects			
Field of view (F)	1	1.31	.26
Error	27		
Within subjects			
Memory task (M)	2	2.37	.095
M × F	2	2.38	.09
Error	525	(219674)	

Note. Value enclosed in parentheses represents mean square error.

Other interesting findings

As in Study 1, a model was created to investigate the effect of FOV group and sex on self-reported adaptation, performance, and immersion scales. An ANOVA showed no significance so those results are not reported here. Individual spatial-ability differences revealed by the PTSOT did have some interesting relationships to locomotion performance.

Perspective taking and spatial orientation test

The PTSOT answers were scored as in Study 1, with an average angle of deviation from the correct answer recorded for each participant. Only 4.89% of the questions were unattempted in this study. As in Study 1 and in Kozhevnikov et al. (2007), participants with scores in the bottom quartile (8.58°–14.5°) were placed in the "high" ability category (7 males, 2 females) and those with scores in the upper quartile (37.82° – 109.11°) were placed in the "low" ability category (5 males, 3 females). All participants in the middle two quartiles were eliminated from the analysis that follows.

A three-factor mixed-model analysis, shown in Table 4.2, was conducted on the start time for left and right translation tasks, adding PTSOT ability as an additional variable to the model used for the left/right start time analysis above. In this new model, memory task [$F(2, 263) = 5.72$, $p < .004$] and the interaction between memory task and PTSOT ability [$F(2, 263) = 4.44$, $p = .01$] are both significant. We can see in Figure 4.5 that the pattern of results is potentially interesting. First, observe that participants with low perspective-taking ability started slower when given a concurrent spatial task than when given a verbal task or no task. This makes sense, as somebody with a lower spatial ability should be expected to perform worse on some types of spatial tasks and it is reasonable that planning and initiating bodily movements

Table 4.2 ANOVA table for Study 2 start time on left and right translation tasks, with PTSOT ability and associated interactions.

Source	*df*	*F*	*p*
Between subjects			
Field of view (F)	1	<0.01	.96
PTSOT ability (P)	1	0.08	.78
F × P	1	0.03	.87
Error	12		
Within subjects			
Memory task (M)	2	5.72^{**}	.004
M × F	2	1.08	.34
M × P	2	4.44^{*}	.01
M × P × F	2	1.60	.20
Error	263	(179983)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$, $^{**}p < .01$.

might require spatial resources. The results for participants with high perspective-taking ability need more interpretation. First notice that these users do not seem to exhibit the same detriment from a concurrent spatial task that was present in the low-ability category. The big difference here is the slower performance when given no concurrent memory task. This may indicate that users with high spatial abilities use more time-consuming strategies when planning and initiating locomotion movements, if resources are not being used for another simultaneous task.

Another three-factor mixed-model analysis, seen in Table 4.3, was conducted on movement time. This analysis revealed a significant interaction between PTSOT ability and memory task [$F(2, 380) = 5.26$, $p = .01$]. The plot is shown in Figure 4.6. More research is needed to draw conclusions, but the pattern of means indicates that different strategies were employed according to PTSOT ability. The plot in Figure 4.7 indicates that the benefits of having a high perspective-taking ability are lessened when the FOV is restricted, though there is no significant interaction between PTSOT ability and FOV.

A final three-factor mixed-model analysis, seen in Table 4.4, was conducted on translation distances, revealing a significant effect of PTSOT ability [$F(1, 12) = 21.79$, $p < .001$]. The plot, shown in Figure 4.8, indicates that participants with low perspective-taking scores traveled greater average distances than those with high scores, regardless of FOV group. These results support the premise that users with a low spatial ability perform worse on semi-natural locomotion tasks.

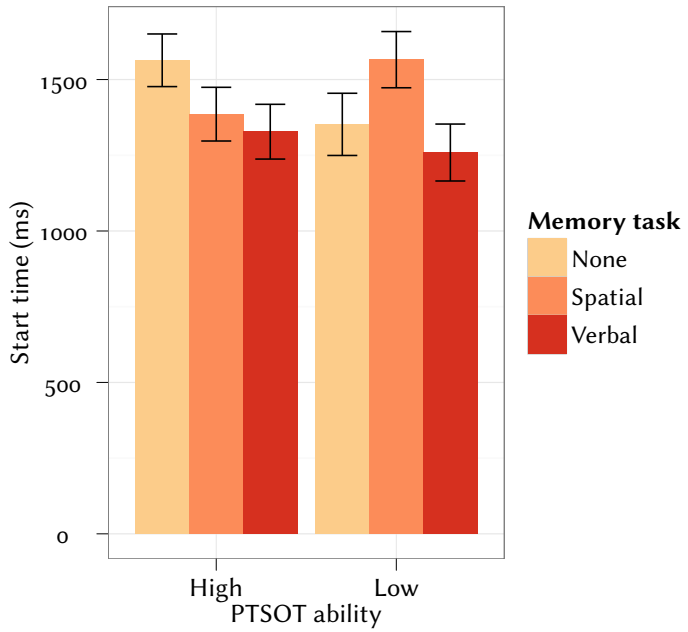

Figure 4.5 Study 2 mean start time for left and right translation tasks as a function of PTSOT ability and memory task. Error bars show ±1 standard error of the mean.

Table 4.3 ANOVA table for Study 2 movement time, with PTSOT ability and associated interactions.

Source	*df*	*F*	*p*
Between subjects			
Field of view (F)	1	2.17	.17
PTSOT ability (P)	1	2.02	.18
F × P	1	0.17	.69
Error	12		
Within subjects			
Memory task (M)	2	1.40	.25
M × F	2	0.15	.86
M × P	2	5.26*	.01
M × P × F	2	0.17	.84
Error	380	(2093304)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$.

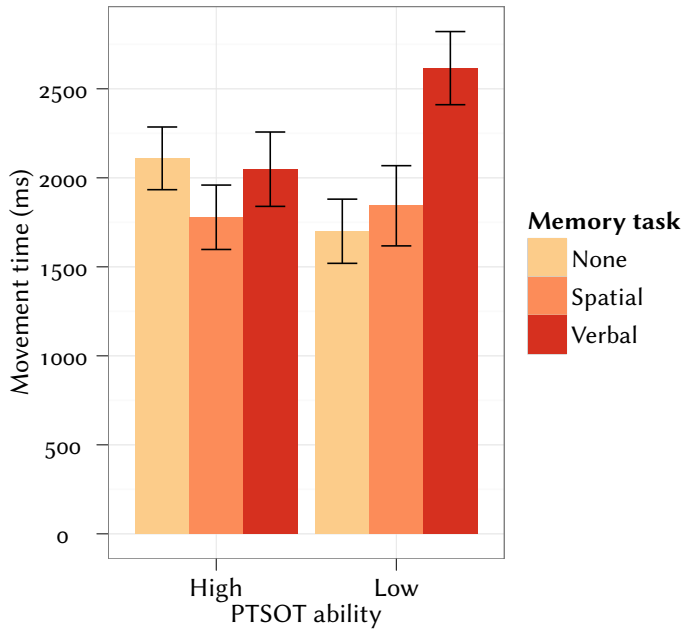

Figure 4.6 Study 2 mean movement time as a function of PTSOT ability and memory task. Error bars show ±1 standard error of the mean.

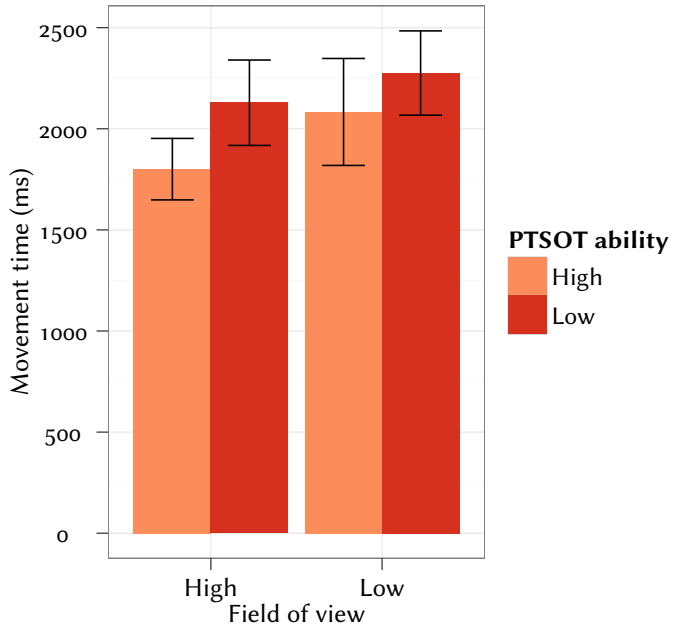

Figure 4.7 Study 2 mean movement time as a function of field of view and PTSOT ability. Error bars show ±1 standard error of the mean.

Table 4.4 ANOVA table for Study 2 distance, with PTSOT ability and associated interactions.

Source	*df*	*F*	*p*
Between subjects			
Field of view (F)	1	3.02	.11
PTSOT ability (P)	1	21.79**	<.001
F × P	1	1.60	.23
Error	12		
Within subjects			
Memory task (M)	2	1.67	.19
M × F	2	1.33	.27
M × P	2	0.80	.45
M × P × F	2	0.52	.59
Error	504	(44.76)	

Note. Value enclosed in parentheses represents mean square error. $^{**}p < .01$.

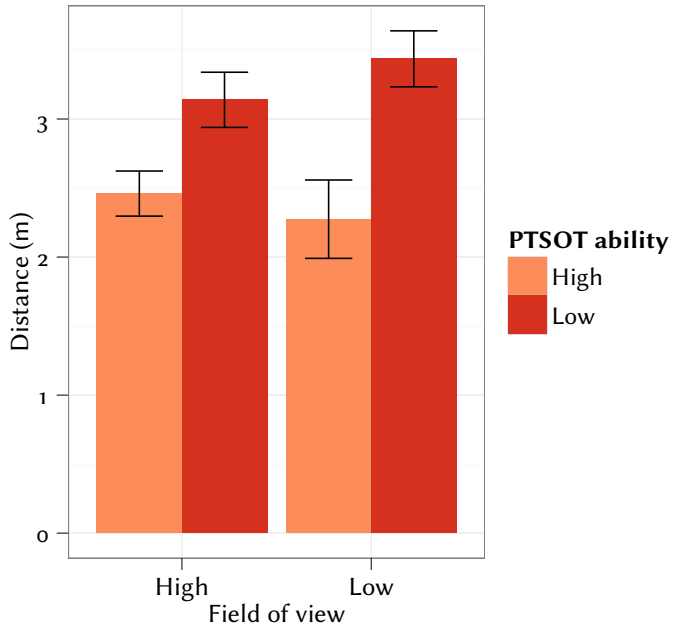

Figure 4.8 Study 2 mean distance as a function of field of view and PTSOT ability. Error bars show ±1 standard error of the mean.

Sex and locomotion performance

Recall that in Study 1, there was reason to believe that PTSOT score may have been a proxy for sex. Based on the number of males and females in each PTSOT ability in this study, there was little reason for concern. However, for completeness, the analyses from the previous section were re-run with sex replacing PTSOT ability in the models. No significant main effects or interactions were found. Recall that the connection between PTSOT ability and sex in the first study was uncertain. These results indicate that spatial ability in itself is more relevant when considering cognitive resource usage during virtual locomotion, as opposed to sex.

Conclusions

The study described in this chapter showed that virtual locomotion with a constrained FOV causes a nearly equivalent detriment to performance on a concurrent spatial or verbal task, beyond the problems due to semi-natural locomotion already seen in Chapter 3. This symmetric performance decrease across memory tasks indicates that it is likely that general attentional resources are in use, as opposed to either the spatial or verbal resource pool. An alternate theory is that locomotion with a small FOV requires an equal amount of both verbal and spatial resources. In either case, the reason for this additional resource usage when FOV is reduced may be due to alternate strategies employed by users in the absence of high-fidelity sensory feedback. Even though the resource usage may be the same, for real-world applicability it may be useful to note that the verbal detriment increases by a much larger percentage than spatial when going to a reduced FOV. As in Study 1, this set of results highlights the importance of gaining a greater understanding of and mitigating problems stemming from resource competition during virtual locomotion with a reduced FOV. In the scenario described here, performance dropped on a contrived representative task. In a real-world use case, performance might be sacrificed on a cognitively-demanding task that is critical to success in the application domain.

Additionally, there may be an impact of a concurrent working memory task on start time when using a locomotion interface with a limited FOV. These results were not quite significant, but the data patterns follow what might be expected, particularly in light of the memory task results. It makes sense that planning or initiating movements may require spatial working memory resources, as these locomotion tasks are inherently spatial in nature. However, previous research did not show an additional demand under low-FOV conditions. This difference could indicate a change in planning strategies when the FOV is reduced. The relationship between PTSOT and start time also points to a possible difference in strategies between users

with high perspective-taking ability and those with low ability. It seems that users with high spatial abilities may use more extensive movement planning strategies, unless a concurrent task requires general attention resources. Users with low spatial ability, on the other hand, seem to employ different planning strategies that are interrupted by a concurrent spatial task. The relation of cognitive resource usage to planning and initiation of movements should be investigated further because it could have implications for the design of VR systems and applications. In addition, future work should investigate if similar effects exist for other aspects of sensory fidelity, such as resolution.

These results may be used in system design, as described above, but they can also inform the design of an adaptive system that attempts to mitigate these problems. Such a system is described in Chapter 5.

Chapter 5. Design of Fuzzy Navigation Engine

The previous chapters have described studies and findings involving the impact of two types of simultaneous cognitive tasks (verbal and spatial) on the different aspects of locomotion performance. These studies showed that the cognitive task performance can also be affected by the competition for resources during semi-natural virtual locomotion. The second objective of this research is to use that knowledge to mitigate the dual-task detriment. To some extent, in domains with a well-defined set of tasks, it may be possible to choose a locomotion interface according to the expected cognitive demands of those tasks. Because no interface is perfect, and choosing will always involve trade-offs, an adaptive interface should be beneficial. The final portion of this work involves the creation and basic testing of an interface that is able to make use of information about a user's current working memory load to modify parameters of the locomotion interface. The adaptive system described in this chapter, referred to as "Fuzzy P2V," is based on the P2V interface from previous chapters, but with the addition of the "Fuzzy Navigation Engine," which incorporates a fuzzy inference system. Figure 5.1 shows a general conceptual model of how the system works to adjust the dead-zone radius for the new interface, not yet including details on how the system learns.

Relevant study findings

Recall that Studies 1 and 2 were designed to identify specific movement problems that result from dual-task competition for resources. Because participants seemed to sacrifice performance on the cognitive tasks to maintain a high level of performance on the locomotion tasks, little information was attained about specific movement problems. However, there were findings that were informative when designing the adaptive system.

First, results show that the more unnatural the locomotion interface a user is required to use, the lower performance will be on a simultaneous spatial task. This means that when a user is performing a spatial task, it should be beneficial if locomotion becomes more natural. This clearly amounts to a trade-off, as the system cannot always allow completely natural movement, so it will be important to lower the naturalness once the concurrent spatial task has been completed.

Second, results also show that users have more trouble stopping with the P2V interface when they are performing a concurrent spatial task than when they are performing a verbal task. Although the results above also showed that users with no task stopped slower than those who had a spatial or verbal task, it is likely that this pattern will not hold in real-world use cases. Users in real world scenarios are likely to sacrifice less on the cognitive task than they did in these studies, as the other task will not be contrived but will be of real-world importance to a user's objectives in the environment. Thus, if the type of load matters, as the spatial-verbal difference indicates, a user with a necessary resource truly loaded should not exhibit higher performance, regardless of factors such as motivation.

Third, the results of Study 2 indicate that performing locomotion tasks with a restricted field of view requires additional spatial and verbal resources. This provides us with evidence that some aspects of locomotion, particularly when sensory fidelity is reduced, may be hindered by a concurrent spatial or verbal load.

Finally, while a one-size-fits-all approach to adaptation may be better than a non-adaptive baseline interface in terms of the results discussed above, the findings related to individual differences make it apparent that such an approach will be sub-optimal. Also, both studies above examined first-time users. As users learn to use an interface more effectively it will become more natural, meaning that less adjustment should be needed as users improve. An effective system should adapt its parameters through learning.

