MACBETH:
Management of Avatar Conflict By Employment of a Technique Hybrid

by
Eric Burns

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill
2006

Approved by:

Advisor: Frederick P. Brooks, Jr.

Reader: Mary C. Whitton

Reader: Abigail T. Panter

Leonard McMillan

Bernard D. Adelstein

Dennis R. Proffitt

ABSTRACT

Eric Burns

MACBETH: Management of Avatar Conflict By Employment of a Technique Hybrid

(Under the direction of Frederick P. Brooks, Jr.)

Since virtual objects do not prevent users from penetrating them, a virtual environment user may place his real hand inside a virtual object. If the virtual environment system prevents the user's hand avatar from penetrating the object, the hand avatar must be placed somewhere other than the user's real hand position. I propose a technique, named MACBETH (Management of Avatar Conflict By Employment of a Technique Hybrid) for managing the position of a user's hand avatar in a natural manner after it has been separated from the user's real hand due to collision with a virtual object. This technique balances visual/proprioceptive discrepancy in position and velocity by choosing each so that they are equally detectable.

To gather the necessary information to implement MACBETH, I performed user studies to determine users' detection thresholds for visual/proprioceptive discrepancy in hand position and velocity. I then ran a user study to evaluate MACBETH against two other techniques for managing the hand avatar position: the rubber-band and incremental-motion techniques. Users rated MACBETH as more natural than the other techniques and preferred MACBETH over both. Users performed better on a hand navigation task with MACBETH than with the incremental-motion technique and performed equally well as with the rubber-band technique.

To my parents, George and Mary Burns,
who have always been my model of unconditional love

ACKNOWLEDGMENTS

I feel richly blessed to have had the opportunity to spend four and a half years at the University of North Carolina to work on this research. Though I know none of them personally, I want to thank the people at the Office of Naval Research, the NIH National Institute for Biomedical Imaging and Bioengineering, and the NIH National Institute for Research Resources for providing funding, without which I would never have had this opportunity.

In a very special way I would like to thank my doctoral committee members. Fred Brooks and Mary Whitton accepted me into their research group when I was a new and inexperienced graduate student and had little to offer them. They encouraged me when I was discouraged, showed understanding during my failures, and supported me through all I encountered. They have taught me so much about what it means to be an honest, dedicated researcher and a generous, caring mentor. I desire nothing more than to live up to their example.

I feel deeply grateful to the contributions of my other committee members, as well. Abigail Panter, Leonard McMillan, and Dennis Proffitt all donated their time to lend me their unique perspectives and expertise. Bernard Adelstein welcomed me to his laboratory at the NASA Ames Research Center for a summer to work in his laboratory. The work I did for him pales in comparison to all that he taught me about user study design. He has continually offered his help whenever I hit a bump in the road and has become a good friend.

My committee helped me greatly, but I wouldn't be where I am now if, when I first arrived at UNC, Sharif Razzaque hadn't invited me to lunch with him to discuss research ideas. The research presented in this dissertation is a result of that conversation. Sharif is the type of person who delights in helping everyone around him, and I cannot express how thankful I am that I was one of the people that he was delighted to help.

I also owe a multitude of thanks to Paul Zimmons. During my first two years at UNC Paul was there every time I ran into a dead-end in my research. He spent long hours helping me find my way out of those dead-ends, and I wonder if I didn't postpone his graduation date because of all the time he spent helping me. I remember several occasions when I sent a panicked e-mail to him in the middle of the night because something was going wrong with my research, only to have him show up in person at the laboratory within 15 minutes. At the beginning of Paul's dissertation defense, Mary Whitton commented on how much he had helped all the other members of our research group. Paul responded that when he first began graduate school many people had helped him, and he just tried to return the favor. I cannot imagine that he ever received as much help as he gave me, or that I could ever hope to help others to the degree he helped me. For his help I am very thankful, and I will never forget how he kept us laughing always.

I would also like to thank Christopher Wiesen at the Odum Institute for his help with statistics. My work would have been much more difficult without him.

I want to thank Luv Kohli, primarily for being such a good friend during my entire stay at Chapel Hill, but also for helping me work through numerous logical errors and programming bugs, even when he had his own very pressing work to attend to. I also want to thank Luv for never being afraid to tell me in what areas, professional or otherwise, he thought I could improve.

I want to thank Christopher Oates for his help with 3D modeling and Christopher VanderKnyff for his technical help in the many forms he provided it. I want to thank Tabitha Peck for being my constant cheerleader as we shared an office during the final year of this research. And I want to thank the entire Effective Virtual Environments research team for being my friends and for providing a place for me to learn.

I also want to thank the UNC Computer Science Department Support Staff, especially David Harrison, for coming to my aid during numerous crises and helping me avoid disaster.

I want to thank my best friend, Andrew, for proofreading this dissertation even though it was outside his field of interest and for flying 400 miles to surprise me at my dissertation defense.

I want to thank my girlfriend, Erin, for giving me a reason to want to finish this research, for encouraging me through its final months, and for acting as my constant support, without which I would still be working.

I cannot even begin to thank my parents for all they have done for me to make this all possible. They have been my advisors, counselors, and creditors, but most importantly, they have been my friends. And I wish I could somehow express to them just how grateful I am, but words will always fail me in that respect.

And beyond all else, I am thankful to God for the gifts he has given me, imperfect steward of them, though I am; for all the people he has put in my life, both friends and enemies; and for loving me in a way I can only begin to fathom.

CONTENTS

List of Tables

List of Figures

Chapter 1: **Introduction**

> Is this a dagger which I see before me,
> The handle toward my hand? Come, let me clutch thee.
> I have thee not, and yet I see thee still.
> Art thou not, fatal vision, sensible
> To feeling as to sight? or art thou but
> A dagger of the mind, a false creation,
> Proceeding from the heat-oppresséd brain?
> I see thee yet, in form as palpable
> As this which now I draw.
> Thou marshall'st me the way that I was going;
> And such an instrument I was to use.
> Mine eyes are made the fools o' the other senses,
> Or else worth all the rest. . . . [Shakespeare, *Macbeth* 2.1.33-45]

In this soliloquy, Macbeth is tormented by a vision of a dagger floating in front of him. Reaching for it, he ends up with nothing but a fistful of air. Most users of virtual environments (VEs) can sympathize with Macbeth: most large VEs do not offer any haptic feedback, and users are left reaching for objects that they can see but cannot touch.

When a head-mounted display user with a hand avatar (a graphical object representing the tracked position of the real hand in the VE) reaches out for a virtual object and encounters this "dagger-of-the-mind" problem, unless some special provision is made he sees his hand avatar penetrate the virtual object (Figure 1-1). Lindeman, Sibert, and Templeman found that this penetration makes it difficult for users to perform precise tasks, and that user performance improved when visual interpenetration was prevented by using simulated surface constraints [Lindeman, Sibert, & Templeman, 2001].

However, preventing visual interpenetration requires that the user's hand avatar sometimes appear somewhere other than where the user's real hand is (Figure 1-2). Preventing the interpenetration thus creates a discrepancy between the user's sensory cues from vison and proprioception – the internal sense of body position and motion. For virtual environments in which a psychological state of presence is desirable, the choice to prevent or not prevent avatar interpenetrations with virtual objects should be made based on whether the visual interpenetration or the visual/proprioceptive discrepancy is less likely to be noticed by the user.

Figure 1-1. Interpenetration problem: A user may see his hand avatar penetrate a virtual object when the object he is reaching for does not exist in the real world.

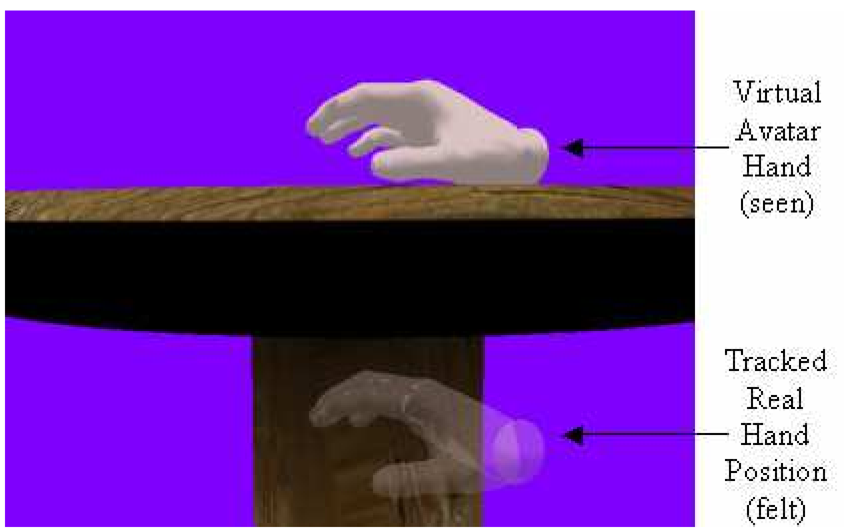

Figure 1-2. Sensory discrepancy problem: Preventing visual interpenetration requires that the user's hand avatar sometimes appear somewhere other than where the user's real hand feels according to the proprioceptive sense.