Approach

Informed by the study results listed above, the Fuzzy Navigation Engine adjusts the size of the dead zone according to the user's current cognitive task load. As the dead-zone radius increases, the outer extent radius increases at the same rate, capping maximum possible speed only if it is larger than the CAVE. Because the inside of the dead zone represents an essentially natural interface, this increases the extent to which movements are natural. Additionally, it is plausible that a larger dead zone will be easier to find, which will facilitate stopping. Because both spatial and verbal resources have been shown to be used for locomotion, the Fuzzy Navigation Engine considers both types of load when determining an appropriate dead-zone radius. However, because spatial load plays a greater role, it has been given more influence on the dead-zone radius than verbal load.

The basic input-output flow of the Fuzzy Navigation Engine is shown in Figure 5.1. For a given memory-load-changed event (i), the system takes two inputs, spatial load (S_i) and verbal load (V_i), and produces a dead-zone radius (D_i).

The Fuzzy Navigation Engine learns at discrete times when it receives a learn event. It

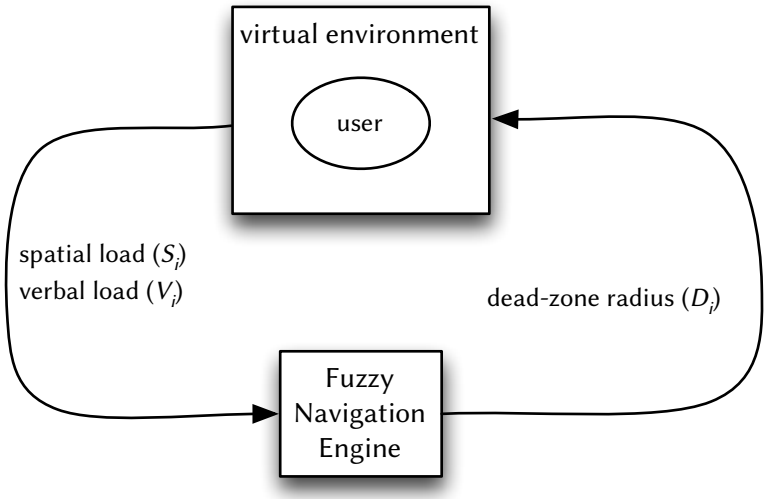

Figure 5.1 Basic input-output flow of the Fuzzy Navigation Engine.

then checks to see if certain key aspects of locomotion performance are outside of the desired ranges. To facilitate learning, a set of error values are introduced. These error values are combined and used to adjust the membership functions for the fuzzy input variables.

Fuzzy inputs

Hardware specifications rarely change when an application is running, so attributes such as FOV can not usually provide a meaningful input to an adaptive system. The abilities of the user with respect to the locomotion technique will change as the user learns how to move about more effectively, but it is not clear how to detect and quantify skill acquisition. However, the system does have access to some level of information about concurrent tasks that the user is being asked to complete. This knowledge is used as input to the Fuzzy Navigation Engine. Because it is difficult to effectively measure task load, for the purposes of this research concurrent tasks were simple but well validated in the cognitive psychology domain. In fact, they were nearly identical to the working memory tasks that were used in Studies 1 and 2. For this research, the following input variables were used to drive the fuzzy inference system:

- number of spatial items currently being remembered (S_i) and
- number of verbal items currently being remembered (V_i).

Using the tasks from the previous studies, it is straightforward to obtain a rough estimate of load using the number of items in a given domain (spatial or verbal) that a user is currently required to maintain in working memory.

Before a fuzzy logic solution can be used, appropriate sets must be defined to map numerical values to fuzzy linguistic terms. To define set membership functions (μ), the Fuzzy Navigation Engine implements overlapping trapezoid functions, with a general template defined as follows, where, for example, x represents a specific verbal or spatial load (S_i or V_i), m_{L} represents the slope of the left-hand side of the trapezoid, and b_{L} represents the y-intercept of the left-hand side of the trapezoid:

$$\mu_{\text{trap,L}} = \begin{cases} m_{\mathrm{L}} \times x + b_{\mathrm{L}} & \text{if } x_1 \leq x \leq x_2 \\ 1 & \text{if } x_2 < x \leq x_3 \end{cases}$$

$$\mu_{\text{trap,R}} = \begin{cases} 1 & \text{if } x_1 \leq x \leq x_2 \\ m_{\mathrm{R}} \times x + b_{\mathrm{R}} & \text{if } x_2 < x \leq x_3 \end{cases}$$

$$\mu_{\text{trap}} := \mu_{\text{trap,L}} \cup \mu_{\text{trap,R}}$$

The Fuzzy Navigation Engine uses these overlapping trapezoid functions to define each linguistic level of memory load (low, medium, and high); for both spatial and verbal memory loads: $\mu_{S,\text{LOW}}$, $\mu_{S,\text{MEDIUM}}$, $\mu_{S,\text{HIGH}}$, $\mu_{V,\text{LOW}}$, $\mu_{V,\text{MEDIUM}}$, and $\mu_{V,\text{HIGH}}$.

Thus membership functions, $\mu(S)$ and $\mu(V)$, map numeric values to linguistic (fuzzy) terms. In the Fuzzy Navigation Engine, fuzzy sets are configurable using an XML file. These input-variable membership functions change as the system learns.

$$\mu_S := \{\ \mu_{S,\text{LOW}}, \mu_{S,\text{MEDIUM}}, \mu_{S,\text{HIGH}}\ \}$$
$$\mu_V := \{\ \mu_{V,\text{LOW}}, \mu_{V,\text{MEDIUM}}, \mu_{V,\text{HIGH}}\ \}$$
$$\widetilde{S} := \{\ S_i, \mu_S \mid S_i \in S\ \}$$
$$\widetilde{V} := \{\ V_i, \mu_V \mid V_i \in V\ \}$$

Fuzzy rules

Fuzzy logic is a good choice for systems that are difficult to model in precise mathematical terms but which can be described linguistically by experts. In the Fuzzy Navigation Engine, the fuzzy rules are configurable with an XML file. For the user study that follows, they are defined according to the findings from the first two studies. Locomotion performance is thought to continually decrease as spatial load increases but performance often is not affected much by verbal load. It is likely that, in some cases, general attention resources of some sort are required and in this case verbal load may have an effect on locomotion performance. The rules have been defined accordingly, with spatial load being given a greater influence

on dead-zone size. The production rules (R_i) used in the studies described in this chapter are based on the results of Studies 1 and 2. They are defined as:

$$R_1: \text{IF } S_i \in \mu_{S,\text{LOW}} \text{ THEN } D_i \in \mu_{D,\text{SMALL}}$$
$$R_2: \text{IF } S_i \in \mu_{S,\text{MEDIUM}} \text{ THEN } D_i \in \mu_{D,\text{MEDIUM}}$$
$$R_3: \text{IF } S_i \in \mu_{S,\text{HIGH}} \text{ THEN } D_i \in \mu_{D,\text{HIGH}}$$

$$R_1: \text{IF } V_i \in \mu_{V,\text{LOW}} \text{ THEN } D_i \in \mu_{D,\text{SMALL}}$$
$$R_2: \text{IF } V_i \in \mu_{V,\text{MEDIUM}} \text{ THEN } D_i \in \mu_{D,\text{SMALL}}$$
$$R_3: \text{IF } V_i \in \mu_{V,\text{HIGH}} \text{ THEN } D_i \in \mu_{D,\text{HIGH}}$$

$$R_S := \{ R_1, R_2, R_3 \}$$
$$R_V := \{ R_4, R_5, R_6 \}$$

Fuzzy outputs

The Fuzzy Navigation Engine adjusts the radius of the dead zone based on the current values of the input variables S_i and V_i. Figure 5.2 shows a top-down depiction of a CAVE, with circles illustrating how changing the dead-zone radius affects velocity. Increasing the size of the dead zone while increasing the outer extent of the P2V region by the same amount accomplishes the following three things.

- It provides a greater area in which movement is completely natural. In effect, a greater percentage of movements will be natural when accomplishing navigation tasks. Because the C6 is exactly ten feet in width, it is useful to describe the dead-zone size in feet. In this way, it is easy to see that a dead zone with a radius of one foot provides natural movement for 1/5 of the distance from the center of the CAVE to the wall. For this reason, all dead-zone sizes in this chapter are listed in feet, as opposed to meters. According to the conclusions in Studies 1 and 2, an interface with a greater proportion of natural movement will require a smaller quantity of cognitive resources, leaving them available for concurrent tasks.

- A larger dead zone provides a larger area to return to when stopping. This should make stopping easier and finding the dead zone should require a smaller quantity of spatial cognitive resources, leaving them for concurrent tasks. A disadvantage is that a user may not truly be in the center when stopped, potentially hindering the start of next movement.

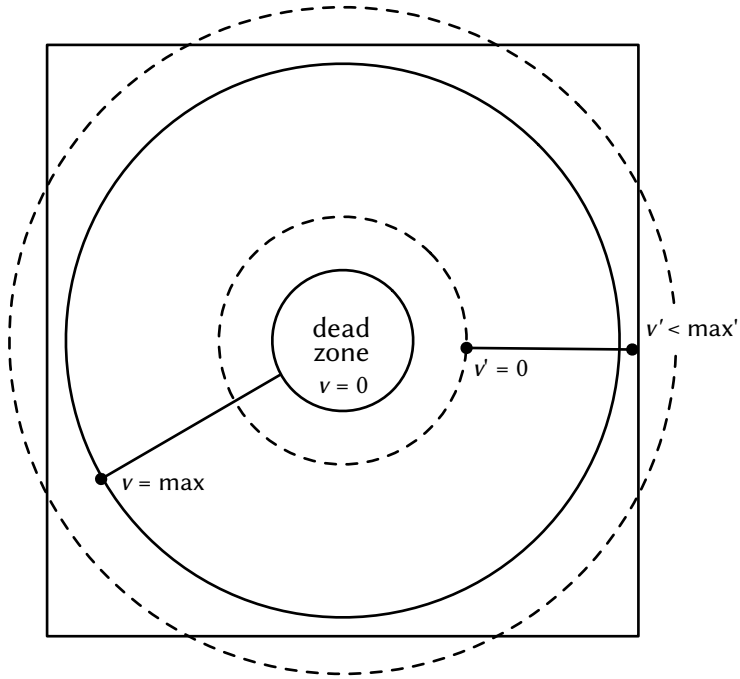

Figure 5.2 Top-down depiction of a CAVE, with circles illustrating how changing the dead-zone radius affects velocity. *v*: velocity; max: maximum possible velocity.

- It limits the maximum velocity at the outer edge of the P2V region. This will act as a sort of "training wheels" when the user cannot expend the necessary cognitive resources but will allow for higher performance when the user is capable. Note that this could increase the risk of running into physical walls if the outer bounds provide a speed that is much slower than the user desires.

A reasonable starting point, and the one chosen for Fuzzy P2V, is to use a symmetric dead zone. The following fuzzy output terms have been used in the fuzzy inference system (low, medium, and high) defined by the following triangle membership functions: $\mu_{D,\text{SMALL}}$, $\mu_{D,\text{MEDIUM}}$, and $\mu_{D,\text{LARGE}}$.

These triangle functions are a special case of the more general trapezoid function described above in the Input Variables section, so they can be specified using the same parameters.

Operation of the Fuzzy Navigation Engine

Every time that the user's assigned load changes, the Fuzzy Navigation Engine calculates a new dead-zone radius. The new radius is then immediately updated in Fuzzy P2V and reflected in the user interface. In the experimental scene described below, the dead-zone is

surrounded by a red circle drawn in the center of the CAVE floor, indicating the current size. The calculation is performed as follows, using the fuzzy-lite fuzzy logic library.

Recall from Chapter 2 that a fuzzy inference system evaluates the premise of each rule according to the set membership functions for the input variables. Because each of the rules used in this system has a single premise, no combination is necessary to determine the output of a given rule. However it is still likely that multiple rules may fire. After a firing set has been constructed based on the results of rule evaluation, inference is performed. In the inference step, an OR composition is used to combine the outcomes of all fired rules. After a combined set has been constructed, defuzzification finds the center of gravity to return an actual numeric value for use as the dead-zone radius. The center-of-gravity method was selected for this system because it is commonly used and it has the convenient property that output values vary smoothly along the output scale as degrees of membership change.

Figure 5.3 shows the operation of the Fuzzy Navigation Engine through an example with sample input values. In the example scenario, the current spatial load is 6 items, the current verbal load is 4 items, and the sets have been configured as shown in Table 5.1. In the fuzzification step, the system determines that spatial load $S_i = 6$ is a member of the medium set with degree 0.4 and of the high set with degree 0.6. Because of these memberships, Rules R_2 and R_3 fire in the inference step, with a firing strength of 0.4 and 0.6, respectively. In parallel, fuzzification is performed on the current verbal load ($V_i = 4$), which reveals a membership of degree 1.0 in the verbal medium set. Because of this membership, Rule R_5 fires with a firing strength of 1.0. The sets resulting from the firing of the spatial and verbal rules now must be combined using an OR operation. Finally, in the defuzzification step, the center of gravity is computed, resulting in a defuzzified output value of 1.92 feet (58.52 cm), which will immediately become the dead-zone radius.

Learning

When a scenario begins, the initial dead-zone radius must be set and all fuzzy sets must be configured with some initial values. To prevent surprising the user, conservative values are chosen. No two users of an interface will have the same ability level so the system must learn at runtime how to better adjust its parameters. As a scenario unfolds, the rules stay the same but the system shifts the membership sets based on user performance. The system uses windowed averages so a large amount of error in a short period of time indicates that adjustment is needed.

If a user has problems with locomotion then it may indicate, regardless of how the current implementation has adjusted the navigation, that the dead zone may be too small. If the user

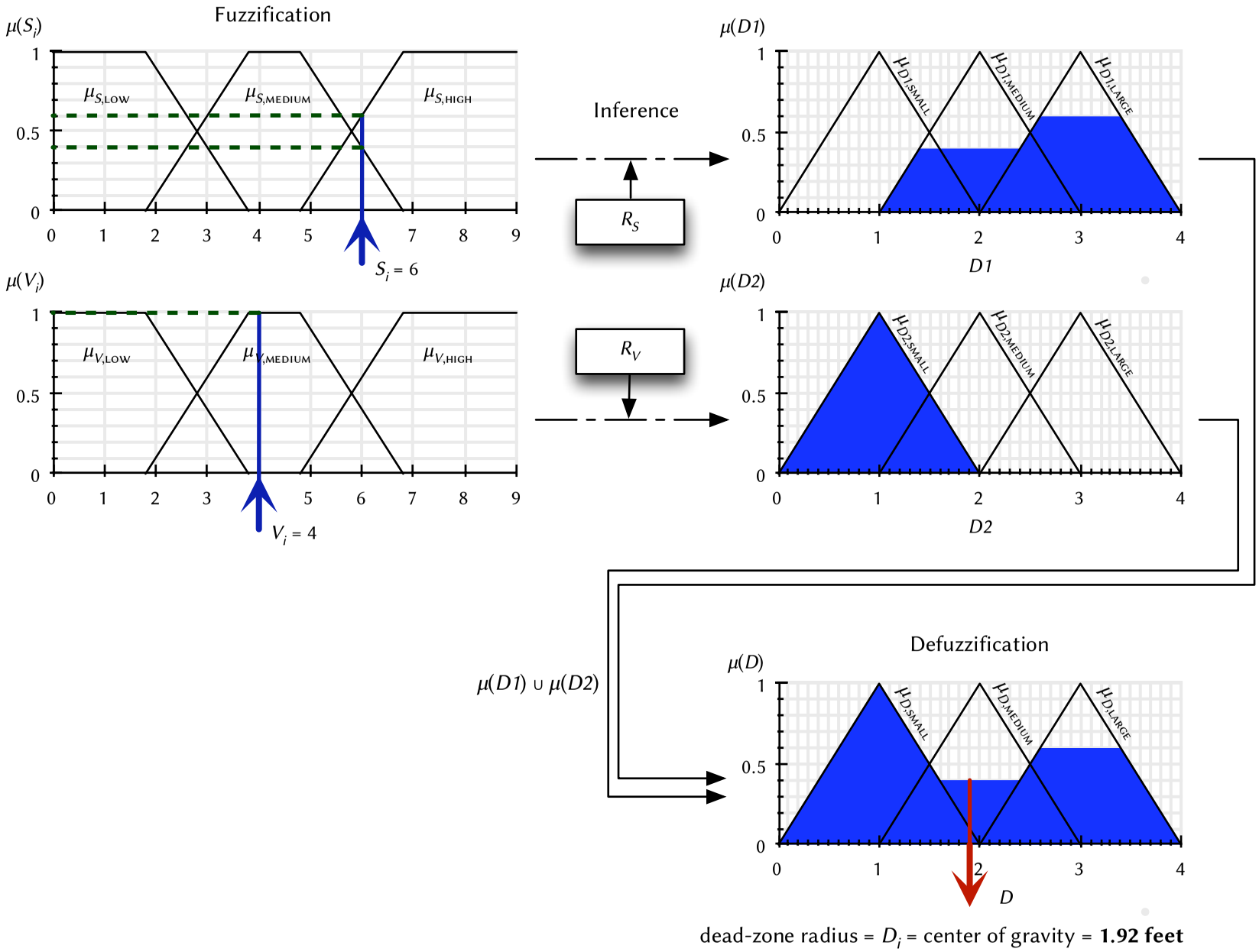

Figure 5.3 Operation of the Fuzzy Navigation Engine, with sample input values.

Table 5.1 Parameter configuration for the example scenario.

	trap,L			trap,R		
Function	m	b	range	m	b	range
$\mu_{S,\text{LOW}}$	–	–	–	−0.5	1.9	1.8–3.8
$\mu_{S,\text{MEDIUM}}$	0.5	−0.9	1.8–3.8	−0.5	3.4	4.8–6.8
$\mu_{S,\text{HIGH}}$	0.5	−2.4	4.8–6.8	–	–	–
$\mu_{V,\text{LOW}}$	–	–	–	−0.5	1.9	1.8–3.8
$\mu_{V,\text{MEDIUM}}$	0.5	−0.9	1.8–3.8	−0.5	3.4	4.8–6.8
$\mu_{V,\text{HIGH}}$	0.5	−2.4	4.8–6.8	–	–	–
$\mu_{D,\text{SMALL}}$	1.0	0.0	0.0–1.0	−1.0	2.0	1.0–2.0
$\mu_{D,\text{MEDIUM}}$	1.0	−1.0	1.0–2.0	−1.0	3.0	2.0–3.0
$\mu_{D,\text{LARGE}}$	1.0	−2.0	2.0–3.0	−1.0	4.0	3.0–4.0

has a concurrent task, it may also be concluded that it is possibly due to competition for resources. If so, this means that the system's current fuzzy sets are inappropriately sized and that the user's load should be viewed as greater than the linguistic fuzzy terms indicate. Adjusting the sets using a negative correction term makes the fuzzy inference system view a given numeric load as higher linguistically. If the sets for the input variables are corrected in this way, assuming the same production rules described above are in place, then the output variable (the dead-zone radius) will tend to be larger.

Sometimes a user may be having very few problems with the interface and could benefit from a smaller dead zone, to increase the maximum possible velocity. This also means that the fuzzy sets may be inappropriately sized. In this case it seems that the user's load should be less than the linguistic fuzzy terms currently indicate. Adjusting the sets using a positive correction term will make the fuzzy inference system view a given numeric load as lower linguistically. When the input sets are corrected in this way, the dead-zone radius will lend to be smaller.

Learning is accomplished by adding a correction term to the input variable as shown by example in Figure 5.4. The amount of correction is determined according to the outputs of the error function described below after a description of the error terms and associated calculations. The Fuzzy Navigation Engine only adjusts the correction term on a working memory resource (verbal or spatial) that was in use at the time that the error was measured. The correction term "expands" or "shrinks" the trapezoid membership functions. For example in the case of spatial correction, the x-intercept for the right side of the low term and the x-intercept for the left side of the medium term will both shift by c_s. The x-intercept for the right side of the medium term and the x-intercept for the left side of the high term will both

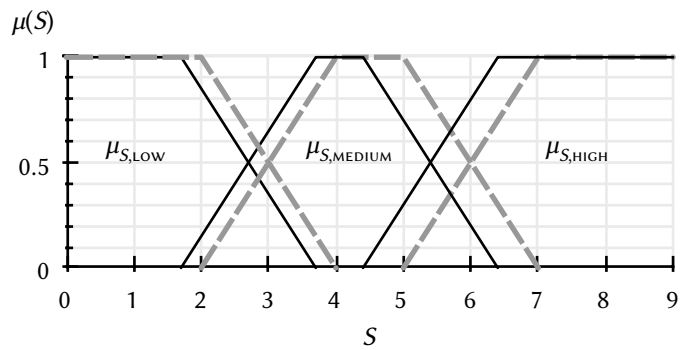

Figure 5.4 Example of variable correction. The spatial-load input variable is being corrected by 0.3.

shift by $c_s \times 2$. In the figure, the spatial sets have all been corrected by -0.3. The solid lines are used to depict the new membership functions while the dashed lines depict the original membership functions (pre-correction).

Adaptive system pilot study

To determine initial parameters and ranges for learning, a pilot study was conducted with learning disabled. The fuzzy inference system was able to adjust the dead-zone radius properly, but it did not learn from user error. During the pilot study, many problems were discovered involving the experiment procedures. The design was iterated as issues were encountered. For example, the length of the rounds and the number of tasks that each participant was asked to complete were adjusted. The results from this pilot study are referenced below in the discussion about error metrics.

Configuration of Fuzzy Navigation Engine

For the pilot study and the formal study, the fuzzy subsets were configured as shown in Figures 5.5, 5.6, and 5.7. Table 5.2 shows the parameters for the trapezoid membership functions. As described above, the output values for the dead-zone radius are in feet. The correction magnitude (c_m) was set to 0.1 and the minimum correction (c_{min}) was set to -1.0.

Participants

Ten participants were recruited through word of mouth and announcements in an undergraduate course. All participants were undergraduate students at Iowa State University. Data from one participant were not used due to a head-tracker malfunction.

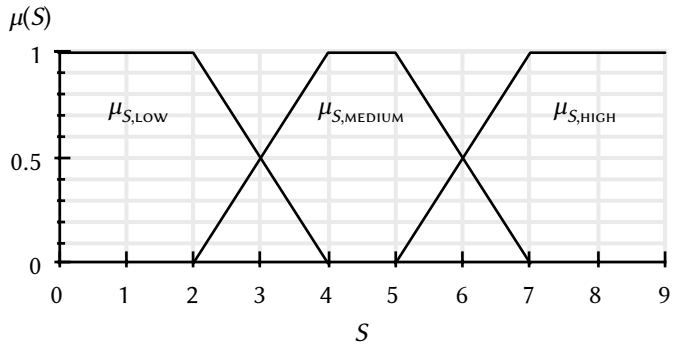

Figure 5.5 Initial fuzzy sets used in Study 3 for the spatial load input variable.

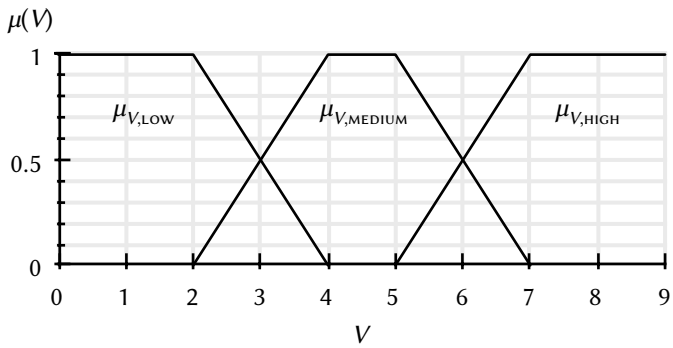

Figure 5.6 Initial fuzzy sets used in Study 3 for the verbal load input variable.

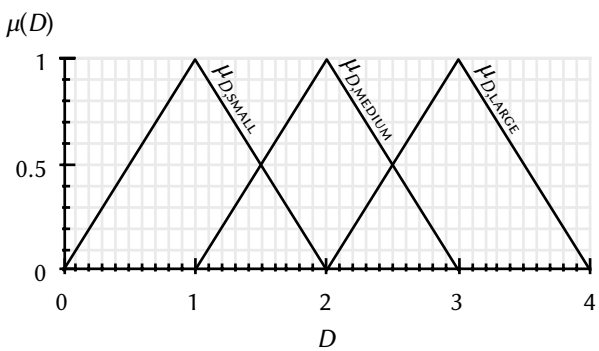

Figure 5.7 Fuzzy sets used in Study 3 for the dead-zone radius.

Table 5.2 The parameter configuration for the adaptive system pilot study and Study 3.

	trap,L			trap,R		
Function	m	b	range	m	b	range
$\mu_{S,\text{LOW}}$	–	–	–	−0.5	2.0	2.0–4.0
$\mu_{S,\text{MEDIUM}}$	0.5	−1.0	2.0–4.0	−0.5	3.5	5.0–7.0
$\mu_{S,\text{HIGH}}$	0.5	−2.5	5.0–7.0	–	–	–
$\mu_{V,\text{LOW}}$	–	–	–	−0.5	2.0	2.0–4.0
$\mu_{V,\text{MEDIUM}}$	0.5	−1.0	2.0–4.0	−0.5	3.5	5.0–7.0
$\mu_{V,\text{HIGH}}$	0.5	−2.5	5.0–7.0	–	–	–
$\mu_{D,\text{SMALL}}$	1.0	0.0	0.0–1.0	−1.0	2.0	1.0–2.0
$\mu_{D,\text{MEDIUM}}$	1.0	−1.0	1.0–2.0	−1.0	3.0	2.0–3.0
$\mu_{D,\text{LARGE}}$	1.0	−2.0	2.0–3.0	−1.0	4.0	3.0–4.0

Methods

A new scene, CogScene, has been created which allows users to traverse a brick corridor, shown in Figure 5.8, while periodically seeing and reciting memory tasks, intended to simulate the existence of concurrent cognitive tasks. The corridor walls were tall enough that users could not see over them, meaning that a participant only saw a small portion of the environment at a given time. Figure 5.9 shows the basic task flow that was repeated some number of times (n). During each trial, the participant should always be moving through the corridor until the stop-sign card is displayed. At the end of each trial, when the stop-sign card is displayed, the participant should come to a stop as quickly as possible. The memory tasks are similar to those used in the previous studies, except for changes to the spatial recall which is described below.

In this study, participants were asked to remember spatial and verbal items at the same time. However, the verbal recall of the spatial positions that was used in the last two studies was not ideal. For this reason, all spatial recall was done using the Logitech Wingman Cordless gamepad that was used for locomotion in Study 1. The six buttons on the right side were covered with red tape so the letters were occluded but the buttons would be clearly visible. The task presentation was a sequence of virtual cards with each circle corresponding to a button on the gamepad. For example, in Figure 5.10, the highlighted circle corresponds to the "Y" button. When it was time to recall the spatial sequence, the card shown in Figure 5.11 was displayed and participants were tasked with pressing the same sequence buttons.

Before the training and experimental scenarios began, each participant completed the same pre-questionnaire used in Studies 1 and 2 (included in Appendix A). Next the user was

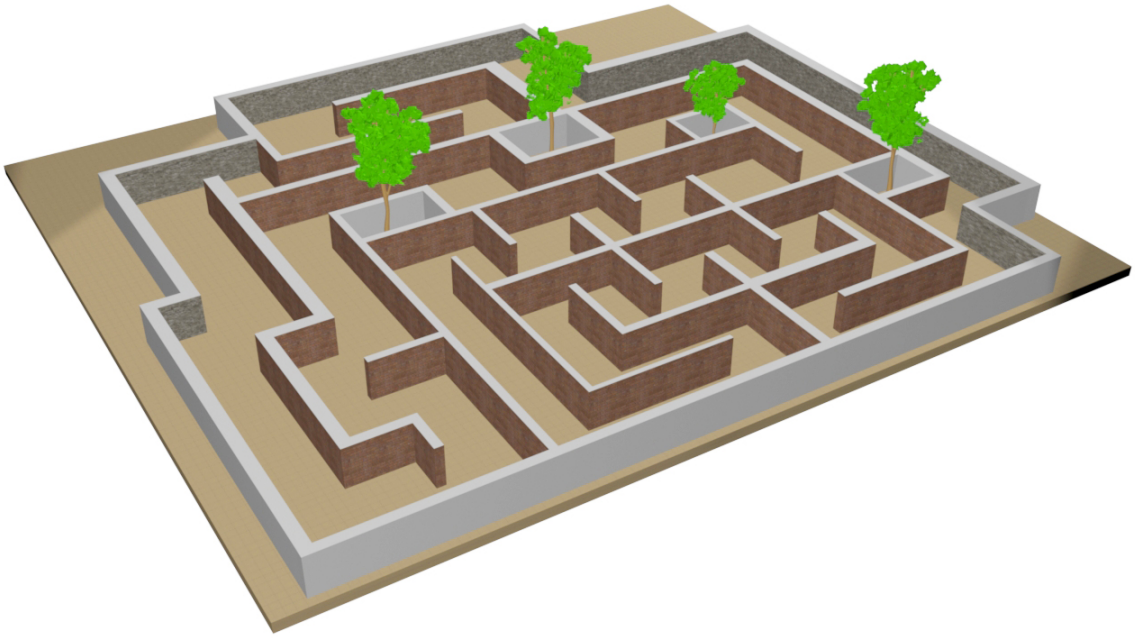

Figure 5.8 The brick corridor traversed in Study 3.

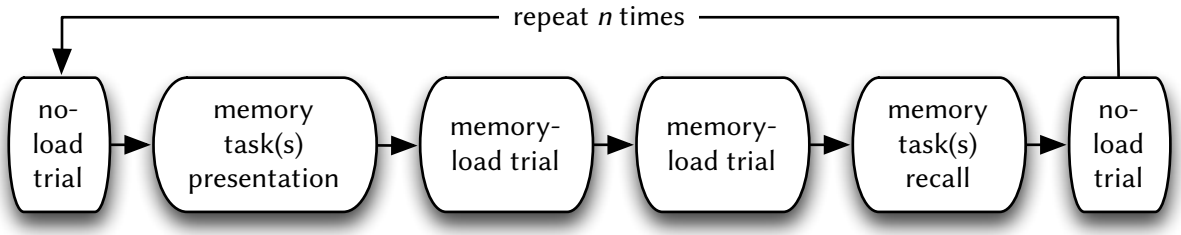

Figure 5.9 Basic task flow for the adaptive system pilot study.

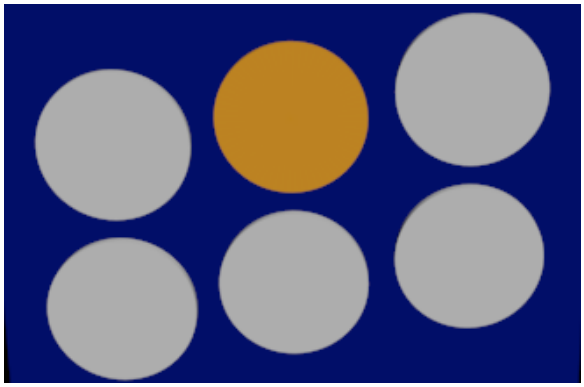

Figure 5.10 Study 3 spatial task presentation card, with the highlighted circle corresponding to the "Y" button on the gamepad.

Figure 5.11 Study 3 spatial recall card.

given a demo of how to move about and complete working memory tasks in the corridor environment. The importance of success on the memory tasks was stressed and the user was also told that movement efficiency, collisions, and stop times would be recorded. Next the participant completed a sequence of practice working memory tasks in the corridor scene, with locomotion disabled. These tasks were intended to give users practice at the memory tasks so they would be comfortable with them in the experimental scenario.

After completing the practice memory tasks, the experimental trials began. One objective of the pilot study was to determine how many trials a participant could comfortably complete in the allotted time (one hour minus questionnaires and training). Because many users exceeded the one-hour time frame or became fatigued, the number of experimental trials was adjusted as the study progressed. For the pilot study, each trial was 20 s long. Experience during the pilot study indicated that incorporating an intermission would allow the participant to complete a greater number of tasks.

When a memory sequence is presented or recalled, CogScene fires an event indicating a change in cognitive resource requirements by the primary tasks. The levels of spatial and verbal resource usage are passed as parameters with the event, indicating the number of spatial and verbal items, respectively, that are currently being remembered by the user.

The Fuzzy P2V interface is currently designed to adjust the dead-zone radius only when the load levels change. Coupled with the logic in CogScene, this means that learning takes place at the end of each trial (using the error function described below), but the dead zone only changes size when a new memory task is presented or an old one is recalled. For the purposes of the study this is ideal for two reasons. First, the user should be standing still in the center of the CAVE during memory task presentation and recall, so there was no change in velocity during active movement or confusion about how far back one must step in order to stop.