1.1 Thesis statement: Part I

> Users are more likely to notice visual penetration of virtual objects by the hand avatar than the discrepancy in visual and proprioceptive hand-position cues introduced by preventing such penetration.

Psychologists have studied intersensory discrepancy for decades. J. Gibson [1933] found that when vision and proprioception disagree, participants tend to perceive their hand position to be where vision tells them it is, a phenomenon called *visual dominance* or *visual capture*. Many researchers have explored visual dominance and other aspects of sensory integration. Welch [1986] compiled an excellent survey of the literature prior to 1986. Van Beers, Sittig, and Denier van der Gon [1999] and van Beers, Wolpert, and Haggard [2002] are notable examples of research since then. These studies dealt primarily with the perception of hand position under sensory discrepancy and not whether participants detected the discrepancy itself. The first part of the thesis statement concerns the latter. I performed a study to test this thesis and found it to be true (Chapter 2).

1.2 The question raised by preventing visual interpenetrations

When visual interpenetrations are prevented, the user's hand avatar can no longer appear where the user's hand is. If a user's avatar hand cannot be placed at the location of his real hand at every simulation time step, then where should it be placed?

Two commonly-used approaches to managing the avatar position after collision with a virtual object are the rubber-band method and the incremental-motion method [Zachmann and Rettig, 2001]. The *rubber-band method* minimizes the position discrepancy between the real and virtual hands at every simulation time step, as if they were connected by a rubber band. However, the result is velocity discrepancy, as the virtual hand sometimes sticks to surfaces when the real hand is moving (Figure 1-3), and when the hand avatar slides off the edge of a virtual object, it sometimes pops to the position of the real hand when the real hand is not moving.

The *incremental-motion method* faithfully preserves the motion of the real hand. For each increment of movement the real hand makes, the virtual hand is moved the same amount. However, the result is a position discrepancy that may grow unboundedly if the hand avatar repeatedly collides with virtual objects (Figure 1-4). The rubber-band method minimizes position discrepancy while disregarding velocity discrepancy; the incremental-motion method minimizes velocity discrepancy while disregarding position discrepancy.

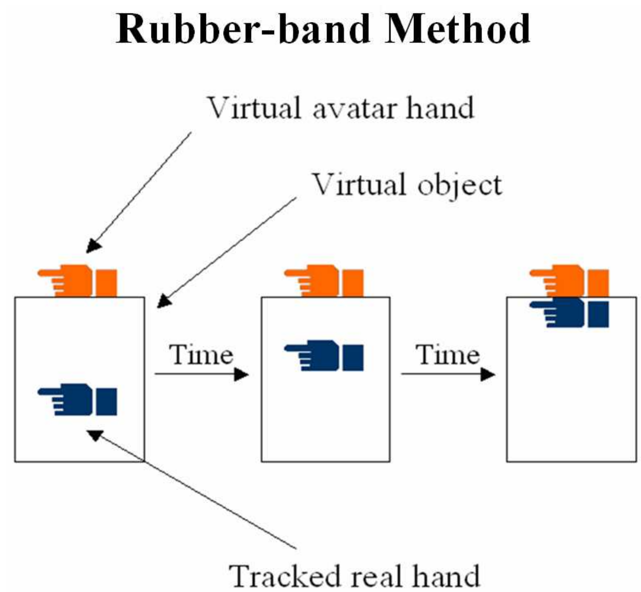

Figure 1-3. Under the rubber-band method, when a user backs his hand out of a virtual object, the hand avatar stays as close as possible to the user's real hand, sticking to the surface while the real-hand is moving, until the penetration is cleared.

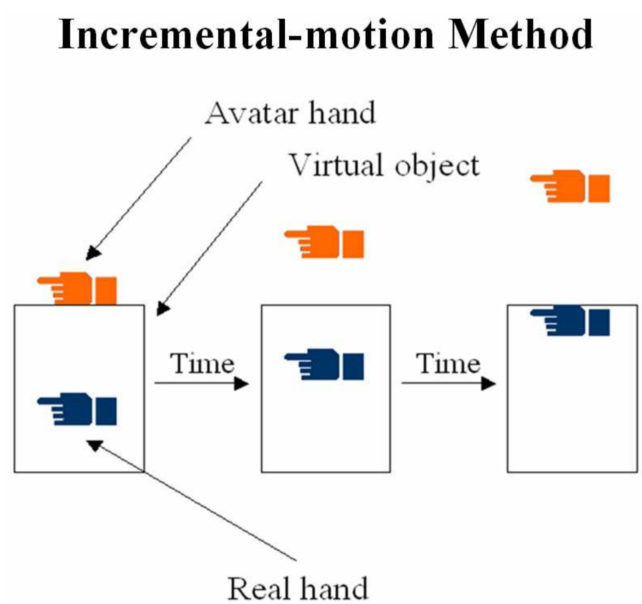

Figure 1-4. Under the incremental-motion method, the hand avatar faithfully preserves the movement of the user's real hand but has no provision to reduce the position discrepancy between the real and avatar hands. With repeated collisions, this position discrepancy can grow unboundedly.

1.3 An idea

I postulate that a method that combines the ideas of the rubber-band and incremental-motion techniques to minimize sensory discrepancy in both hand position and hand velocity will be better than either technique alone. I propose the following: almost preserve the velocity of the real hand, like the incremental-motion method, but introduce some velocity discrepancy to reduce the position discrepancy over time (Figure 1-5). This technique would ensure that:

1) the hand avatar moves when the real hand moves
2) the hand avatar returns to the position of the real hand

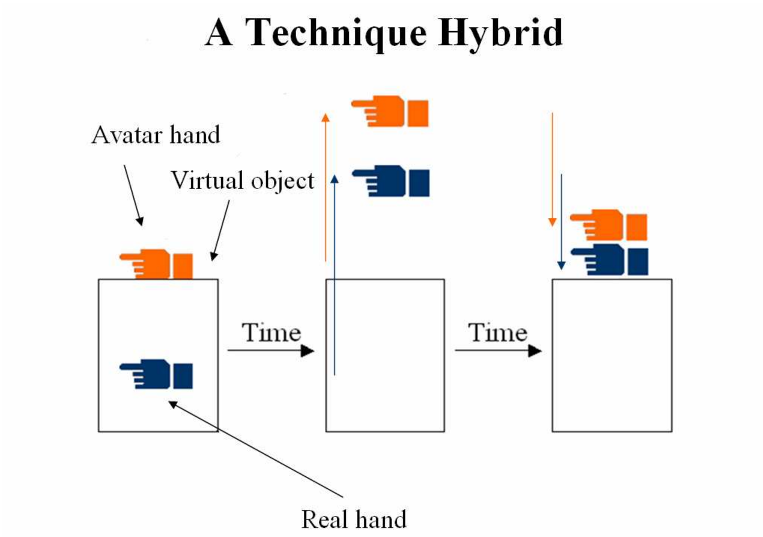

Figure 1-5. Position discrepancy can be reduced by moving the hand avatar slower than the real hand when the user is moving his real hand toward the hand avatar's position (center) and faster when he is moving his real hand away (right).

Others have pursued this area. Colgate, Stanley, and Brown [1995] suggested calculating forces for a haptic device by conceptually connecting virtual objects to their real counterparts by a damped spring. This technique could also be used to bring virtual objects (with an assigned mass) back to the position of a real object. This technique would not remove position discrepancy instantaneously, as does the rubber-band method, but would do so over time. Therefore, velocity discrepancy would be less than under the rubber-band method. Zachmann and Rettig [2001], in the description of the incremental-motion method, actually state, "when the [real object] has moved by a certain delta the [virtual] object will try to move *about* the same delta" (emphasis added). Moving *about* the same amount as the real object can reduce the position discrepancy. The damped-spring model is discussed in Chapter 4.

What is yet to be done is to decide how best to balance the two discrepancies. I propose a method that starts from the incremental-motion technique, in which the user's real and avatar hands have position discrepancy and no velocity discrepancy (other than that created by collisions). Velocity discrepancy of some "proper" amount can then be added to reduce the position discrepancy. I propose that the proper amount be chosen systematically, according to principles:

1) A velocity discrepancy should never be introduced that is more detectable than the existing position discrepancy because its addition would make the overall manipulation more detectable.
2) Position discrepancy should be reduced as quickly as possible. In other words, the largest velocity discrepancy possible should be introduced without violating the first principle.

These principles dictate that the level of the velocity discrepancy introduced should be exactly as detectable as the level of the existing position discrepancy.

Creating such a technique hybrid requires knowing the levels at which users detect position discrepancy and velocity discrepancy. The first study yielded detection thresholds for position discrepancy (Chapter 2). A second study yielded user's velocity discrepancy detection thresholds (Chapter 3). These thresholds were then used to implement the proposed method, called MACBETH (Management of Avatar Conflict By Employment of a Technique Hybrid) (Chapter 4).

1.4 Thesis statement: Part II

> If a user's hand avatar is rejoined to the real hand so that sensory discrepancy in position and velocity are equalized, one or more of the following will result:
>
> - The user will rate the technique as more natural.
> - The user will prefer his virtual environment experience.
> - The user will perform better on tasks in a virtual environment.