Second, changing the dead-zone radius only when the cognitive task changed allowed for more straightforward analysis because a given dead-zone radius could be linked with a given task difficulty and performance measurements. If the system changes very conservatively it may be possible to adjust parameters on the fly but, for many systems, waiting until the user is known to be stopped may be the best solution.

Error and performance metrics

Some error metrics were devised and tested for feasibility in the pilot study. The results from the pilot study also provided an indication of the ranges of values to be expected. The Fuzzy Navigation Engine uses the following three raw error metrics to drive learning:

- the number of collisions in a 15 s window (r_c);
- the time to stop after a stop sign was presented (r_s); and
- the average percent of C6 used over 15 s window (r_p).

Overall locomotion efficiency is the ratio of virtual distance to physical distance. While this ratio is used as a measure of efficiency when evaluating the Fuzzy P2V interface, it is problematic for use in system learning because the range of possible efficiency values is affected by the current dead-zone radius, which is what the fuzzy inference system adjusts.

Collision error term (e_c)

If a user collides with the virtual walls frequently, virtual movement distance is likely to be limited. Also collisions are unlikely to be intended and in some domains locomotion precision may be critical to successful task completion. For these reasons, the number of collisions within a 15 s window is used as a metric for system learning. An examination of the plot of virtual distances versus collisions in Figure 5.12 shows that the above premise seems to be true. The users that achieve very high distances do not typically have a high number of collisions. On the other side of the plot, it appears that some participants had very few collisions but they did not go very far. These users may have been more careful, not pushing the limits of their locomotion abilities.

In Figure 5.13, the variation of dead-zone radius versus number of collisions is presented. This plot generally supports the idea that a larger radius is associated with fewer collisions. In particular notice that participants did not have six collisions in the time window except when they had a very small dead zone (1-foot radius). When moving with a very large dead zone (2.5-foot radius), participants had two or fewer collisions.

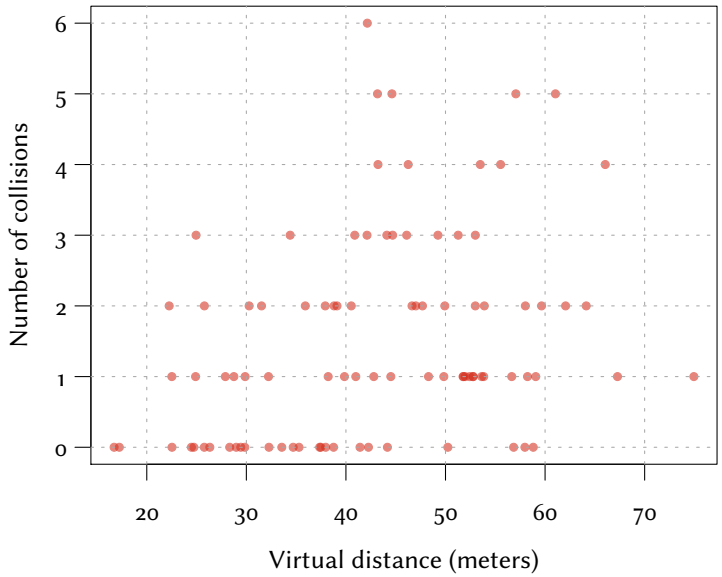

Figure 5.12 Number of collisions as a function of virtual distance traveled in the adaptive system pilot study.

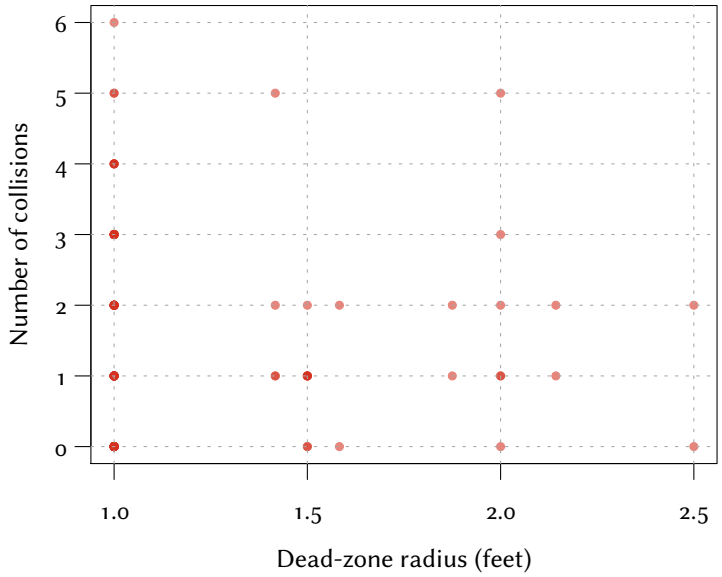

Figure 5.13 Number of collisions as a function of dead-zone radius in the adaptive system pilot study.

Based on these results, a goal was set to make it possible for participants in the study to be able to travel a virtual distance 45.72 m (150 feet) in 15 s. This is higher than the mean, but many users did exceed it in the pilot study. This objective led to capping the upper bound of the error function at 5 collisions. Because the third quartile ends at 2 collisions, that is where the lower bound on the error function was set. This range should effectively capture the outliers, mapping the collision counts of 2–5 to error values of 0–1.

$$e_c = \begin{cases} 0 & \text{if } r_c \leq 2 \\ \dfrac{r_c - 2}{3} & \text{if } 2 < r_c < 5 \\ 1 & \text{if } 5 \leq r_c \end{cases}$$

Adding collision events is fairly straightforward with the existing VirtuTrace experiment platform. The RealWorldPhysics class was modified to keep track of all physics bounding boxes that are currently overlapping. Whenever a new overlap is detected, a collision event is fired. This event is received by the FuzzyNavEngine class, which tracks all locomotion problems.

Stop error term (e_s)

It is known from the results of Study 1 that users have problems stopping quickly using the P2V interface when completing a simultaneous spatial memory task. It was expected that increasing the radius of the dead zone will help users stop more quickly. A plot was constructed (see Figure 5.14) showing the dead-zone radius versus stop time for all pilot trials. The plot shows that the mean is similar for all dead-zone sizes, but when the dead zone is smaller there seem to be more outliers with large stop times. These outliers are what the fuzzy inference system design will seek to prevent and, based on this plot, it seems that adjusting the dead-zone radius will help.

The third quartile of pilot study stop times ends at 4931 ms so it makes sense to start the error function near there. Only three data points are above 30 000 ms so that is a reasonable number for the error function upper bound in the formal study. Based on this rationale, the function for the stop error term in the learning system linearly maps stop times of 5000–30 000 to error values of 0–1.

$$e_s = \begin{cases} 0 & \text{if } r_s \leq 5000 \\ \dfrac{r_s - 5000}{25000} & \text{if } 5000 < r_s < 30000 \\ 1 & \text{if } 30000 \leq r_s \end{cases}$$

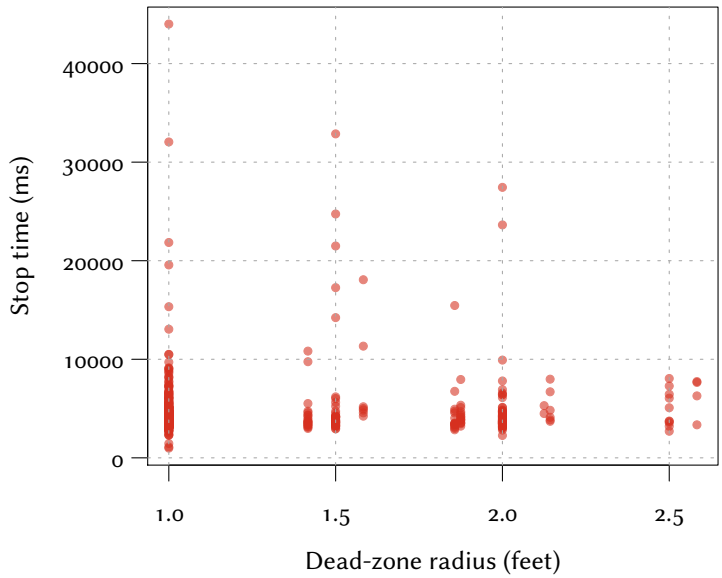

Figure 5.14 Stop time as a function of dead-zone radius in the adaptive system pilot study.

Percent-of-CAVE error term (e_p)

The final metric that used for learning is a windowed measure of the extent to which the horizontal movement area of the C6 is being utilized. This is equal to the windowed average percent of the 1.524 m (5-foot) distance between the center of the C6 and each wall. It is possible for this value to be greater than 100% because a user could move toward the corner of the CAVE and have an average distance of greater than 1.524 m from the center.

Figure 5.15 shows the relationship between dead-zone radius and percent-of-CAVE measurements from the pilot study. The largest dead-zone radius possible in the fuzzy inference system was 3 feet (91.44 cm). When the dead zone is set to this largest size, a user must use 60% of the C6 in order to activate translation in the P2V locomotion interface. The plot seems to reflect this expectation and larger dead-zone sizes generally lead to higher values for this metric. Because the objective is to capture outliers, the lower bound for the error function was set at 70%. It is important to prevent users from running into the physical walls and because the corridor scene is axis-aligned, meaning that movements to the corners of the CAVE are infrequent, the upper bound on the error function was set to 90%, or 15.24 cm (0.5 feet) from the CAVE wall.

$$e_p = \begin{cases} 0 & \text{if } r_p \leq 0.7 \\ \dfrac{r_p - 0.7}{0.2} & \text{if } 0.7 < r_p < 0.9 \\ 1 & \text{if } 0.9 \leq r_p \end{cases}$$

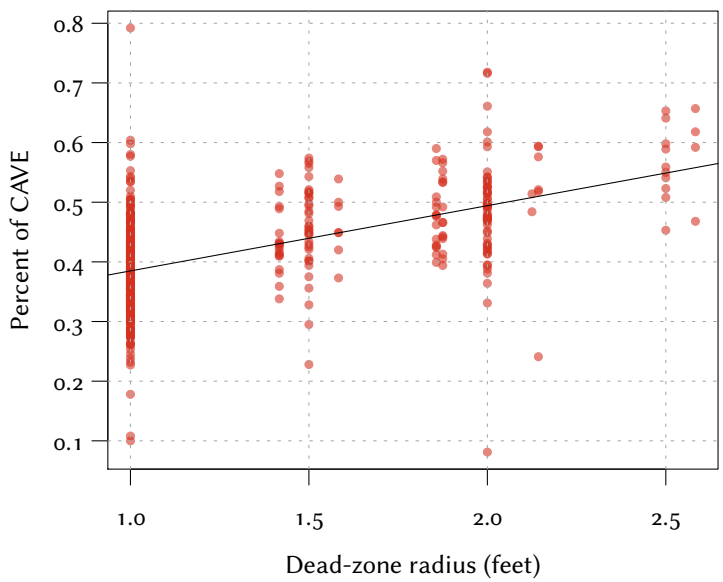

Figure 5.15 Percent of CAVE as a function of dead-zone radius in the adaptive system pilot study.

Combining the error terms

With the above error metrics, the basic input-output flow of the Fuzzy Navigation Engine is shown in Figure 5.16.

The three error terms above are always between 0 and 1. These terms must be combined in a meaningful manner in order to drive learning. The stop error (e_s) and collision error (e_c) terms indicate that the user is performing poorly, so these can be thought of as user error (e_u). The percent-of-CAVE (e_p) error term indicates that the user may be restricted by an overly large dead zone, so this term can be thought of as interface error (e_i) and should serve to counteract user error. The user error and interface error are combined and multiplied by the correction magnitude (c_m). The result is added to the total verbal correction if the current verbal load (l_v) is greater than zero, with the restriction that total verbal correction must never be less than c_{min}. If the current spatial load (l_s) is greater than zero, the result is added to the total spatial correction, also with the restriction that total spatial correction must never be less than c_{min}.

$$e_u = e_c + e_s$$

$$e_i = e_p$$

$$c_{temp} = c_m \times (e_i - e_u)$$

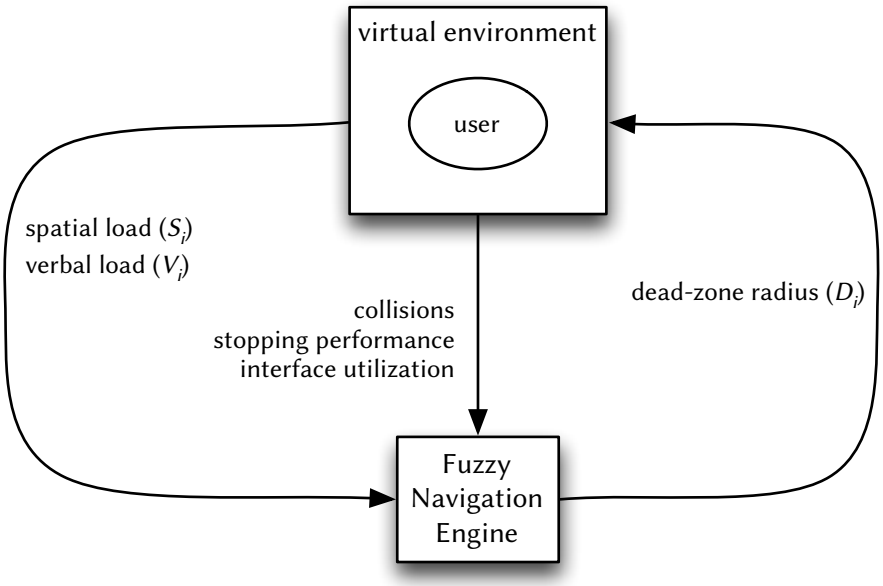

Figure 5.16 Basic input-output flow of the Fuzzy Navigation Engine, including error metrics.

$$c_v = \begin{cases} c_v + c_{temp} & \text{if } l_v > 0 \text{ and } c_v + c_{temp} > c_{min} \\ c_{min} & \text{if } l_v > 0 \text{ and } c_v + c_{temp} \leq c_{min} \\ c_v & \text{if } l_v = 0 \end{cases}$$

$$c_s = \begin{cases} c_s + c_{temp} & \text{if } l_s > 0 \text{ and } c_s + c_{temp} > c_{min} \\ c_{min} & \text{if } l_s > 0 \text{ and } c_s + c_{temp} \leq c_{min} \\ c_s & \text{if } l_s = 0 \end{cases}$$

Study 3

A third formal study was conducted, testing the Fuzzy P2V interface with users to verify that it is beneficial. The adaptation should be considered a success if users of the new fuzzy inference system are able to outperform users of the baseline system at basic locomotion tasks during a concurrent cognitive task load. It was considered likely that results could be mixed, with some benefits and some problems resulting from the use of such a system. Participants in this study were divided into two groups according to the locomotion system in use: Fuzzy P2V (Fuzzy) and P2V. As they were moving through the VE, a series of concurrent spatial and verbal span tasks of random difficulties were presented and recalled.

The ranges for the error metrics were configured as described above. The c_m was set to 0.1 and c_{min} was set to -1.0.

Experiment design

This study was a 2×3 design with two locomotion systems (Fuzzy and P2V) and three levels of working memory load (spatial, verbal, none). Locomotion system was a between-subjects variable due to time constraints and the expected impact of learning. Working memory load involved relatively quick tasks that could all be performed within the environment, so it made sense for that variable to be manipulated within subjects.

There were two between-subject groups: Fuzzy and P2V. Participants in the Fuzzy group used the new fuzzy inference system (with learning enabled) for the entire time that they were in the VE. The fuzzy system started with default baseline settings and adjusted itself in response to problems exhibited by the participant. The P2V group used the same P2V interface that was used in Study 2. It was configured with settings that remained unchanged throughout the participant's travels through the VE.

Methods

Participants

Twenty-six undergraduate students were recruited from the Department of Psychology research participant pool, word of mouth, and announcements in undergraduate courses. Participants came from multiple departments and majors across campus. All participants were required to have 20/20 (corrected) binocular vision.

Procedures

First, participants were asked to complete a pre-questionnaire with topics involving demographic information and video game experience. Then they entered the C6 and were given instructions and a demonstration of how to complete working memory tasks in the VE. There were no pre-assessment memory tasks in this study, because tasks of varying difficulty (1–7 items) were presented during the experimental phase. However, because it was desirable for memory task performance to be high, users were given an opportunity to practice so they would be comfortable and confident when remembering the items during the experimental phase.