A third study tested MACBETH against the rubber-band and incremental-motion methods (Chapter 5). Overall, MACBETH was rated by users as statistically significantly more natural than both the rubber-band and incremental-motion techniques and was statistically significantly preferred to both methods. On a task which I considered an average case, users performed as well with MACBETH as they did with the rubber-band technique, and statistically significantly better than they did with the incremental-motion technique.

Chapter 2: **Study 1 – Sensitivity to Visual Interpenetration vs. Visual-proprioceptive Position Discrepancy**

This chapter is a modified form of an article published in *Presence: Teleoperators and Virtual Environments* [Burns, Razzaque, Panter, Whitton, McCallus, & Brooks, 2006].

Figure 2-1. This participant believes he is aiming at a virtual game board directly in front of him.

2.1 Questions and Hypotheses

This study explored three questions:

1) Are users more sensitive to visual interpenetration or to visual-proprioceptive position discrepancy?
2) When users are *expecting* visual-proprioceptive discrepancy, how much more sensitive are they than when they are *not expecting* it?
3) Do users report that visual interpenetration or visual-proprioceptive position discrepancy is easier to detect?

My hypotheses were:

1) Visual-proprioceptive discrepancy detection thresholds are higher than visual interpenetration detection thresholds; interpenetration is easier to detect.
2) Visual-proprioceptive discrepancy detection thresholds are higher when users are not expecting discrepancy.
3) Users will report that visual interpenetration is easier to detect than visual-proprioceptive position discrepancy.

The study confirmed all three hypotheses with statistical significance of $p \leq 0.05$.

2.2 Study Design

Forty right-handed introductory psychology students (19 males and 21 females) participated in this study. All gave consent and were given class credit for their participation.

The study consisted of three parts. Part I measured reaction time. Part II measured detection thresholds for visual-proprioceptive discrepancy. Part III measured detection thresholds for visual interpenetration. All participants completed Part I first, but the order of Parts II and III were assigned randomly. After the three main parts, users were given an exit questionnaire and then interviewed.

Parts II and III used a partial method-of-limits design to find users' detection thresholds. A complete method-of-limits design consists of an ascending series (starting with no stimulus and increasing it until the user perceives it) and a descending series (starting with a detectable stimulus and decreasing it until the user no longer perceives it). These two series balance each other because ascending series overestimate detection thresholds, and descending series underestimate detection thresholds. However, in a real scenario either stimulus would start from zero when the hand avatar first contacted a virtual object and then grow until it was detected. Since, the goal is to determine how large these stimuli can *grow* before being noticed the ascending-series design is appropriate and does not overestimate the desired threshold.

2.2.1 Part I – Reaction time

A detection threshold is the magnitude of a stimulus at the time of its detection. One can measure the stimulus magnitude only at a user's time of report, one reaction time later:

$$t_{report} = t_{detect} + t_{react}$$

Therefore, I measured participants' reaction times so as to estimate their detection thresholds.

In this part of the study, each participant sat in front of a black computer screen and held a joystick in the right hand. At random intervals the screen turned white, at which point the participant clicked the joystick button as quickly as possible. The interval was recorded as the reaction time. Participants performed this task 45 times.

I assumed that performing this task would not significantly affect users' subsequent performance because:

1) The task was dissimilar from those following, so it was unlikely to produce a significant training effect.
2) The task was short enough (less than five minutes) that it was unlikely to produce fatigue effects.

2.2.2 Part II – Detection threshold for visual-proprioceptive position discrepancy

Part II measured participants' detection thresholds for visual-proprioceptive position discrepancy. Each participant wore a Virtual Research Systems V8 HMD and held a joystick in the right hand. Both the head and hand were tracked using a 3rdTech Hiball 3000. The participant sat in a chair (Figure 2-1) and was visually immersed in a virtual room with four large colored panels on the front wall. The participant's hand avatar held a remote control (Figure 2-2).

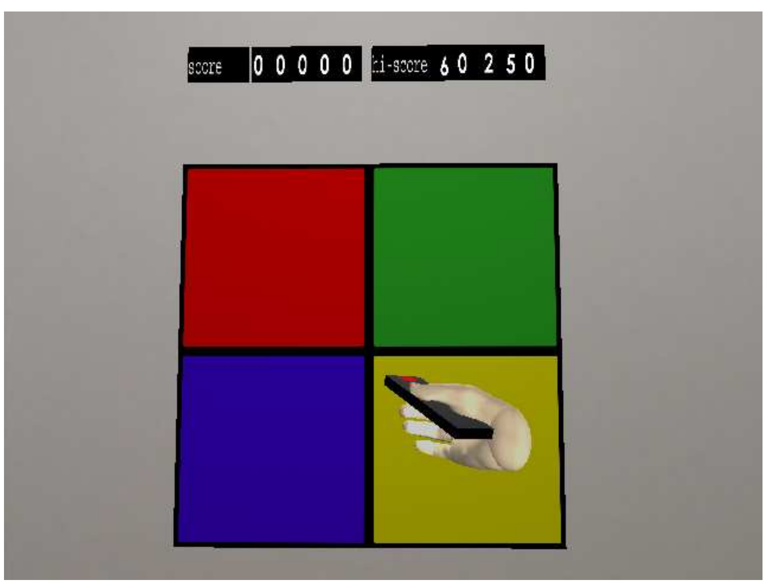

Figure 2-2. The user's view of the virtual room with the Simon game board on the wall – The user's hand avatar, holding a TV-like remote control, is in the foreground.

Participants played a game similar to Hasbro's Simon®. Participants watched as panels lit up successively in a random sequence of length five. Participants then duplicated the sequence by aiming at the appropriate panels in turn and clicking the joystick button. After a participant completed each sequence correctly or made an error, a new sequence began. To keep participants engaged, I scored their performance. The score was displayed on the wall over the colored panels, together with the top score of all participants to date.

Before the game began, I told participants that the study was about perception and performance in a VE and therefore, it was very important for them to report if they noticed anything odd about the VE experience, by holding down the joystick button for five seconds. I then gave three examples of events they would want to report: the game stopping, the computer display having problems, or the virtual hand having drifted away from the real hand.

This part of the study was divided into two sections. In part IIA, participants were not directly primed to expect visual-proprioceptive discrepancy; in part IIB, they were.

2.2.2.1 Part IIA – Unprimed threshold

The Simon® game began, and after a geometrically distributed random interval, averaging 25 seconds, the participant's hand avatar was made to drift from the real hand position. The hand drifted left along a cylinder centered at the participant's estimated shoulder position (a fixed offset from the head tracker) (Figure 2-3).

To investigate position discrepancy, I needed to be certain that participants noticed the *extent* of the drift and not the motion itself. Therefore, I needed to execute the drift such that it was imperceptible.

Pre-study piloting showed that participants could detect even a very slow drift if they held their hands completely still and watched for it. Therefore, during the study, the hand avatar drifted only if the user's hand was moving faster than 5 cm/s. When the user's hand was so moving, the hand avatar drifted 0.46 degrees/s (5 mm/s for someone with a 63.5 cm arm). With these values, none of the pre-pilot participants detected the drift.[1]

The hand avatar drifted until the participant reported noticing the discrepancy or until it reached 60 degrees. If the participant did not report the discrepancy, I asked if he had noticed anything odd. If not, I told him that something odd had happened and asked

[1] This method of gradually increasing the sensory discrepancy is essentially Howard's [1968] method of *discordance shaping*, used to induce perceptual adaptation.

him to guess what it was. If he did not guess correctly, I told him that the hand had drifted and asked again if he had noticed.

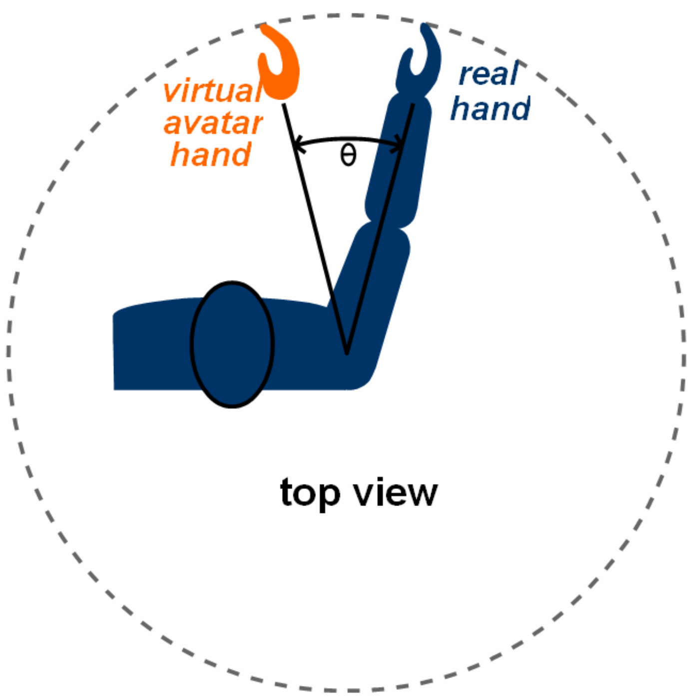

Figure 2-3. The participant's hand avatar drifted left about the shoulder.