Before the experimental phase, the participant was given instructions and a detailed demonstration of the P2V interface and all locomotion tasks. The demonstration took place in a corridor scene similar to the one used in the experimental phase. The user was not informed about the adaptation of the navigation system parameters. The participant was not allowed to practice the locomotion tasks, but there was a run-through in which the experimenter

demonstrated what the user would be required to do. The decision to not allow locomotion practice was made to prevent any learning from taking place before the actual experimental tasks. This would maintain the unnaturalness of the movements and probably the extent to which cognitive resources would be required.

In the experimental phase, the participant was required to traverse a winding corridor, completing memory tasks along the way. The corridor traversal was intended to simulate the types of basic navigation tasks that a user might encounter in a "real-world" VE. Scenes were not switched from locomotion task to memory task as they were in the previous two studies. The memory tasks were similar to those experienced in Studies 1 and 2, except in this case they popped up along the path as the participant traveled through the corridor. The recall cards were also similar to those from previous studies and they were also presented within the same scenario. Participants were instructed to move through the corridor whenever there was no memory card displayed, to stop whenever a stop-sign shape appeared, and to stay stopped whenever a memory task was being presented or recalled. Participants were told that the memory tasks were the most important priority and movement performance would also be recorded. Specifically participants were told that movement efficiency, stop times, and the number of collisions with virtual walls would be recorded.

The study was divided into two halves, with an intermission in between. The task flow of one half is pictured in Figure 5.17. Both halves were identical in structure but all memory tasks were of random difficulty and the sequences were random as well. For the memory task presentation, one of the following was displayed (randomly):

- a spatial task;
- a verbal task;
- a spatial task followed by a verbal task; or
- a verbal task followed by a spatial task.

Two of each of the preceding possibilities was experienced during each half, for a total of eight task loads in each half. Each memory task was of random difficulty (length), containing between one and seven items. Each half took approximately 15 minutes to complete. In the intermission, the scene was paused and participants were given the opportunity to rest if needed. They were also asked to complete just the Likert scale portion of the post-questionnaire included in Appendix A.

After both halves (64 trials) were complete, the participant exited the C6 and was asked to complete a post-questionnaire identical to that used in Studies 1 and 2 (included in Appendix A). After completing the questionnaire, the experimenter asked some questions in

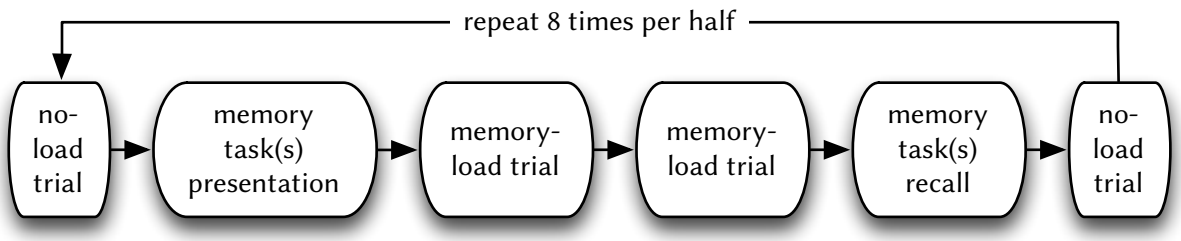

Figure 5.17 Task flow for one half of Study 3. This flow was repeated twice with an intermission in between.

an unstructured interview. Topics covered involved strategies used on the working memory tasks as well as overall opinions and suggestions regarding the locomotion interface.

Response variables

In this study there were only two types of task: memory and locomotion. No attempt was made to differentiate translation from rotation. All participant responses on the spatial and verbal memory tasks were recorded and checked for correctness. The following response variables were used for the locomotion tasks.

Stop time Time from presentation of the stop card until the user was completely stopped.

Number of collisions Number of collisions in a 15 s window preceding presentation of each stop card.

Percent of CAVE Average percentage of CAVE used in a 15 s window preceding presentation of each stop card.

Physical distance Total physical distance traveled in a 15 s window preceding presentation of each stop card.

Virtual distance Total virtual distance traveled in a 15 s window preceding presentation of each stop card.

As described above, the physical distance and virtual distance measurements were used to calculate efficiency.

Logging

Recall that there were two halves with 32 movement trials each, for a total of 64 trials. The participant's head position was logged in every frame and the experimenters watched the

scenario on the head node for a subjective interpretation of problems encountered by the user.

Due to the length of the experimental phase and the heat generated by the 24 projectors, several participants experienced simulator sickness and were excused early. Even in these cases, participants completed many trials so the measurements up to that point were still included in the analysis. One note, however, is that this led to fewer participants completing trials in round two than in round one. This affected the Fuzzy group slightly more than the P2V group but it was not clear if the Fuzzy P2V interface led to increased simulator sickness. In the P2V group, 12 participants (out of 13) made it to the second round and 10 of those completed the entire scenario. In the Fuzzy group, 8 participants (out of 13) made it to the second round and 7 made it to the end of the scenario. The total number of trials with the P2V interface was 747 and the total number with the Fuzzy P2V interface was 654.

There were many instances of the tracking system losing track of a participant's physical position in the CAVE. In some cases this was isolated to interference with cellular phones and in other cases it may have been because of strategies employed by a participant involving rapid movements near the extents of the tracked area. For example, some users employed a lunging technique which sometimes evaded the tracker because the user's head was near the physical wall and relatively low to the ground. This position is far from the optimal tracking area in the center of the CAVE. The XML experiment log files were parsed with a Python script which was capable of identifying likely head-tracker malfunctions. The script inspected the participants' alleged head positions to see if they were within reasonable bounds. If not, the affected results were flagged and discarded if they corresponded with subjective observations by the experimenters. Note however that even if the data points were removed, they were still used to drive system learning in some cases. For example, if the head tracker failed and caused a user to collide with a virtual wall, that collision was still included in the online error calculations.

Results

Data analysis focused on verifying that the Fuzzy Navigation Engine was functioning properly, checking for an improvement in user performance with the Fuzzy P2V interface over the P2V interface, and assessing the choice of error metrics and ranges.

Dead-zone radius

The dead-zone radius was fixed at 1.5 feet (45.72 cm) for all trials of the P2V group. This radius was chosen as a nominal "best" based on experience from past studies and pilot studies. The

Table 5.3 ANOVA table for Study 3 efficiency.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	1	3.95	.06
Error	24		
Within subjects			
Round (R)	1	0.91	.34
R × I	1	<0.001	.98
Error	1326	(6.61)	

Note. Value enclosed in parentheses represents mean square error.

mean dead-zone radius for trials in the Fuzzy group was 1.292 feet (39.38 cm). The results reported in the following subsections should be taken in context with this information. All other things being equal, it is generally a positive thing for the dead zone to be small because it gives users more performance capability, i.e., higher potential maximum velocity.

Efficiency

Because the dead-zone radius for the P2V group was larger, on average, than that for the Fuzzy group, it was expected that those users would travel larger physical distances. The larger dead-zone radius also meant that the P2V region was slightly smaller due to the CAVE boundaries, though it was unlikely to have much impact because users rarely get close to the physical walls.

A two-factor mixed-model analysis was conducted on efficiency with fixed effects for locomotion interface group and round combinations (4 means) and a random effect for subject. The results are shown in Table 5.3. The analysis showed a marginally significant main effect of interface [$F(1, 24) = 3.95$, $p = .06$]. The means are plotted in Figure 5.18, showing that users of the Fuzzy P2V interface moved more efficiently than users of the P2V interface.

To better understand the efficiency results above, additional information was needed regarding the physical and virtual distance. First a two-factor mixed-model analysis was conducted on physical distance. The analysis, shown in Table 5.4, revealed a significant main effect of round [$F(1, 1326) = 7.23$, $p = .007$], a significant interaction between interface group and round [$F(1, 1326) = 8.10$, $p = .004$], and a marginally significant main effect of group [$F(1, 24) = 3.18$, $p = .09$].

Next, a two-factor mixed-model analysis was conducted on virtual distance. The analysis, shown in Table 5.5, revealed a significant interaction of interface group and round

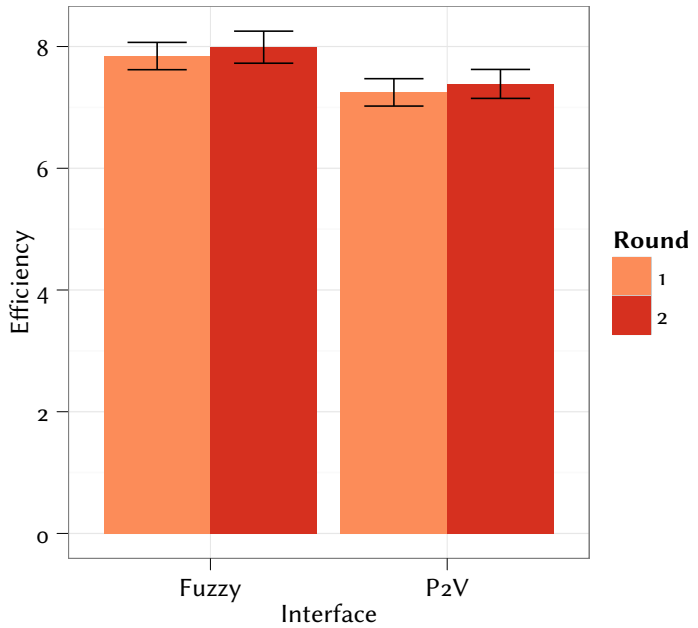

Figure 5.18 Study 3 mean efficiency as a function of interface and round. Error bars show ±1 standard error of the mean.

Table 5.4 ANOVA table for Study 3 physical distance.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	1	3.18	.09
Error	24		
Within subjects			
Round (R)	1	7.23**	.007
R × I	1	8.10**	.004
Error	1326	(23.70)	

Note. Value enclosed in parentheses represents mean square error.
**$p < .01$.

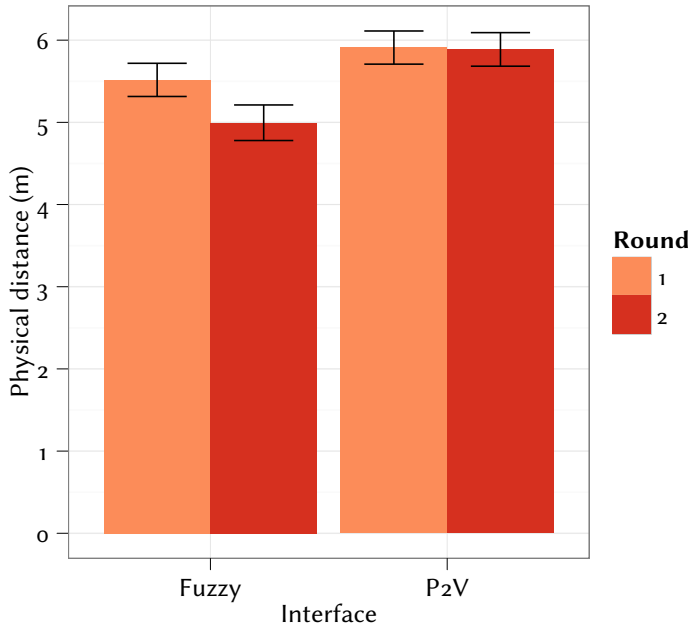

Figure 5.19 Study 3 mean physical distance as a function of interface and round. Error bars show ±1 standard error of the mean.

[$F(1, 1326) = 8.32$, $p = .004$], but the main effects were not significant. A plot of the means is shown in Figure 5.20.

This pattern of distance and efficiency results indicates that users of the Fuzzy P2V interface were more efficient. It seems that participants used less physical input and, even with the lower virtual distance, achieved a higher efficiency. It is unclear why the interaction exists in the physical and virtual distance data. It appears that users in the Fuzzy group may not have tried to (and thus did not) move as far in the second round. Taken together with second-round performance improvements described below, a trade-off may exist, warranting additional research.

Stop time

If all other aspects were equal, the larger mean dead-zone radius in the P2V group should have made stopping easier, leading to lower stop times. However, one objective of the Fuzzy P2V interface was to increase the dead-zone radius when needed due to the user's concurrent task load. This means that the interface should be considered a success (at least with respect to stopping) if stop times are lower.

A two-factor mixed-model analysis was conducted with fixed effects for locomotion interface and round (4 means) and a random effect for subject. This analysis, shown in Table 5.6,

Table 5.5 ANOVA table for Study 3 virtual distance.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	1	0.11	.74
Error	24		
Within subjects			
Round (R)	1	0.37	.54
R × I	1	8.32**	.004
Error	1326	(1033.5)	

Note. Value enclosed in parentheses represents mean square error.
**$p < .01$.

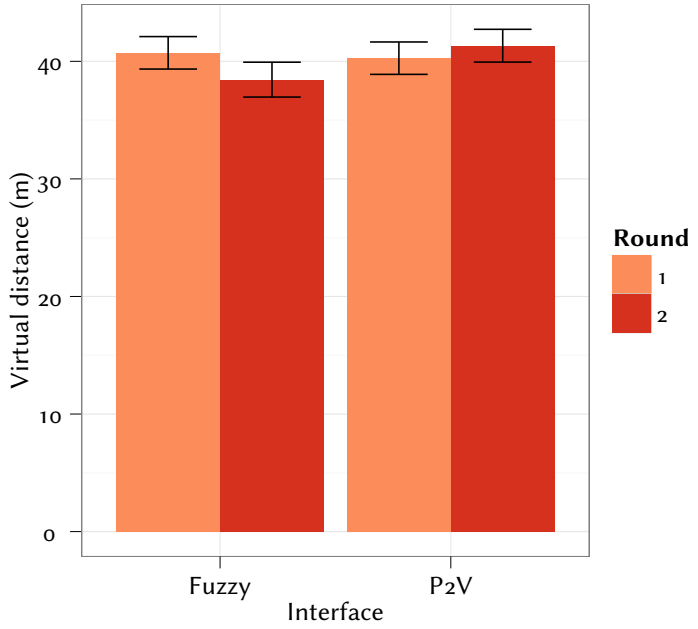

Figure 5.20 Study 3 mean virtual distance as a function of interface and round. Error bars show ±1 standard error of the mean.

Table 5.6 ANOVA table for Study 3 stop times.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	1	3.04	.09
Error	24		
Within subjects			
Round (R)	1	4.88*	.03
R × I	1	0.43	.51
Error	1331	(5138378)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$.

Table 5.7 ANOVA table for Study 3 stop error.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	1	4.57*	.04
Error	24		
Within subjects			
Round (R)	1	5.55*	.01
R × I	1	1.30	.25
Error	1332	(0.006)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$.

indicates a significant main effect of round [$F(1, 1331) = 4.88$, $p = .03$] and a marginally significant main effect of interface group [$F(1, 24) = 3.04$, $p = .09$]. A corresponding plot is shown in Figure 5.21.

Recall that the system was not configured to directly lower the mean stop time, but to reduce the occurrence of outliers which were were quantified using an error term that linearly mapped the range 5000–30000 to 0–1. These outliers represent users who had particular problems using the P2V interface to stop. Another two-factor mixed-model analysis was conducted with this error term as the response variable, revealing significant main effects of interface [$F(1, 24) = 4.57$, $p = .04$] and round [$F(1, 1332) = 5.55$, $p = .01$], as shown in Table 5.7. The pattern of means in Figure 5.22 indicates that the stop error term is lower for users of the Fuzzy P2V interface than for users of the P2V interface.

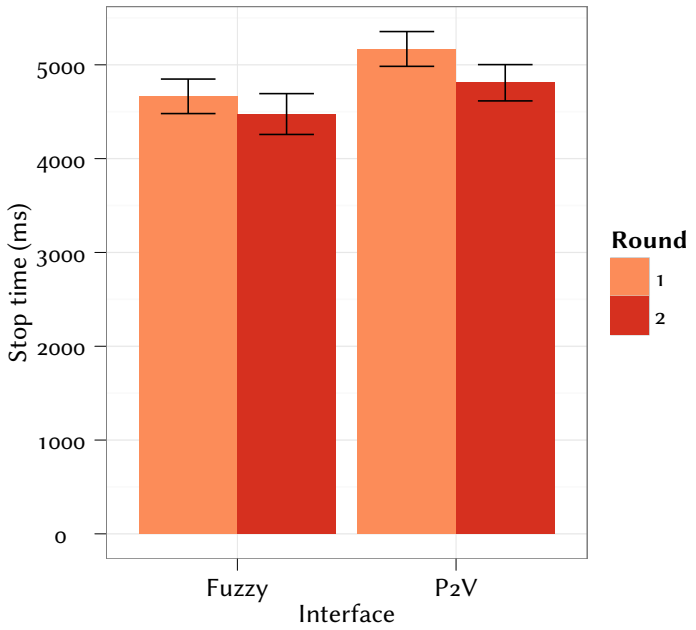

Figure 5.21 Study 3 mean stop time as a function of interface and round. Error bars show ±1 standard error of the mean.