2.2.2.2 Part IIB – Primed threshold

When Part IIA ended, I told participants that the rest of Part II was divided into eight trials of the Simon® game, and in each trial the hand avatar would have a 50 percent chance of drifting. In one trial each, the hand avatar drifted left, right, up, and down with respect to the real hand. In the other four trials, the hand did not drift. The order of the drift conditions was selected from an 8x8 balanced Latin square matrix (each order was used five times over the 40 participants). These drift conditions correspond to the position discrepancy that would be introduced when a real hand penetrated a virtual object from its left, right, top, and bottom surfaces, respectively.

I instructed participants to report drift as soon as they noticed it and to report the drift direction. I told them that it was much more important to report the drift immediately than to get the direction correct. I then told them they would be rewarded with bonus Simon® points for correctly identifying drift, regardless of whether they chose the correct direction but would be penalized the same number of points for reporting drift when none occurred. The points were awarded so users would not ignore the drift recognition task in favor of the Simon® game score. The penalty motivated users not to report drift when they did not detect it.

2.2.2.3 Measures

In both IIA and IIB, I measured the maximum angular offset between the virtual and real hands at the time of report, as well as the maximum linear distance between them (for comparison to the visual interpenetration thresholds). I recorded what the participant reported as odd (if anything) in Part A and which direction the user believed the hand drifted after every drift report in Part B. I recorded the mean point score per second in Part A and on each trial of Part B.

2.2.3 Part III – Visual interpenetration detection threshold

Part III measured each participant's visual-interpenetration detection threshold. Participants wore the same HMD and held the same joystick as in Part II. In this part, the user's real hand movement did not control the virtual hand. Instead, when the user clicked the button at the beginning of a trial, the virtual hand moved under simulation control toward a planar virtual object (either a tabletop or a wall). I told participants that in each trial the virtual hand had a 50 percent chance of penetrating the virtual object. They were instructed that if the hand penetrated the object, they must click the button as soon as they noticed. Participants repeated this task 40 times.

The hand speed was varied so participants could not use time alone to judge when the hand would penetrate the object. Penetration and hand-speed orders were selected from independent 40x40 balanced Latin square matrices. I told participants that they were free to look around the room and gather depth cues from the other walls, but I asked them not to move their heads to view the hand from a different angle. If at any point the user's head moved more than 15 cm from its starting position, the user's view went blank and recorded audio instructions asked the user not to move his head position during the task. The user then clicked the joystick button to continue.

Viewing hand penetration from different angles and with different backgrounds affects the difficulty of this task. Detection is easiest from a viewing angle perpendicular to hand motion because the closing gap between the virtual hand and object are directly visible. Conversely, detection is most difficult from a parallel viewing angle because the point of contact is obscured by the hand itself until it becomes extreme, so the user must rely on depth cues to detect the penetration (Figure 2-4).

I originally chose a study condition that I felt represented a commonplace occurrence in VEs, named the vertical-motion condition (Section 2.2.3.1). However, since I hypothesized that sensory discrepancy is harder to detect than visual interpenetration, I feared that my choice of visual penetration condition would be biased toward making penetration detection easy. Hence, I added another, more difficult condition, named the horizontal-motion condition (Section 2.2.3.2). Eighteen participants were randomly assigned to the vertical-motion condition and 21 were assigned to the horizontal-motion condition. One participant's data was accidentally lost.

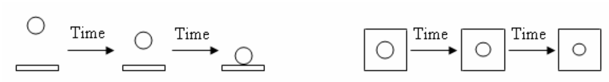

Figure 2-4. Detecting the collision of a ball with the ground is easier when viewed from the side (perpendicular to motion direction), left, than when viewed from above (parallel to motion direction), right.

2.2.3.1 Vertical Motion

In the vertical-motion condition, participants viewed a virtual hand holding a cylinder above a wood-textured tabletop that stood 0.74 meters off the ground (Figure 2-5). The hand was placed based on the height of the user's head so that its point of impact with the table was 45 degrees below the user's horizontal view direction. When the participant clicked the button at the beginning of each trial the virtual hand began moving down toward the tabletop. This condition mimics a common scenario in which a person is seated at a table and places a hand on top of it with arm outstretched.

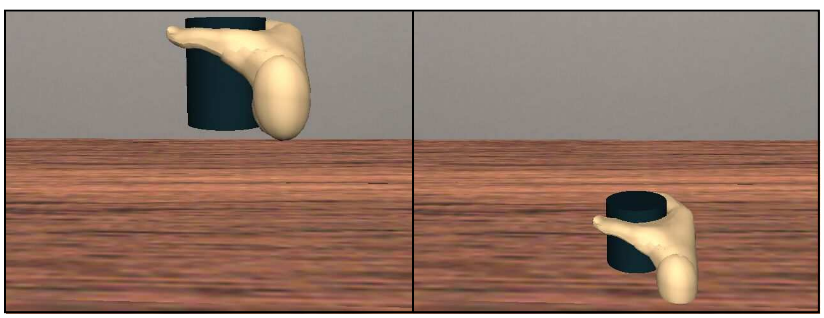

Figure 2-5. The vertical-motion condition: Participants viewed a hand holding a cylinder above a tabletop. Left – the hand's starting position; Right – the hand after penetrating 2 cm.

This condition matches the *up* condition in Part II because each represents a possible outcome of a user moving a hand down through a virtual tabletop. *Without* simulated surface constraints, the virtual hand penetrates the tabletop, as in this condition. *With* the constraints, the virtual hand stays on top of the table, creating a position discrepancy in the up direction with respect to the real hand, as in the *up* condition in Part II.

2.2.3.2 Horizontal Motion

In the horizontal-motion condition, participants viewed the virtual hand 20 cm in front of a wall that was approximately 40 cm from the viewer. The wall was featureless so as to offer minimal depth cues (Figure 2-6). When participants clicked the button to start each trial the hand began moving toward the wall.

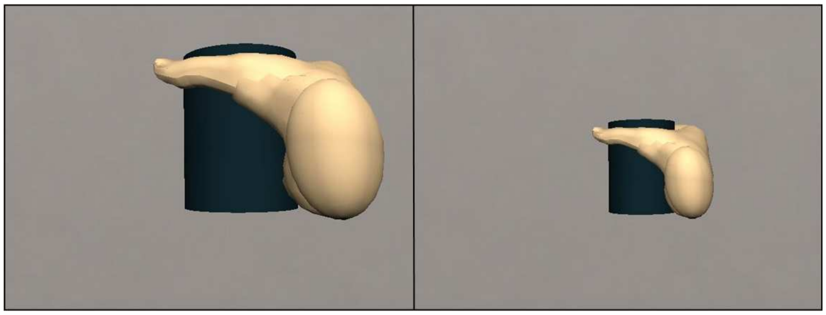

Figure 2-6. The horizontal-motion condition: Participants viewed a hand holding a cylinder in front of a wall. Left – the hand's starting position; Right – the hand after penetrating 2 cm.

2.2.3.3 Measures

In each condition, I recorded the hand penetration depth at the time of the user report.

2.3 Results and Analysis

The 40 participants yielded 19 sets of complete data. I lost six sensory discrepancy values due to software malfunctions, 16 due to false alarms on trials in which the hand would have drifted (when the participant reported drift before it began), and 16 because time ran out before completion of the experiment.

2.3.1 A note about statistical analysis

With all t-tests and ANOVAs in this research, I assume normality of the population distributions. This assumption is weak, meaning the results of t-test and ANOVAs are robust if the assumption fails to hold. In most tests, I also assume equivalence of variance of the two populations being sampled. The equivalence of variance assumption is stronger, meaning the results of the tests depend more heavily on the assumption. Whenever I have reason to doubt equivalence of variance, I use test variants that do not use a pooled variance for the two sample populations. However, it is important to keep in mind that the strongest assumption I make when using these tests is that the underlying model is additive, meaning that the value of the outcome variable is determined by a linear combination of the independent variables. This assumption is made by all who use these tests, and no method exists to test it.

2.3.2 Simplifying analysis by combining data across drift directions

To simplify data analysis, I wished to treat the sensory discrepancy thresholds for the four drift directions as four different measurements of the same threshold. First, I tested the thresholds for each direction for statistically significant differences.

A repeated-measures ANOVA failed to find a significant difference among the four drift directions for the 19 participants with complete data ($F_{3, 54} = .80$, $p > .49$). However, this analysis ignores the possibility that participants with missing data vary systematically with respect to participants with complete data. Since the participants most likely to have missing data are at the two extremes of performance – under-responders, who took a long time to report and ran out of time before completing the experiment; and over-responders, who reported drift before it actually occurred – I cannot claim that participants with missing data do not vary systematically with respect to participants with complete data.

To include the effect of participants with missing data, I used a Markov chain Monte Carlo multiple imputation method [Yuan, 2000] to generate 30 complete datasets using the mean and covariance structure of the observed data. I did not have a method to combine the results of 30 repeated-measures ANOVAs, so I performed the simpler two-tailed t-test on the six individual direction pairs for each dataset and combined the results to produce the statistics shown in Table 2-1. These pairwise t-tests are more susceptible to type I error (finding a statistically significant difference when none exists) for individual large differences than a repeated-measures ANOVA. However, none of the pairs produced a statistically significant difference, so no large differences are likely to exist between any of the drift direction pairs. The inability to use the repeated-measures ANOVA sacrificed its added power to find small differences across all drift directions, but if these differences exist, they are small.

Table 2-1. Results of the two-tailed t-test for each direction pair on the multiply-imputed data set of sensory discrepancy thresholds.

Direction pair	***Magnitude of position discrepancy difference (m)***	t_{39}	p
left / right	.043	1.92	.063
left / up	.026	1.05	.30
left / down	.013	.34	.74
right / up	.017	-1.03	.31
right / down	.030	-.803	.43
up / down	.013	-.29	.77

Neither the test on the complete datasets nor the test on the imputed datasets showed statistically significant differences, but not finding a difference does not automatically imply that one does not exist, especially since each test has an issue that calls its credibility into question:

- The test on complete datasets excludes participants whose data may vary systematically.
- The test on multiply-imputed datasets requires substituting values for a high percentage of missing data (20 percent).

Therefore, I cannot conclude that no difference exists between the drift directions. However, the lack of statistically significant differences is evidence that differences among drift directions are small enough that I may combine the four thresholds into a mean discrepancy threshold for each user.

2.3.3 Detection threshold comparison

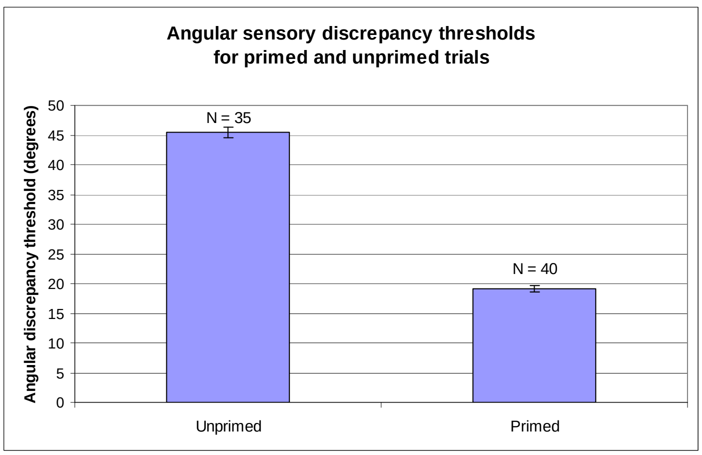

Figure 2-7. Mean angular visual-proprioceptive discrepancy thresholds – Bars represent a 95 percent confidence interval for the mean.

Figure 2-7 shows the mean angular unprimed and primed discrepancy thresholds from Part II. Figure 2-8 shows the mean linear unprimed and primed discrepancy thresholds from Part II alongside the mean visual interpenetration thresholds from Part III. These values represent the estimated stimulus levels at the time of detection, calculated from report times and reaction times (mean reaction time = 260 ms, standard deviation = 20 ms) as follows:

$$thresh_{detect} = pos_{report} - t_{react} * v_{hand}$$

where $thresh_{detect}$ is the detection threshold, pos_{report} is the position discrepancy or penetration depth at the time of report, t_{react} is the user's reaction time, and v_{hand} is the hand speed. I discarded false alarms prior to calculating the mean detection thresholds.

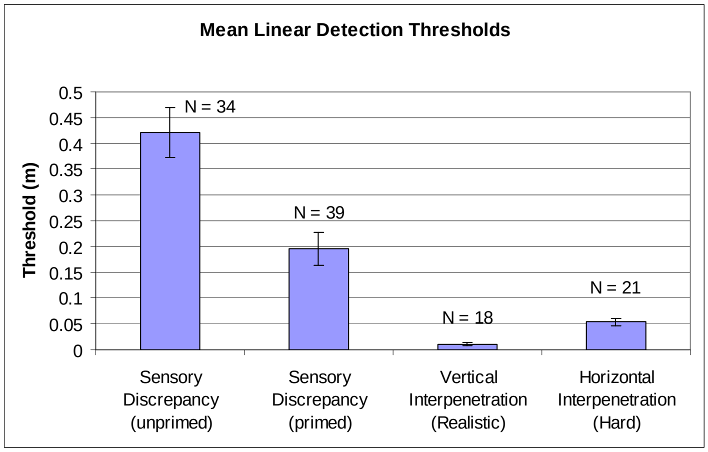

Figure 2-8. Mean detection thresholds for visual-proprioceptive discrepancy and visual interpenetration – Bars represent a 95 percent confidence interval for the mean.

Because the visual-proprioceptive discrepancy and visual interpenetration detection thresholds may have different variances, I analyzed them using MANOVA. The analysis showed a significant difference between primed sensory-discrepancy thresholds and visual-interpenetration thresholds for both the vertical-motion condition ($F_{1,\ 17} = 61.74$, $p < .001$) and the horizontal-motion condition ($F_{1,\ 20} = 322.23$, $p < .001$). Comparing the unprimed sensory discrepancy thresholds against the visual interpenetration thresholds (when users were primed to expect visual interpenetration) is not meaningful because of the different user expectations.

The sensory-discrepancy thresholds were higher than the visual interpenetration thresholds even though they were underestimated for two reasons. First, I assumed that the hand avatar was moving throughout the duration of the participant's reaction time. If the user held his hand still or removed it from his field of view, the hand would not have moved during this time, and the reaction distance value subtracted from the discrepancy would be too large, resulting in a reported threshold that is too small. Second, the mean detection threshold ignores the false alarm rate of the participants. Figure 2-9 shows mean detection thresholds as a function of the number of false alarms reported by the participant. A linear regression of mean detection threshold on number of false alarms yielded a statistically significant downward trend (intercept = 0.227m, slope = -0.0217m, $F_{1,\ 31} = 8.68$, $p < .006$), meaning that the participants with the lowest thresholds had the most false alarms. Their low detection thresholds suggest that they performed the task well. However, their high false alarm rates reveal that these participants were not consistently able to discriminate sensory discrepancy from its absence. Therefore, their low thresholds are misleading.

Often researchers ascertain the discriminability of the stimulus by analyzing receiver-operator characteristics [Heeger, 2003]. However, receiver-operator characteristic analysis requires a study design that allows users to miss a stimulus by reporting that it does not exist when it does [Coren, Ward, & Enns, 1999]. Since the *method of limits* design used in this study increases the stimulus level until the stimulus is detected, it is impossible for a participant to miss a stimulus. I instead used the data from all participants without regard to false alarm rates to estimate a mean detection threshold. The resulting estimate is conservative because the data from participants with high false alarm rates artificially lowers the mean.

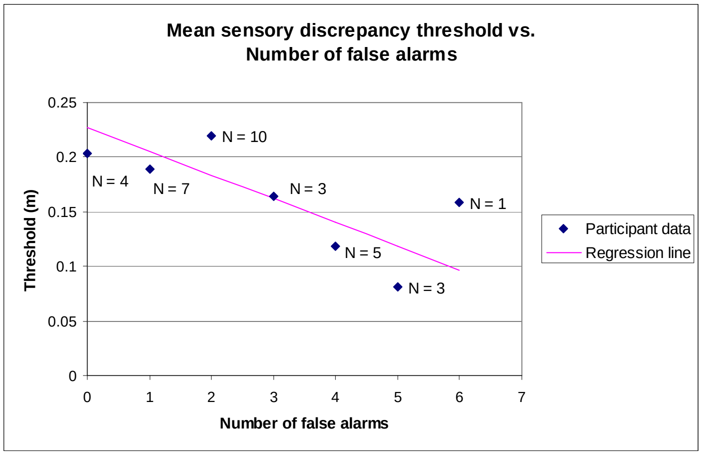

Figure 2-9. Mean unprimed sensory discrepancy thresholds as a function of the participant's number of false alarms – N values represent the number of participants with the given number of false alarms.

2.3.4 Sensory discrepancy detection threshold comparison with respect to priming

Figure 2-10 shows a top-down view of the unprimed and primed mean visual-proprioceptive discrepancy thresholds. A correlation test showed that these two thresholds are mildly correlated, with $r = .352$, $p < .042$. A repeated-measures t-test showed the unprimed thresholds to be statistically significantly higher than the primed thresholds with $t_{33} = 9.008$, $p < .001$. However, the systematic underestimation of the primed detection thresholds, indicated by the high false alarm rate (Figure 2-9), calls this result into question. I cannot assume that unprimed detection thresholds are subject to the same underestimation, because a linear regression of unprimed threshold on false alarm rate failed to find the same trend as that found with primed threshold.

However, the mean unprimed detection threshold is underestimated for a different reason. Seventeen participants did not report an odd event on the unprimed trial. Instead, the trial ended when the hand avatar reached a 60-degree offset from the real hand. These participants were then asked if anything was odd about their experience:

- Five immediately mentioned the hand drift, though they had not reported it. These participants were not included in the statistics for the unprimed trial, because their lack of reporting was likely due to a misunderstanding of the instructions rather than a lack of detecting the sensory discrepancy.
- Eight could not guess what was odd about the experience when told that I had introduced a manipulation. However, when asked if they noticed that the virtual hand had drifted, they said they *did* notice. One of these participants volunteered his understanding of where his real hand was in relation to his avatar hand, but did so incorrectly.
- Four said they did not notice at all that the hand avatar had drifted.