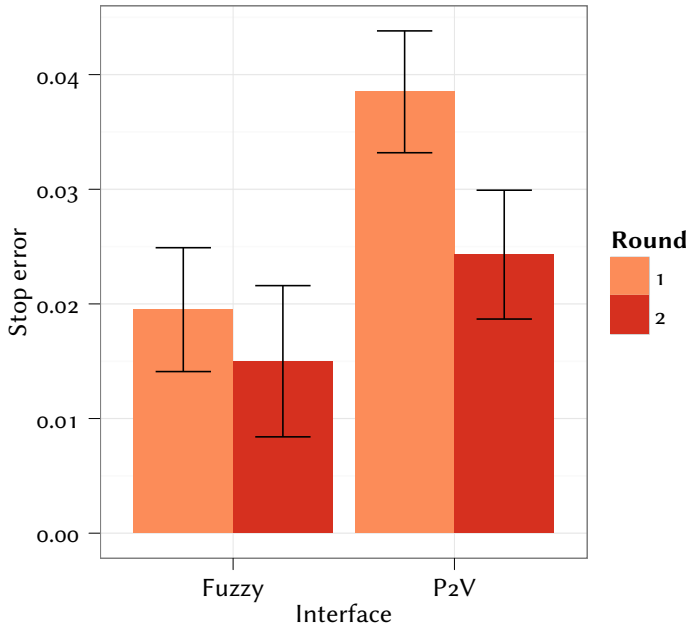

Figure 5.22 Study 3 mean stop error as a function of interface and round. Error bars show ±1 standard error of the mean.

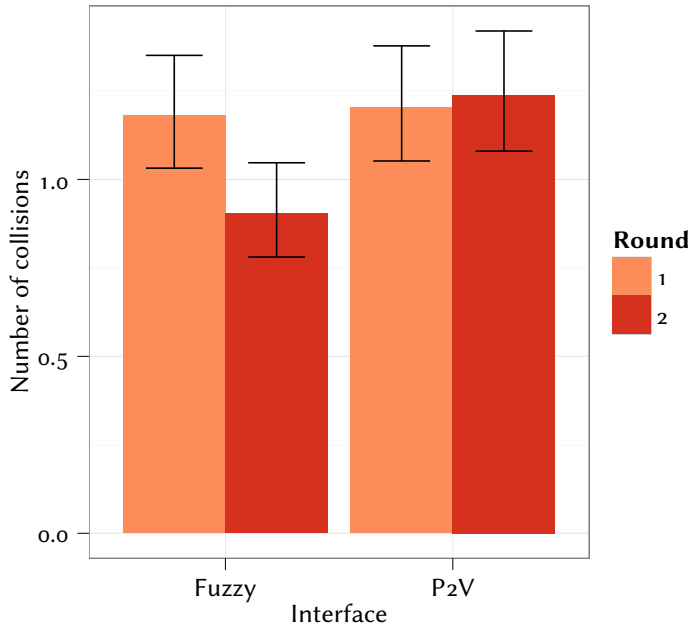

Figure 5.23 Study 3 mean number of collisions as a function of interface and round. Error bars show ±1 standard error of the mean.

Number of collisions

A mixed-model analysis was conducted on the number of collisions, treated as a Poisson response. This revealed a significant main effect of round [$F(1, 1328) = 4.96$, $p = .03$] and a significant interaction between interface group and round [$F(1, 1328) = 7.57$, $p = .01$]. A plot of these results, shown in Figure 5.23, indicates that participants in the Fuzzy group reduced collisions in round two while those in the P2V group did not.

While it can be argued that any number of collisions is often a bad thing and they should be minimized, recall that the system was configured to specifically prevent collision counts greater than two, which were mapped to an error term. A two-factor mixed-model analysis was conducted with this collision error term as the response variable. As shown in Table 5.8, the analysis revealed a significant interaction between interface group and round [$F(1, 1328) = 7.04$, $p = .01$]. The plot in Figure 5.24 shows that the significant interaction seems to be due to a reduction in collisions from round one to round two in the Fuzzy group while the opposite pattern exists in the P2V group. It seems that participants in the Fuzzy group did better at learning to avoid collisions, perhaps because more cognitive resources were available to be allocated for this purpose. Alternately, the changing dead-zone size may have made collision avoidance easier by restricting the users' maximum speed when

Table 5.8 ANOVA table for Study 3 collision error.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	1	0.13	.72
Error	24		
Within subjects			
Round (R)	1	0.25	.62
R × I	1	7.04**	.008
Error	1328	(0.042)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .01$.

resources were in demand by concurrent tasks. In any case, collisions impede virtual travel so this is a promising result.

Incorrect memory sequences

In this study, there was a large range of possible memory spans (1–7). For this reason, the analysis treated the responses as binomial, simply reflecting correctness of the entire sequence.

A two-factor mixed-model analysis was conducted on incorrectness of spatial tasks with fixed effects for locomotion interface and group (4 means), and a random effect for subject. The analysis revealed no significant main effects or interactions. It seems that using the Fuzzy P2V interface had no impact on a participant's ability to remember a spatial sequence. However, the performance was very low in both groups, possibly indicating problems with using the gamepad to respond. For example, it is possible that some users did not press the buttons firmly enough so that all responses were recorded.

Another two-factor mixed-model analysis was conducted, this time with incorrectness on the verbal memory tasks as the response variable. The analysis revealed a significant main effect of interface group [$F(1, 24) = 12.34$, $p = .002$] and a marginally significant effect of round [$F(1, 233) = 2.89$, $p = .09$] and of the interaction between round and interface group [$F(1, 233) = 3.10$, $p = .08$]. The plot in Figure 5.25 indicates that these differences are driven primarily by extremely low performance in round two of using the Fuzzy P2V interface. As mentioned above, it should be noted that only eight participants (out of 13 total) in this group made it to round two without being dismissed early due to simulator sickness. Previous results fail to predict this pattern of means so it is possible that the inaccuracy is indicative of fatigue, boredom, or reduced motivation in the later trials.

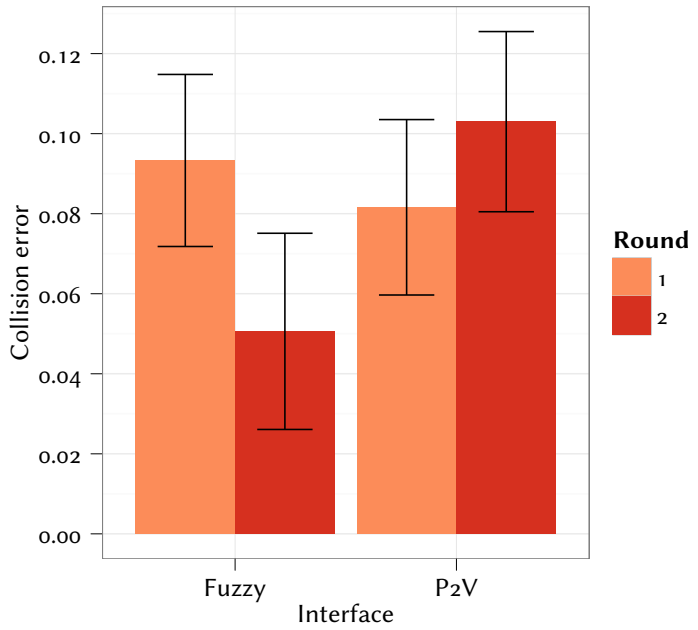

Figure 5.24 Study 3 mean collision error as a function of interface and round. Error bars show ±1 standard error of the mean.

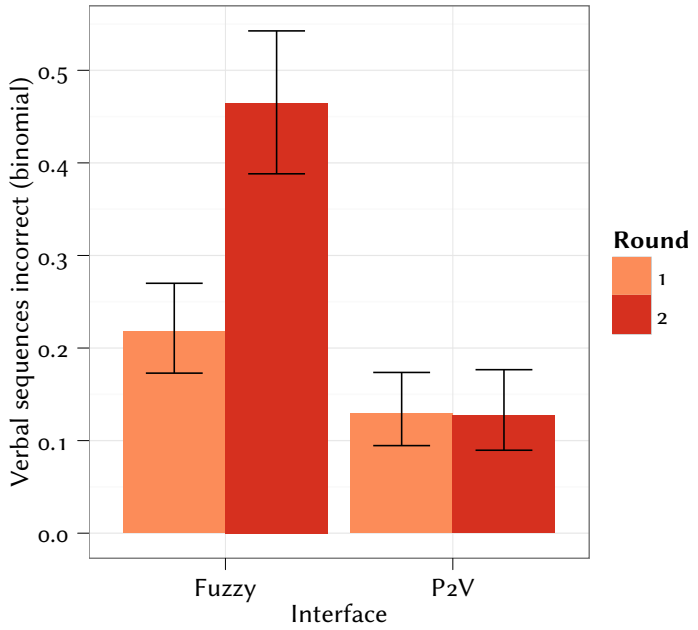

Figure 5.25 Study 3 verbal sequences incorrect (binomial) as a function of interface and round. Error bars show ±1 standard error of the mean.

Effectiveness of learning

Participants in the formal study did not tend to use as much of the C6 as was used during the pilot study. This resulted in very few participants ever having percent-of-CAVE values greater than zero. Of those trials that did have larger values, there was often a known or suspected head-tracker malfunction. For this reason, no analysis was performed on the percent-of-CAVE metric. As mentioned above, the pilot study led to changes in various aspects of the flow of the experiment and also a different corridor model was used in the formal study than was experienced by most pilot users. Some combination of these factors may have led to users not needing or wanting to use as large a physical area.

The results above provide evidence that adjusting the dead-zone radius according to the defined fuzzy rules and sets has been generally helpful in terms of locomotion performance, but these analyses have not directly assessed the extent to which the system was effective at improving itself. One way to measure how well the system learned is to look at the absolute value of both the new verbal and spatial set corrections for each trial. As the system converges on an optimal setting, the absolute value of new correction in each trial should tend to decrease, meaning that values should be lower in the second round if the system is learning effectively.

Recall that there were two broad types of error described above: participant error (collision error and stop error) and interface error (percent-of CAVE error). User error means that the dead zone should be larger while interface error means that the dead zone should be smaller. Unfortunately, because interface error was rare, lower absolute new correction values may really mean that the user error is decreasing. In this way, participant learning may be confounded with system learning. However, a lower absolute correction value in round two than in round one would still reflect positively on the system.

Recall that error was calculated in each trial but correction was only calculated for a given memory type (spatial or verbal) if a task was assigned in that trial. A two-factor mixed-model analysis was conducted on absolute new spatial correction in only those trials where spatial correction was possible (a spatial task was assigned), with fixed effects for locomotion interface and round (4 means), and a random effect for subject. The analysis, shown in Table 5.9, revealed no significant main effects but a significant interaction was found between interface group and round [$F(1, 496) = 4.92$, $p = .03$]. As seen in Figure 5.26, there is a large drop in correction values from round 1 to round 2 in the Fuzzy group. Lower absolute new correction values are an indication that the system may be converging on more appropriate fuzzy input sets. This reduction is not seen in the P2V group.

An identical model was set up for an analysis of absolute new verbal correction values.

Table 5.9 ANOVA table for Study 3 absolute new spatial correction.

Source	*df*	*F*	*p*
Between subjects			
Interface (I)	1	1.49	.23
Error	24		
Within subjects			
Round (R)	1	0.52	.47
R × I	1	4.92*	.03
Error	496	(0.0004)	

Note. Value enclosed in parentheses represents mean square error. $^{*}p < .05$.

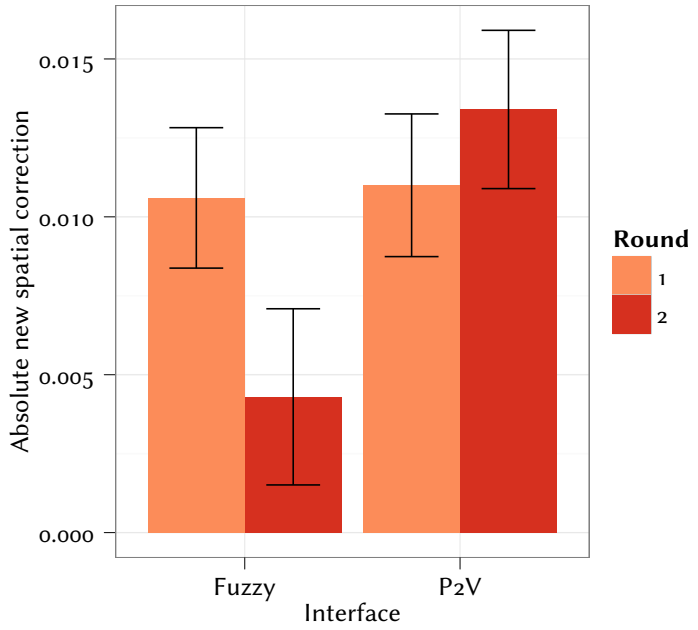

Figure 5.26 Study 3 mean absolute new spatial correction as a function of interface and round. Error bars show ±1 standard error of the mean.

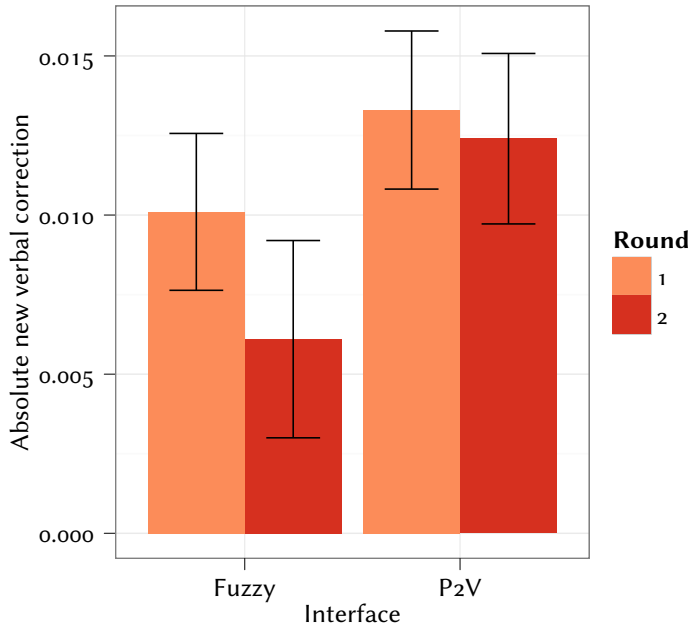

Figure 5.27 Study 3 mean absolute new verbal correction as a function of interface and round. Error bars show ±1 standard error of the mean.

The pattern of means, shown in Figure 5.27, looks similar to that seen above for the absolute new spatial correction, however the analysis revealed no significant main effects or interaction.

Follow-up trials

After analyzing the results from the fuzzy system study, some parameters were modified and two more participants used the system. The following changes were made:

- the dead-zone membership functions were configured as shown in Figure 5.28 and Table 5.10; and
- the r_p range for the percent-of-CAVE error term was changed to 0.4–0.8, as shown below.

$$e_p = \begin{cases} 0 & \text{if } r_p \leq 0.4 \\ \dfrac{r_p - 0.4}{0.4} & \text{if } 0.4 < r_p < 0.8 \\ 1 & \text{if } 0.8 \leq r_p \end{cases}$$

The primary objective of these adjustments was to make the range of the percent-of-CAVE error term more appropriate so that positive and negative new correction values would be

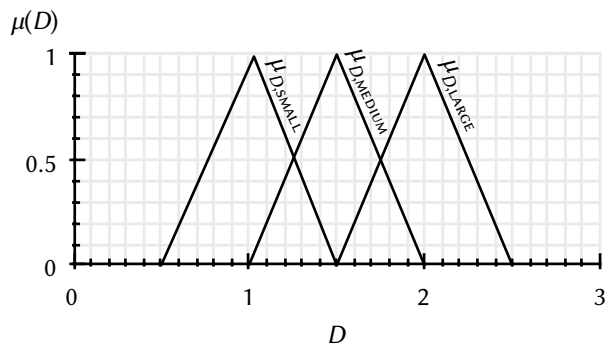

Figure 5.28 Dead-zone membership functions for the follow-up trials.