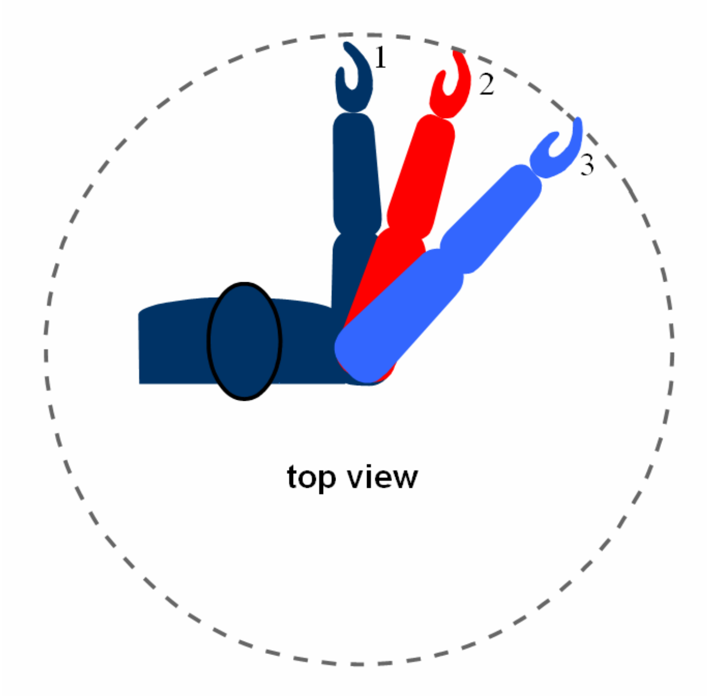

Figure 2-10. An overhead view of hand placements corresponding to the mean thresholds in Figure 2-8: 1) Hand avatar position 2) Mean threshold in primed trials (19.1°) 3) Mean threshold in unprimed trial (45.4°)

In addition to the 12 participants who never reported an odd event on the unprimed trial (not including the five whose data was discarded), eight participants reported some other odd occurrence before they noticed the hand had drifted. Therefore, 20 out of 34 participants (only 34 instead of 40 because, in addition to discarding the five nonresponders who had noticed drift, I lost one unprimed trial due to an equipment malfunction) yielded values that represented lower bounds on their real detection thresholds. I can only be sure that the reported value for 14 out of 34 participants represents an actual detection threshold. Therefore, the resulting unprimed threshold estimate is conservative.

2.3.5 User report of task difficulty

On an exit questionnaire, users rated the difficulty of detecting sensory discrepancy and visual interpenetration on a scale of 1 to 7. Because these data fall into discrete categories which have an inherent order, the parametric ordered multinomial regression test is appropriate. The regression of user report of difficulty on the type of threshold (sensory discrepancy or visual interpenetration) showed that participants rated the task of detecting hand drift significantly harder than that of detecting visual interpenetration with $\chi^2_{1,\,40} = 62.7$, $p < .001$ (Figure 2-11).

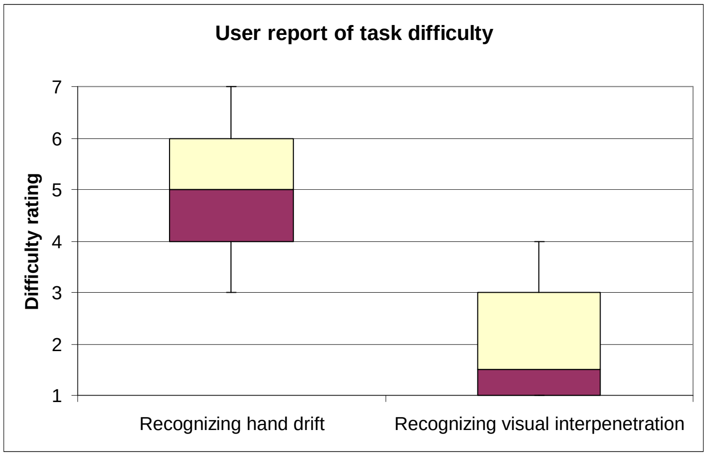

Figure 2-11. User report of task difficulty on a scale of 1 to 7 (1 – easiest; 7 – hardest) – the bottom of each box represents the 25th percentile mark, the mid-line is the median, and the top of the box represents the 75th percentile. Error bars represent the minimum and maximum responses.

2.3.6 Performance effects of visual-proprioceptive discrepancy

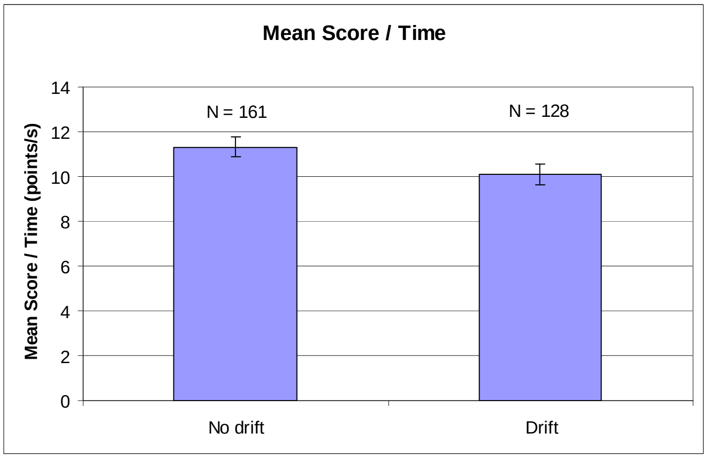

Figure 2-12. Mean score per second on trials in which the hand did or did not drift – Bars represent a 95 percent confidence interval for the mean.

For each trial of the Simon game, I calculated the participant's mean score per second. A repeated-measures t-test showed significantly poorer performance on trials during which the hand drifted than on those during which it did not, with $t_{39} = 3.18$, $p < .003$ (Figure 2-12). This led me to question if visual-proprioceptive discrepancy affected a user's perceived hand position such that performance on a manual task would suffer. I returned to this question in Study 3 (Chapter 5).

2.4 Discussion

The results of Study 1 support my hypotheses: visual-proprioceptive discrepancy thresholds were statistically significantly higher than visual interpenetration thresholds, visual-proprioceptive discrepancy thresholds were statistically significantly higher when users were not expecting it (although, as discussed, this result is not beyond question), and users reported that detecting visual interpenetration was statistically significantly easier than detecting visual-proprioceptive discrepancy.

Lindeman, Sibert, & Templeman [2001] found that simulated surface constraints improve users' speed and accuracy on manual tasks and that users prefer simulated surface constraints to their absence. This study has added to these results by finding that users are less likely to notice the position discrepancy resulting from simulated surface constraints than the visual interpenetration that would otherwise occur.

Chapter 3: **Study 2 – User Sensitivity to Visual/proprioceptive Discrepancy in Hand Velocity**

This chapter is a modified form of a paper presented at the ACM Symposium on Virtual Reality Software and Technology in November 2006 [Burns & Brooks, 2006].

3.1 Question

What is the detection threshold for velocity discrepancy (difference in velocity between the *viewed* virtual hand and the *felt* real hand – I refer to this vector as the *discrepancy vector*)?

3.2 Study Design

3.2.1 Participants

Thirty-three introductory psychology students participated in this study. All gave consent and were given class credit for their participation. Three participants developed symptoms of simulator sickness soon after beginning the study and were excused, leaving 17 males and 13 females.

3.2.2 Equipment

Each participant wore a Virtual Research Systems V8 head-mounted display and held a joystick in the right hand. Both head and hand were tracked using a 3rdTech Hiball 3000. Participants sat in a chair and were visually immersed in a one-room VE that measured 4.6 m by 2.3 m with a 2.7 m by 1.9 m alcove behind them (Figure 3-1).

3.2.3 Stimulus

3.2.3.1 The magnitude of the discrepancy vector

Weber's Law states that the magnitude of the smallest distinguishable difference between two stimuli, or difference detection threshold, is directly proportional to the magnitude of the base stimulus [Fechner, 1966]. In equation form:

$$\Delta I = k * I$$

Where I is the intensity of the base stimulus, ΔI is the difference detection threshold, and k is some constant relating the two.

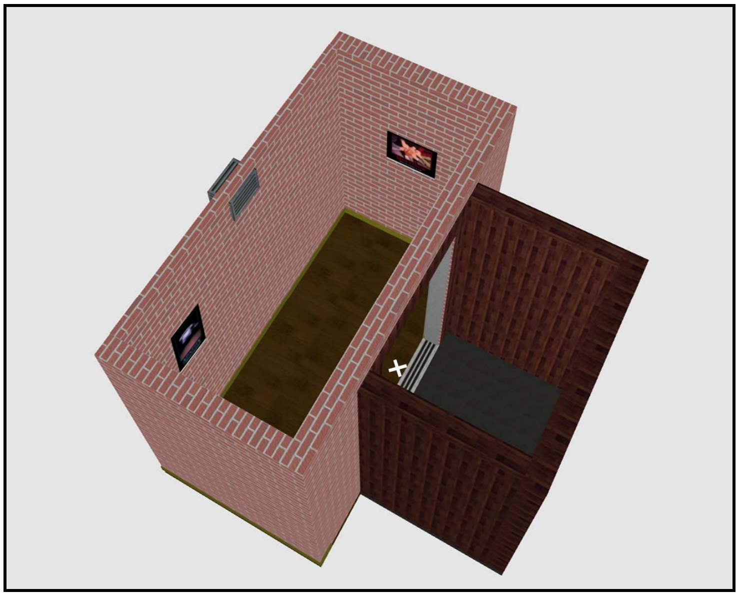

Figure 3-1. View of the VE from above. The white x shows where the user sat, facing the long brick wall.

Weber's Law was empirically developed for *difference* detection thresholds (*difference thresholds*, for short). Although this study concerns *discrepancy* detection thresholds (*discrepancy thresholds*, for short), they are very similar to difference thresholds. A *difference threshold* is the magnitude that one stimulus must differ from another in the *same* modality for a person to be able to distinguish them [Coren, Ward, & Enns, 1999]. A *discrepancy threshold* is the magnitude that a stimulus in one modality must differ from a stimulus in a *different* modality for a person to be able to distinguish them.

Applying Weber's Law, by analogy, to discrepancy thresholds, I have made the simplifying assumption that the velocity discrepancy threshold will be a constant multiple of the base stimulus (in this case the real-hand velocity):

$$\Delta v = k * v_{real}$$

For simplicity, I measure the factor k because, unlike the absolute velocity discrepancy threshold (Δv), k is invariant to changes in the user's real-hand velocity (v_{real}). Therefore, in each trial, the stimulus level is a potential value for k, and the hand-avatar velocity is set to:

$$v_{avatar} = v_{real} + \Delta v = v_{real} + k * v_{real} = (1 + k)v_{real}$$

where k is positive for faster conditions and negative for slower conditions.

I later tested the assumption that the discrepancy threshold is constant with respect to real-hand speed (Section 3.3.3).

3.2.3.2 The orientation of the discrepancy vector

Since the stimulus in this study is a vector, it can vary not only in magnitude but in direction as well. The vector's orientation may affect the discrepancy threshold, as it affects which sensory receptors get excited and how. It is therefore necessary to specify a frame of reference and then to deal with the potential variation of discrepancy threshold with respect to orientation.

3.2.3.2.1 Frame of reference

The frame of reference for the orientation of the discrepancy vector is somewhat complicated, as two coordinate systems are involved:

1) Vision: Based on the position and orientation of the eyes
2) Proprioception: Based on the position and orientation of the muscles transmitting the sensations of motion

I chose the visual frame of reference as the base frame and used the real-hand velocity vector to represent the influence of the proprioceptive frame of reference.

The visual frame of reference has its origin between the user's eyes and is most naturally described in spherical coordinates with θ corresponding to the horizontal placement on the retina, φ corresponding to the vertical placement on the retina, and r corresponding to the distance from the origin (Figure 3-2).

3.2.3.2.2 Studying the potential variation of discrepancy threshold with respect to orientation

The discrepancy threshold cannot be measured for each of the infinite possible discrepancy vector directions. However, if I assume the detection threshold in an arbitrary direction is a linear combination of the detection thresholds in its three spatial component directions, then detection thresholds need only be measured in the component directions.

In reality the assumption of linearity does not hold. When users perform arbitrarily complex movements, the added complexity decreases the accuracy of their proprioceptive feedback. Therefore, this assumption will likely yield conservative velocity discrepancy detection thresholds. In the implementation of MACBETH, such an underestimation means that the introduced velocity discrepancies will be less than those which would have been possible, and the hand avatar will not return to the user's real hand as quickly as it could have. The assumption of linearity allows a tractable solution to the problem of choosing a velocity discrepancy that matches a given position discrepancy. Though the resulting implementation of MACBETH might not remove position discrepancy as quickly as possible, it certainly will not make the situation worse.

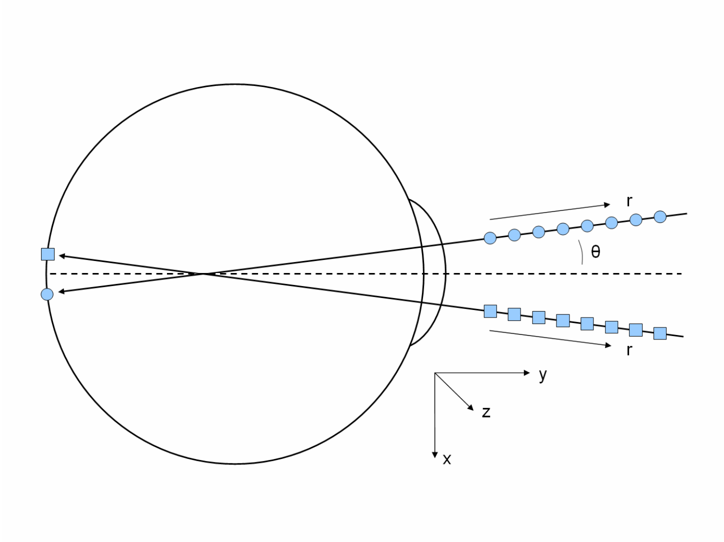

Figure 3-2. The eye viewed from above. Though the individual objects on each of the straight lines have different x, y, and z coordinates in a Cartesian coordinate system, they have the same θ and φ values in spherical coordinates, and their images land in the same position on the retina.

3.2.4 Conditions

Pilot studies showed that movements to the left across the visual field did not necessarily have the same discrepancy detection thresholds as movements right across the visual field (likewise for up and down, and toward and away). This is not surprising, since the muscles are used in different ways to perform each motion. Therefore, for each directional component I measured the detection threshold for hand motion in both the positive and negative directions. For each of these, I measured one detection threshold for when the hand avatar moved more quickly than the real hand and one for when it moved more slowly. This yielded 12 conditions (3 directional components x 2 real-hand motion directions x 2 faster/slower conditions).

3.2.5 The execution of each condition

3.2.5.1 The trial

Participants underwent a series of trials, each of which yielded a single binary data point of whether or not the participant detected discrepancy for a given stimulus level in a given condition.

At the beginning of a trial, a panel on the virtual wall in front of the user indicated which direction the user was to move his hand. The participant clicked the button on the joystick, and a sphere appeared indicating where the user should move his hand to start the trial (Figure 3-3). When the participant moved his real hand to the apparent location of the sphere, his hand avatar disappeared, and his hand movement controlled the movement of the sphere. The sphere's velocity was set as follows:

$$v_{avatar} = (1 + stimulus)v_{real}$$

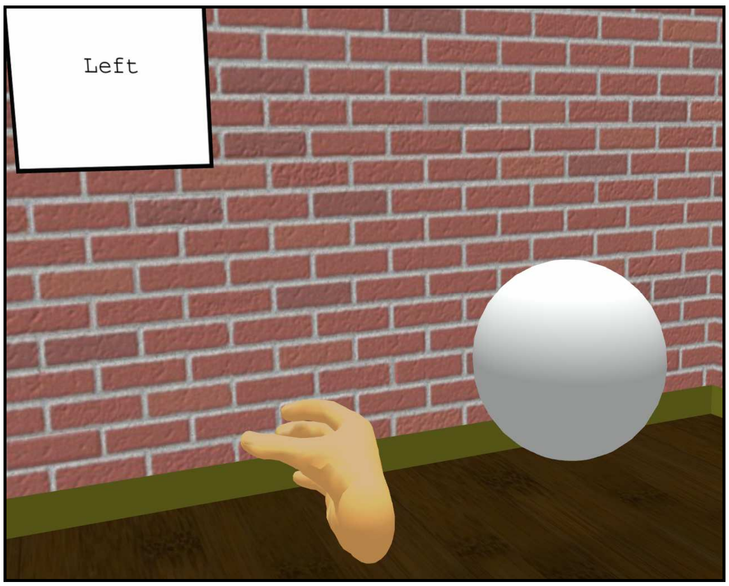

Figure 3-3. At the beginning of each trial, a sphere indicated where the user should move his real hand to start the trial. A panel on the wall indicated the direction the user was to move his hand during the trial.

The user then moved his hand in the direction specified by the panel until the sphere disappeared at an invisible goal position which varied randomly with the trial. The user's mean real-hand speed was recorded from when the sphere was intersected to when the goal was reached. Upon reaching the goal, the sphere disappeared and the panel on the wall changed to a response menu with three panels, allowing the participant to choose whether the movement of the sphere appeared to be *faster*, the *same* speed, or *slower* than the real hand. The user selected a panel with a laser pointer controlled by his gaze direction. When the laser pointer dot passed over a panel, the panel would light up (Figure 3-4). The user made the final selection by clicking the button on the joystick. When the user clicked the button, his hand avatar appeared again with a new sphere to indicate the starting position of the hand and a new panel to specify the direction of motion for the next trial.