Table 5.10 Configuration of dead-zone membership functions for the follow-up trials.

	trap,L			trap,R		
Function	m	b	range	m	b	range
$\mu_{D,\text{SMALL}}$	2.0	−1.0	0.5–1.0	−2.0	3.0	1.0–1.5
$\mu_{D,\text{MEDIUM}}$	2.0	−2.0	1.0–1.5	−2.0	4.0	1.5–2.0
$\mu_{D,\text{LARGE}}$	2.0	−3.0	1.5–2.0	−2.0	5.0	2.0–2.5

generated, allowing the system to begin to converge. The output set sizes were adjusted as well because observations during the formal study indicated that the dead zone may have confused some users by changing too drastically. Follow-up trials were conducted in order to test these new settings. Both participants in these follow-up trials used the Fuzzy P2V interface. The demographics of these users were different than for the formal study described above. The first user was a graduate student who was somewhat familiar with the C6 but had not spent much time in it and had never used the P2V interface. The second user was an undergraduate student who had more experience with the operation of the C6 and had a basic understanding of the P2V interface. He had used the interface previously but in several short segments totaling less than 15 minutes. The performance of the two participants was expected to be different due to the difference in experience with the interface. The following changes were made to the experiment flow:

- only 32 trials were assigned (the first user only made it through 30 before experiencing simulator sickness);
- there was no intermission because there were fewer trials; and
- there were no post-questionnaires.

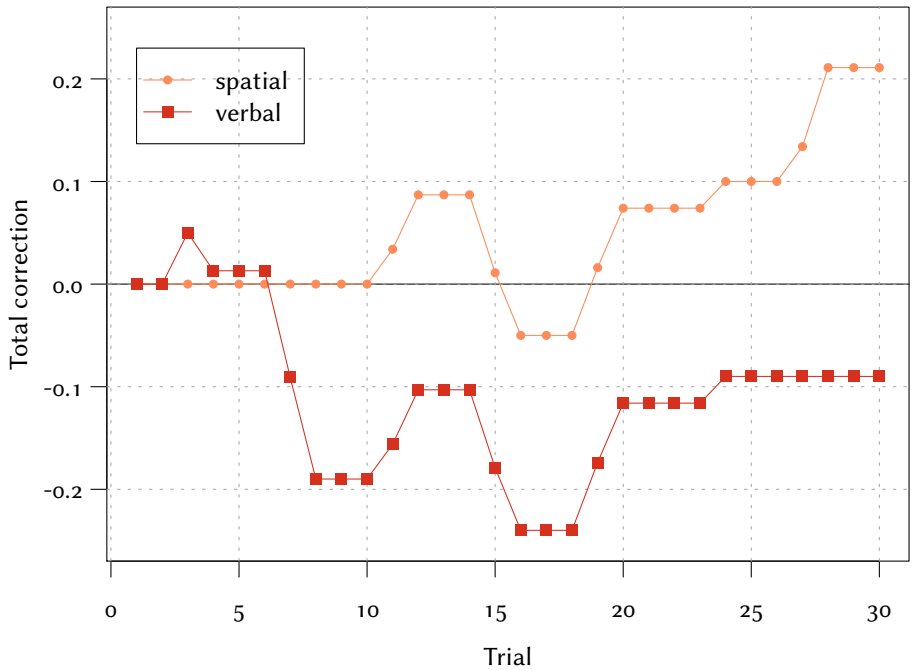

Figure 5.29 One participant's total spatial and verbal correction with the new configuration.

These trials were intended for the sole purpose of examining the total variable correction values after the configuration change. A plot, shown in Figure 5.29, was created to track these total correction values for the first user. The plot shows that the error function drove the total verbal and spatial correction in somewhat different directions. First, recall that there is no change to a total correction value if there is currently none of the respective type of load (verbal or spatial). This is why there is no change to the spatial for the first 10 trials and no change to the verbal for the last 6 trials. The plot shows how this allows the total correction of each variable to behave differently, in this case with total verbal correction being negative and total spatial correction being positive. The plot shows that the participant had some locomotion troubles at first. He generated some collision error and some stop error that counteracted his percent-of-CAVE error, meaning he would potentially benefit from a larger dead zone. After about 15–20 trials (about 8 minutes), it appears that he improved to the point where the percent-of-CAVE error was greater than the sum of the collision error and stop error. Using a large percentage of the interface but not making many mistakes means that he may benefit from a smaller dead zone.

Conclusions

A fuzzy inference system has been created based on the results of Studies 1 and 2. The system adjusts the dead-zone radius for the P2V interface based on knowledge of the user's current cognitive task load. The experiment results described above show that users of the Fuzzy P2V interface performed better on key performance metrics than users of a baseline, P2V interface. On some metrics, it also appears that users of the Fuzzy P2V interface improved more from round one to round two. These results show the potential of the fuzzy inference system to improve users' locomotion performance.

There were some results that are not easily explainable based on past research and expectations. For example, there was inexplicably low verbal memory task performance in the second round when using the Fuzzy P2V interface. If this effect is real, it will be important to do more research to better understand the implications for future systems. It would be interesting to explore if there was a trade-off in which users resorted to a verbally demanding strategy to improve upon stopping and collisions while using the Fuzzy P2V interface. Another possible explanation for this pattern of results is that participants did not perceive resource competition during a verbal task so in the second round they were trying to figure out how the interface was adapting, which may have required verbal resources or general attention resources.

Unfortunately, due to what seems to have been an inappropriate configuration of the percent-of-CAVE error function, learning typically only went one way for participants in the study. Results from an additional follow-up user did seem to follow expected trends after adjusting the problematic settings. Also the analysis of absolute new correction values does provide evidence that the fuzzy inference system may be adjusting itself effectively, thus reducing the amount of needed correction to the input sets. The head-tracker problems that were encountered during the experiment actually demonstrate robustness. Because these new correction values were affected by extraneous head positions, it is encouraging that the absolute correction values are still relatively low. This may be an indication that the system adjusts itself conservatively enough that an occasional outlier does not hinder the learning.

In the future, the same basic fuzzy inference system can possibly be extended by adding additional output variables. One example involves the problems that users had when ducking with a concurrent spatial task. Perhaps a future version of the Fuzzy P2V interface could incorporate a ducking gain of some sort, to facilitate ducking when users were likely to experience problems. The challenge in implementing such an addition will likely be making the application detect when a user intends to duck.

Improvements may also be possible to the dead-zone adaptation demonstrated here. For

example, only the dead-zone radius was manipulated in this research and it was a symmetric adjustment. It is possible that other aspects of the P2V interface could be adjusted, such as the control-display gain, though care must be taken not to hinder the user's learning process. Technically, it is also possible that the dimensions of the dead zone could be asymmetric. For example, perhaps natural walking is more important for forward/backward movement than left/right movement. This presents some possible implementation problems, such as how to differentiate between the direction a user's head is facing, the direction the body is facing, and the direction of intended movement. It may also confuse the user, particularly in the case of rotations performed while outside of the position-to-velocity region.

For real-world use, more research should be conducted to learn how to more accurately assess current utilization of working memory resources. In some domains, such as piloting unmanned aerial vehicles on search and rescue missions, keeping count of the entities that a user must track may be sufficient for a rough estimate of load. However, once unexpected combat occurs, load would become very unpredictable and impossible to estimate using naive methods. A future possibility would be to incorporate pupillometry or other physiological measures, as described in (Grimes, Tan, Hudson, Shenoy, & Rao, 2008; Hirshfield et al., 2009; Yun, Shastri, Pavlidis, & Deng, 2009), for a true augmented cognition system. However, the power of the system described in this chapter lies partly in its use of basic, easily assessed metrics.

Chapter 6. Conclusion

This research shows that competition for cognitive resources can influence the effectiveness of locomotion interfaces for VR systems while reducing performance at concurrent tasks. It then shows that the problems can be mitigated by an adaptive system. This dissertation first described two studies, using the dual-task selective interference paradigm, aimed at understanding specific cognitive demands of locomotion tasks and the associated detriment to performance on locomotion activities as well as concurrent tasks. The problems identified in the studies were then used to inform the design of an adaptive system intended to mitigate those problems. The adaptive system was tested in pilot trials and a formal user study.

Contributions

The contributions from this work lie in two main areas: understanding the impact of concurrent cognitive tasks on locomotion performance when using an unnatural interface; and the design of an adaptive interface in an attempt to mitigate those performance problems.

All three studies made use of the dual-task selective-interference paradigm from the Psychology domain to assess cognitive demands while using a locomotion interface. This in itself is novel. While the paradigm has been used extensively to study basic tasks in the Psychology realm, it has not been used for this type of immersive VR study. Previous research has acknowledged the cognitive demands associated with manipulating an unnatural interface. It is also widely accepted that those demands would interfere with concurrent ongoing tasks. Finally, previous research has identified the problem that virtual locomotion poses with respect to infinite navigation and unnatural interfaces. However, no attempt was previously made to understand the details of these demands. This dissertation has shown that the dual-task selective-interference paradigm is effective for isolating the specific competition for cognitive resources that exists when using a semi-natural locomotion interface.

These findings indicate that unnatural aspects of locomotion require spatial cognitive resources as opposed to verbal resources or general attention resources. First, the work showed that completing simultaneous cognitive tasks does indeed hinder performance using a semi-natural locomotion interface, and the reverse is true as well. Specifically, a clear decrease in

performance at a spatial memory task was shown when using semi-natural interfaces, with interfaces that are generally regarded as being less natural causing greater problems. Additionally, reduced performance was observed on aspects of locomotion while the user was concurrently completing a spatial memory task. Stopping was shown to be slower when using a body-based interface that required locating and returning to the center of the CAVE and ducking to avoid overhead objects was shown to be less successful in all studied interfaces. They also reiterate the importance of understanding specific aspects of resource usage when making interface decisions as a developer, as some interfaces may have a greater impact on a particular type of ongoing task.

The results from the second study are a strong indication that the reduced FOV provided by many VR systems causes additional cognitive resources to be used during basic locomotion tasks. The nearly equivalent decrease in verbal and spatial memory performance indicates that general attention resources are required. An alternate explanation is that both verbal and spatial resources are required. It is likely that the reduced performance resulted from a switch to more cognitively demanding strategies in the absence of high-fidelity sensory feedback, as would be found in the physical world. Because a typical head-mounted display has an FOV even smaller than that in the study, this finding will be important when considering display technology options.

Together, the results from the first two studies can be used to inform the design of future systems. A more complete understanding of the specific movement detriments associated with different types of concurrent tasks can aid in the selection of an appropriate interface. It may also inform the design of other aspects of the system that have an impact on feedback fidelity, specifically FOV.

For the third contribution, it has been shown that a fuzzy inference system is an appropriate solution to adapt a locomotion interface, enhancing the user's movement performance. Further, it was shown that a user's current verbal and spatial load are sufficient to drive such adaptation and that modifying the dead-zone size can increase performance using the P2V interface. Appropriate error metrics and a mechanism for adjusting fuzzy input sets have also been identified. A formal user study showed that the adaptive system does improve locomotion performance in terms of the identified metrics. The realized fuzzy inference system described here can be extended, but it serves as an excellent proof of concept for the use of fuzzy logic to mitigate dual-task locomotion problems. This adaptation can be used to address problems that cannot be solved with the type of careful interface selection described above.

Future work

The dual-task selective-interference paradigm as implemented was useful in answering the questions posed in this research. However, the memory task specifics can be improved. Tasks in these studies were virtual cards displayed graphically in front of a participant, but this type of task can be administered using other senses as well. VR is well suited for such alternate modalities so tasks should be designed to take advantage of the rich multi-sensory experiences that are possible.

Additionally, the dual-task paradigm should be applied to more locomotion interfaces in order to develop a taxonomy of interfaces in terms of cognitive demands. Such a taxonomy would be very useful when selecting locomotion interfaces for use in specific types of projects and for understanding problems that may arise as a result of these choices.

More research is also needed on the interfaces described in this dissertation. For example, it is interesting that ducking was negatively impacted by a concurrent spatial task in Study 1 even when the interface allowed ducking just as in the physical world. It is not surprising that "real-world" activities are impacted by competition for resources. Using a cell phone while driving is a prominent example. However, the basic locomotion tasks that users performed in these studies should have been very natural and did not require any high-level understanding of the scenario.

The studies described here did not explore high-level navigation activities such as path integration or wayfinding, but those tasks may also compete with locomotion interfaces for cognitive resources. Therefore, an experiment should be carried out involving a capture-the-flag task using a very unnatural locomotion interface. This will require the user to construct a mental model of the VE in order to return to the beginning after finding the flag. It is likely that interface performance will suffer as the user's mental model of the scene grows.

More investigation is needed into aspects other than FOV that may impact cognitive resource usage. Locomotion with a reduced FOV requires additional resources, which may be due to reduced optic flow information, but additional studies are needed to verify this. Once verified, perhaps it can be shown that low resolution has a similar impact on cognitive demands. The same ideas can also be applied to other sensory modalities. The cognitive resource usage resulting from reductions in auditory feedback, for example, may cause similar problems to those identified in this paper. These other modalities should be investigated in additional studies similar to those described here.

The results from the first two experiments provide strong confirmation of the importance of individual differences in terms of cognitive strategies during virtual locomotion. While these individual differences were not directly included in the research questions posed, it is

impossible to proceed with this research without making further attempts to isolate differences in locomotion strategies between users with diverse demographics and abilities.

The adaptive system described in this dissertation relied solely on naive task-load metrics (memory span). This was sufficient to increase user performance and such a light-weight solution may be ideal in some settings. However, more advanced metrics will be required in many domains of interest. Such metrics may include physiological measures which have the potential to provide much more information about a user's current task demands, but care must be taken to mitigate the weaknesses of these technologies.

In addition to improving the adaptive system's understanding of a user's concurrent task load, other aspects of the interface should be added to the system as output variables. For example, control-display gain is a possible candidate for adaptation, though it may hinder the user's ability to learn the interface. Also it may be possible to help with ducking, possibly by adding a "ducking gain" to the system, though it may be difficult to detect when the user intends to duck. There are probably many other variables that can be identified through future dual-task studies that may benefit from adaptation. VirtuTrace in general and the Fuzzy Navigation Engine specifically are easily configurable and the existing components can be easily extended to include more input variables, output variables, or rules in the adaptation process.

Appendix. Study Documents

Participants in all three studies completed the following documents.

Pre-questionnaire. This contained questions about basic demographic information, including video game and athletic experiences.

Post-questionnaire. This contained questions about the participant's experiences while completing the tasks. It included Likert scales to assess perceived performance and written questions regarding problems using the interface and suggestions for improvements. In some cases, the answers on this questionnaire steered the questions asked during the unstructured exit interview.

Pre-questionnaire

Locomotion Study: Pre Questionnaire Page 1

Pre Questionnaire

Instructions: Please answer the following questions about your past experiences

1. **Experience**

a. What video games do you currently play?

b. How many hours per week do you play first-person (for example: first-person shooter) video games? _____

c. Do you have any experience using Virtual Reality systems?

Yes No

i. If yes, please explain. _____

d. What sports do you play?

e. How many hours per week do you spend working out? _____

2. **Demographics**

a. What is your age? _____

b. What is your sex? (circle one) Male Female

c. What is your height? _____

Participant_____ Group_____ Date_____

3. **Education**

a. What is the highest educational level you have completed? (circle one)

High School B.S. M.S. Ph.D. Other:_____

b. If you are a student, what is your major? _____

4. If you are not a student, what is your profession? _____

Participant_____ Group_____ Date_____

Post-questionnaire

Locomotion Study: Post Questionnaire Page 1

Post Questionnaire

Instructions: Please answer the following questions pertaining to your experiences in the virtual environment while playing the game.

1. **Immersion**

a. How immersed did you feel in the virtual environment?

Not Very		*Immersion Rating*				Very
Lowest Immersion						Highest Immersion
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐

2. **Overall Performance**

a. Compared to an average person, how would you rate your performance at the movement tasks?

Not Very		*Performance Rating*				Very
Lowest Performance						Highest Performance
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐

b. Compared to an average person, how would you rate your speed at adapting to the environment?

Slow		*Performance Rating*				Fast
Lowest Performance						Highest Performance
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐

3. What problems did you encounter while attempting to complete the movement tasks?

4. What strategies did you use to adapt to the environment?

Participant_____ Group_____ Date_____

5. What problems did you encounter while attempting to complete the cognitive (memory) tasks?

6. What improvements can be made to enhance navigation performance in the virtual world?

Participant_____ Group_____ Date_____

Bibliography

Ackerman, P. L. (1988, September). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, *117*(3), 288–318. DOI: 10.1037/0096-3445.117.3.288

Alfano, P. L., & Michel, G. F. (1990, February). Restricting the field of view: Perceptual and performance effects. *Perceptual and Motor Skills*, *70*(1), 35–45.

Anderson, J. R. (1982, July). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406.

Baddeley, A. D. (1992, January). Is working memory working? The fifteenth Bartlett lecture. *The Quarterly Journal of Experimental Psychology*, *44A*(1), 1–31. DOI: 10.1080/14640749208401281

Baddeley, A. D. (2002, June). Is working memory still working? *European Psychologist*, *7*(2), 85–97. DOI: 10.1027//1016-9040.7.2.85

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York, NY, USA: Academic Press.

Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2005). *3D user interfaces: Theory and practice* (1st ed.). Boston, MA, USA: Addison-Wesley Professional.

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. M. (1992, July). Verbal recoding of visual stimuli impairs mental image transformations. *Memory & Cognition*, *20*(4), 449–455.

Chance, S. S., Gaunet, F., Beall, A. C., & Loomis, J. M. (1998, April). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence: Teleoperators and Virtual Environments*, *7*(2), 168–178. DOI: 10.1162/105474698565659

Chittaro, L., & Ranon, R. (2007, July/August). Adaptive hypermedia techniques for 3D educational virtual environments. *IEEE Intelligent Systems*, *22*(4), 31–37. DOI: 10.1109/MIS.2007.63

Christensen, R. R., Hollerbach, J. M., Xu, Y., & Meek, S. G. (2000, February). Inertial-force feedback for the treadport locomotion interface. *Presence: Teleoperators and Virtual Environments*, *9*(1), 1–14. DOI: 10.1162/105474600566574

Cirio, G., Marchal, M., Regia-Corte, T., & Lécuyer, A. (2009, November 18–20). The magic barrier tape: A novel metaphor for infinite navigation in virtual worlds with a restricted walking workspace. In *Proceedings of the 16th ACM symposium on virtual reality software and technology (VRST 2009), Kyoto, Japan* (pp. 155–162). New York, NY, USA: ACM. DOI: 10.1145/1643928.1643965

Cox, E. (1992, October). Fuzzy fundamentals. *IEEE Spectrum*, *29*(10), 58–61. DOI: 10.1109/6.158640

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

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992, June). The CAVE: Audio visual experience automatic virtual environment. *Communications of the ACM*, *35*(6), 64–72. DOI: 10.1145/129888.129892

Darken, R. P., Cockayne, W. R., & Carmein, D. (1997, October 14–17). The omni-directional treadmill: A locomotion device for virtual worlds. In *Proceedings of 10th annual symposium on user interface software and technology (UIST 1997), Banff, Canada* (pp. 213–221). New York, NY, USA: ACM. DOI: 10.1145/263407.263550

Darling, S., Sala, S. D., & Logie, R. H. (2009, March). Dissociation between appearance and location within visuo-spatial working memory. *The Quarterly Journal of Experimental Psychology*, *62*(3), 417–425. DOI: 10.1080/17470210802321984

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999, September). Pattern span: A tool for unwelding visuo-spatial memory. *Neuropsychologia*, *37*(10), 1189–1199. DOI: 10.1016/S0028-3932(98)00159-6

Dolezal, H. (1982). *Living in a world transformed: Perceptual and performatory adaptation to visual distortion.* New York, NY, USA: Academic Press.

Engel, D., Curio, C., Tcheang, L., Mohler, B., & Bulthoff, H. H. (2008, October 27–29). A psychophysically calibrated controller for navigating through large environments in a limited free-walking space. In S. K. Feiner, D. Thalmann, P. Guitton, B. Fröhlich, E. Kruijff, & M. Hachet (Eds.), *Proceedings of the 15th ACM symposium on virtual reality software and technology (VRST 2008), Bordeaux, France* (pp. 157–164). New York, NY, USA: ACM. DOI: 10.1145/1450579.1450612

Gibson, J. J. (1958, August). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology*, *49*(3), 182–194. DOI: 10.1111/j.2044-8295.1958.tb00656.x

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 1–49). New York, NY, USA: John Wiley & Sons, Inc.

Gray, R. (2004, March). Attending to the execution of a complex sensorimotor skill: Expertise differences, choking, and slumps. *Journal of Experimental Psychology: Applied*, *10*(1), 42–54.

Grimes, D., Tan, D. S., Hudson, S. E., Shenoy, P., & Rao, R. P. (2008, April 5–10). Feasibility and pragmatics of classifying working memory load with an electroencephalograph. In *Proceedings of the 26th international conference on human factors in computing systems (CHI 2008), Florence, Italy* (pp. 835–844). New York, NY, USA: ACM. DOI: 10.1145/1357054.1357187

Hagras, H., Callaghan, V., Colley, M., Clarke, G., Pounds-Cornish, A., & Duman, H. (2004, November/December). Creating an ambient-intelligence environment using embedded agents. *IEEE Intelligent Systems*, *19*(6), 12–20. DOI: 10.1109/MIS.2004.61

Hayashi, Y. (1992, March 8–12). A neural expert system using fuzzy teaching input. In *IEEE international conference on fuzzy systems, San Diego, CA, USA* (pp. 485–491). IEEE. DOI: 10.1109/FUZZY.1992.258661

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175–191. DOI: 10.1016/j.intell.2003.12.001

Hirshfield, L. M., Solovey, E. T., Girouard, A., Kebinger, J., Jacob, R. J. K., Sassaroli, A., & Fantini, S. (2009, April 4–9). Brain measurement for usability testing and adaptive interfaces: An example of uncovering syntactic workload with functional near infrared spectroscopy. In *Proceedings of the 27th international conference on human factors in computing systems (CHI 2009), Boston, MA, USA* (pp. 2185–2194). New York, NY, USA: ACM. DOI: 10.1145/1518701.1519035

Ittelson, W. H. (1973). Environment perception and contemporary perceptual theory. In W. H. Ittelson (Ed.), *Environment and cognition* (pp. 1–19). Oxford, England: Seminar Press.

Iwata, H. (1999, March 13–17). Walking about virtual environments on an infinite floor. In *Proceedings of IEEE virtual reality, Houston, TX, USA* (pp. 286–293). Washington DC, USA: IEEE Computer Society. DOI: 10.1109/VR.1999.756964

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998, July). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, *9*(4), 293–298. DOI: 10.1111/1467-9280.00058

Kline, P. B., & Witmer, B. G. (1996, September 2–6). Distance perception in virtual environments: Effects of field of view and surface texture at near distances. In *Proceedings of human factors and ergonomics society 40th annual meeting, Philadelphia, PA, USA* (Vol. 2, pp. 1112–1116). Santa Monica, CA, USA: HFES. DOI: 10.1177/154193129604002201

Kosko, B. (1994, November). Fuzzy systems as universal approximators. *IEEE Transactions on*

Computers, *43*(11), 1329–1333. DOI: 10.1109/12.324566

Kozhevnikov, M., Motes, M. A., & Hegarty, M. (2007, July/August). Spatial visualization in physics problem solving. *Cognitive Science*, *31*(4), 549–579. DOI: 10.1080/15326900701399897

Kulik, A. (2009, November/December). Building on realism and magic for designing 3D interaction techniques. *IEEE Computer Graphics and Applications*, *29*(6), 22–33. DOI: 10.1109/MCG.2009.115

Lin, C., & Lee, C. (1992, March 8–12). Real-time supervised structure/parameter learning for fuzzy neural network. In *IEEE international conference on fuzzy systems, San Diego, CA, USA* (pp. 1283–1290). DOI: 10.1109/FUZZY.1992.258596

Logie, R. H. (2003). Spatial and visual working memory: A mental workspace. *The Psychology of Learning and Motivation*, *42*, 37–78. DOI: 10.1016/S0079-7421(03)01002-8

MacKenzie, I. S. (1995). Input devices and interaction techniques for advanced computing. In W. Barfield & T. A. Furness III (Eds.), *Virtual environments and advanced interface design* (pp. 437–470). New York, NY, USA: Oxford University Press.

Macuga, K. L., Loomis, J. M., Beall, A. C., & Kelly, J. W. (2006, July). Perception of heading without retinal optic flow. *Perception & Psychophysics*, *68*(5), 872–878.

McCreary, F. A., & Williges, R. C. (1998, October 5–9). Effects of age and field-of-view on spatial learning in an immersive virtual environment. In *Proceedings of human factors and ergonomics society 42nd annual meeting, Chicago, IL, USA* (pp. 1491–1495). Santa Monica, CA, USA: HFES. DOI: 10.1177/154193129804202106

Medina, E., Fruland, R., & Weghorst, S. (2008, September 22-26). Virtusphere: Walking in a human size VR "hamster ball". In *Proceedings of the human factors and ergonomics society 52nd annual meeting, New York, NY, USA* (pp. 2102–2106). Santa Monica, CA, USA: HFES. DOI: 10.1177/154193120805202704

Mohler, B. J., Thompson, W. B., Creem-Regehr, S. H., Pick, H., Jr., & Warren, W. H., Jr. (2007, August). Visual flow influences gait transition speed and preferred walking speed. *Experimental Brain Research*, *181*(2), 221–228. DOI: 10.1007/s00221-007-0917-0

Montello, D. (1998, May 14–16). Thinking of scale; the scale of thought [Summary Report]. In D. R. Montello & R. G. Golledge (Eds.), *Scale and detail in the cognition of geographic information, Santa Barbara, CA, USA* (pp. 11–12).

Nieuwenhuizen, K., Martens, J.-B., Liu, L., & van Liere, R. (2009, November/December). Insights from dividing 3D goal-directed movements into meaningful phases. *IEEE Computer Graphics and Applications*, *29*(6), 44–53. DOI: 10.1109/MCG.2009.121

Openscenegraph. (2011). Retrieved February 23, 2012, from `http://www.openscenegraph.org/`

Péruch, P., May, M., & Wartenberg, F. (1997). Homing in virtual environments: Effects of field of view and path layout. *Perception*, *26*(3), 301–311.

Rada-Vilela, J. (2011). *Fuzzy-lite: A fuzzy logic library written in C++*. Retrieved February 23, 2012, from `http://code.google.com/p/fuzzy-lite/`

Razzaque, S., Kohn, Z., & Whitton, M. C. (2001, September). Redirected walking. In *Proceedings of Eurographics, Manchester, UK*.

Ruddle, R. A., & Jones, D. M. (2001, October). Movement in cluttered virtual environments. *Presence: Teleoperators and Virtual Environments*, *10*(5), 511–524. DOI: 10.1162/105474601753132687

Schwartz, T. J. (1992, June). IEEE gets fuzzy. *IEEE Expert*, 72–73.

Shiffrin, R. M., & Schneider, W. (1977, March). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*(2), 127–190. DOI: 10.1037/0033-295X.84.2.127

Sholl, J. (1998, May 14–16). Are there different psychological scale classes? [Summary Report]. In D. R. Montello & R. G. Golledge (Eds.), *Scale and detail in the cognition of geographic information, Santa Barbara, CA, USA* (p. 16).

Sibert, L. E., Templeman, J. N., Stripling, R. M., Coyne, J. T., Page, R. C., La Budde, Z., & Afergan, D. (2008, September 22-26). Comparison of locomotion interfaces for immersive training systems. In *Proceedings of the human factors and ergonomics society 52nd annual meeting, New York, NY, USA* (pp. 2097–2101). Santa Monica, CA, USA: HFES. DOI: 10.1177/154193120805202703

Simon, H. A. (1983). Why should machines learn. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (pp. 25–37). Palo Alto, CA, USA: Tioga.

Slater, M., Steed, A., McCarthy, J., & Maringelli, F. (1998, September). The influence of body movement on subjective presence in virtual environments. *Human Factors*, *40*(3), 469–477. DOI: 10.1518/001872098779591368

Slater, M., Usoh, M., & Steed, A. (1994, August 23–26). Steps and ladders in virtual reality. In G. Singh, S. K. Feiner, & D. Thalmann (Eds.), *Proceedings of the conference on virtual reality software and technology (VRST 1994), singapore* (pp. 45–54). River Edge, NJ, USA: World Scientific.

Suma, E. A., Finkelstein, S. L., Clark, S., Goolkasian, P., & Hodges, L. F. (2010, March 20–21). Effects of travel technique and gender on a divided attention task in a virtual environment. In *Proceedings of the IEEE symposium on 3D user interfaces (3DUI 2010), Waltham, MA, USA* (pp. 27–34). Washington DC, USA: IEEE Computer Society. DOI: 10.1109/3DUI.2010.5444726

Templeman, J. N., Denbrook, P. S., & Sibert, L. E. (1999, December). Virtual locomotion: Walking in place through virtual environments. *Presence: Teleoperators and Virtual Environments*, *8*(6), 598–617. DOI: 10.1162/105474699566512

Templeman, J. N., & Sibert, L. E. (2006, September). Immersive simulation of coordinated motion in virtual environments: An application to training small unit military tactics, techniques, and procedures. In G. L. Allen (Ed.), *Applied spatial cognition: From research to cognitive technology* (pp. 339–372). London, UK: Psychology Press.

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004, October). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments*, *13*(5), 560–571. DOI: 10.1162/1054746042545292

Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., & Brooks, F. P., Jr. (1999, August 8–13). Walking > walking-in-place > flying, in virtual environments. In *Proceedings of the 26th annual conference on computer graphics and interactive techniques (SIGGRAPH 1999)* (Vol. 33, pp. 359–364). Los Angeles, CA, USA: ACM. DOI: 10.1145/311535.311589

Vainio, A.-M., Valtonen, M., & Vanhala, J. (2008, March/April). Proactive fuzzy control and adaptation methods for smart homes. *IEEE Intelligent Systems*, *23*(2), 42–49. DOI: 10.1109/MIS.2008.33

Vergauwe, E., Barrouillet, P., & Camos, V. (2009, July). Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1012–1028. DOI: 10.1037/a0015859

Wandell, B. A. (1995). *Foundations of vision.* Sunderland, MA, USA: Sinauer Associates.

Wang, Z., Bauernfeind, K., & Sugar, T. (2003, March 22–23). Omni-directional treadmill system. In *Proceedings of the 11th symposium on haptic interfaces for virtual environment and teleoperator systems (HAPTICS 2003), Los Angeles, CA, USA* (pp. 367–373). Washington DC, USA: IEEE Computer Society. DOI: 10.1109/HAPTIC.2003.1191314

Warren, W. H., Jr., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001, February). Optic flow is used to control human walking. *Nature Neuroscience*, *4*(2), 213–216.

Wells, M., Peterson, B., & Aten, J. (1996, March 30–April 3). The virtual motion controller: A sufficient-motion walking simulator. In *Proceedings of the IEEE virtual reality annual international symposium (VRAIS 1996), Santa Clara, CA, USA* (pp. 1–8). Los Alamitos, CA, USA: IEEE Computer Society.

Wickens, C. D., & Baker, P. (1995). Cognitive issues in virtual reality. In W. Barfield & T. A. Furness III (Eds.), *Virtual environments and advanced interface design* (pp. 514–541).

New York, NY, USA: Oxford University Press.

Willemsen, P., Colton, M. B., Creem-Regehr, S. H., & Thompson, W. B. (2009, February). 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), 8:1–8:14. DOI: 10.1145/1498700.1498702

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Review*, *monograph supplement 3*, 1–114.

Yun, C., Shastri, D., Pavlidis, I., & Deng, Z. (2009, April 4–9). O' game, can you feel my frustration?: Improving user's gaming experience via stresscam. In *Proceedings of the 27th international conference on human factors in computing systems (CHI 2009), Boston, MA, USA* (pp. 2195–2204). New York, NY, USA: ACM. DOI: 10.1145/1518701.1519036

Zanbaka, C., Lok, B., Babu, S., Ulinski, A., & Hodges, L. F. (2005, November/December). Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics*, *11*(6), 694–705. DOI: 10.1109/TVCG.2005.92