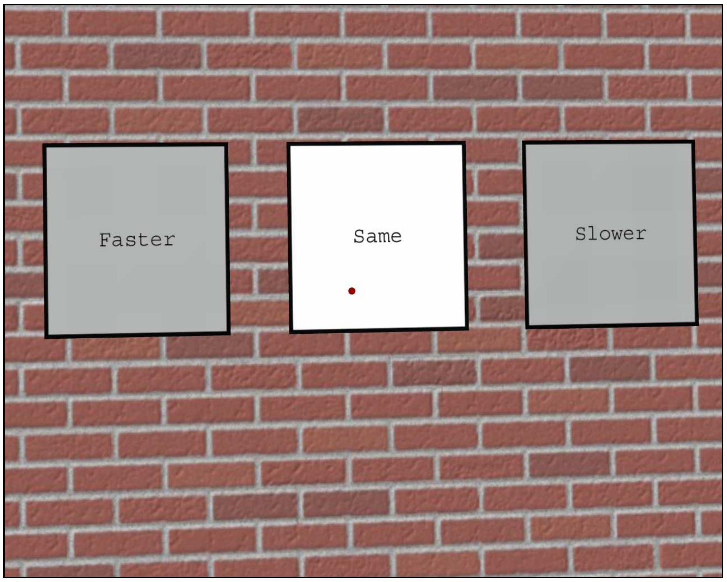

Figure 3-4. At the end of the trial, the user selected whether the movement of the hand avatar appeared faster, slower, or the same speed as the real hand.

3.2.5.2 Velocity discrepancy detection vs. position discrepancy detection

I was concerned that instead of comparing the velocity of the real hand to that of the virtual hand, participants might notice the accumulating position discrepancy between the real and virtual hands. If I assume that participants pay attention to velocity discrepancy (because I have asked them to), position discrepancy is only an issue if it is *more* detectable than velocity discrepancy. If it is *less* detectable, I can be sure that in any trial in which the participants noticed position discrepancy, they would have also noticed velocity discrepancy, so their reports are correct for the purpose of this study. However, if position discrepancy is *more* detectable, there might be times when participants did *not* notice velocity discrepancy, but reported that the sphere moved faster or slower because they noticed the position discrepancy. The results of Study 1 suggest that the position discrepancy is likely not a concern because its detection threshold is very large.

However, as a precaution I undertook to make position discrepancy harder to recognize. These efforts focused on the end of the hand motion, since:

1) At the end of the hand motion, the hand avatar is farthest from the real hand.
2) I feel that users are most cognizant of position discrepancy at the end of the hand motion because their attention shifts from the motion of the hand (which has stopped) to the position of the hand.

To make position discrepancy harder to recognize at the end of the hand motion, the invisible goal position of the hand varied from trial to trial, so that users would not know where the end of the hand motion would be. They instead had to move the sphere until it disappeared. Because they required reaction time to stop their hands, the final resting position of the hand was not directly comparable to the final visible position of the virtual sphere. Therefore, position discrepancy was not directly assessable.

However, despite my efforts at throwing users off from detecting position discrepancy and the fact that humans are bad at detecting it anyway, I cannot be certain that participants are not, in fact, noticing the position discrepancy rather than the velocity discrepancy. In a worst-case scenario, however, if position discrepancy is what they notice, I know that their detection thresholds to velocity discrepancy are higher than the current discrepancy, so my measurements are a conservative estimate of their actual detection thresholds.

3.2.5.3 The adaptive staircase

For each trial, the stimulus magnitude – the potential value of k that determined the velocity discrepancy – was selected according to a 1-up, 1-down adaptive staircase method. Staircase designs focus the majority of trials in the stimulus region of most interest (around the areas where participants sometimes answer one way, but sometimes answer another). *Adaptive* staircases refine the step size to help the staircase converge to a detection threshold faster.

The first trial had either a small discrepancy magnitude (the bottom of the staircase, in this case, 0.0) or a large discrepancy magnitude, chosen based on the results of a pilot study (1.0 for faster conditions and -0.6 for slower conditions). The next trial's stimulus level was increased or decreased by one step of the staircase, depending on whether the participant reported the discrepancy correctly or incorrectly (for faster conditions, a step up was in the positive direction; for slower conditions it was in the negative direction). The stimulus level would not advance beyond the extremes of the range (-1.0 to 1.0). The beginning step size was 0.2. Each time the participant responded the opposite of the previous trial, the step size was halved until a minimum step size was reached. The minimum step size was 0.1. Each staircase continued until the participant had made 10 reversals or had completed 50 trials.

3.2.5.4 Groups of staircases

The goal of all the trials was to find, for each participant, the detection threshold for each of the 12 conditions. The detection threshold was found by fitting a Gaussian ogive to the participant's detection rate at every stimulus level (the percentage of the presentations of that stimulus level that the participant detected) by minimizing the weighted sum of square differences of the data values to the ogive fit values divided by the ogive fit values at every point. This minimization was accomplished by varying the mean and variance of the ogive. This method minimizes the chi-square of the Gaussian ogive fit to create an estimate of the participant's *psychometric function* (the cumulative distribution function of a user's probability of detecting the stimulus – Figure 3-5 is a sample). From the psychometric function I extracted an *absolute detection threshold* (the stimulus level at which the participant had 50 percent accuracy, also known as the *point of subjective equality* or PSE).

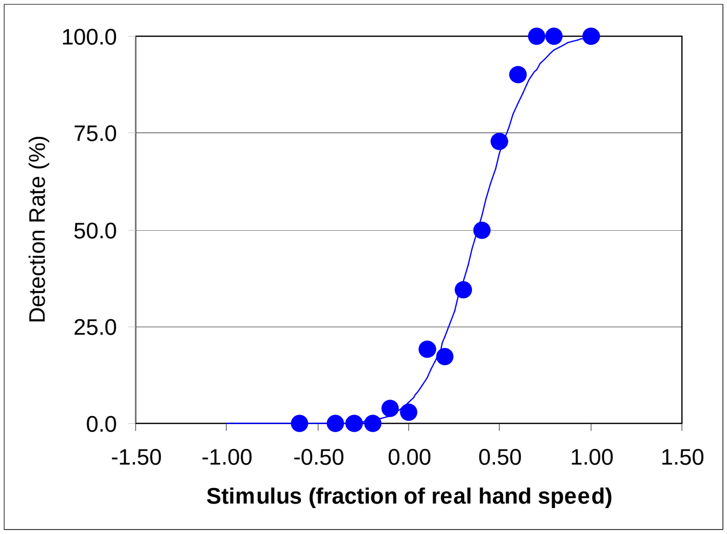

Figure 3-5. A sample psychometric function fit to a user's data points for the left/faster condition

The number of trials needed to create a good psychometric function can be achieved by:

1) One long staircase that requires many reversals before ending
2) Several shorter staircases that require fewer reversals to end

Choosing several shorter staircases has two advantages:

1) After the first few trials, participants may recognize the staircase nature of the presentation of the stimuli. Several staircases may be randomly interleaved so as to make it difficult for a participant to determine where he is on the staircase.
2) Though I wish to concentrate the data in the center of the participant's psychometric function, it is desirable to have more than one data point at the extremes. Each staircase guarantees data at the starting point, which is a high or low extreme.

I concurrently ran six staircases for each condition, three starting low and three starting high, to ensure three data points at each of the extremes.

3.2.6 Participant groups

Participants were randomly placed into three groups. The participants in each group experienced 4 of the 12 conditions. The four conditions were chosen so the participant would have two opposite hand motions (left and right, up and down, or toward and away) with a pair of faster and slower conditions for each. The staircase for each trial was chosen randomly with the requirement that its real-hand motion be in the direction opposite to the last hand motion. This requirement was added because participants have a tendency to compare the hand-avatar velocity to the previous hand-avatar velocity, rather than to the velocity of the real-hand. By making the real-hand motion the opposite of the previous trial, it was more difficult for users to make this mistake.

3.2.7 Data

I used the 50% detection threshold for each of the participants that experienced a condition to construct a confidence interval for the population's mean detection threshold. I used the hand-speed measurements to test the assumption that the detection threshold follows Weber's Law.

3.3 Results

3.3.1 Psychometric functions

I created 120 psychometric functions, one for each participant (N = 30) for each of four conditions (Figure 3-5 is an example). One participant's data in the toward/slower condition was erratic to an extent that the correlation coefficient of the data with the subsequent ogive fit was not statistically significant. Thus, the estimated psychometric function did not yield a dependable detection threshold. That participant's data for that condition was discarded. Eight of the remaining 119 sets of data yielded psychometric functions that had a detection threshold greater than 1.0 (Figure 3-6 is an example). Since the greatest stimulus for which data was collected was 1.0, the detection thresholds for these datasets lay outside the region of collected data and were extrapolated from data that comprised less than half of the psychometric function's region of most interest. For this reason, these values are at high risk of containing large amounts of error. I decided it would be safer to replace these detection thresholds with the value of 1.0, recognizing that this represents a lower bound on the real detection threshold. Therefore, my reported detection thresholds are conservative.

3.3.2 Mean detection thresholds

Mean 50% detection thresholds for all 12 conditions are shown in Figure 3-7.

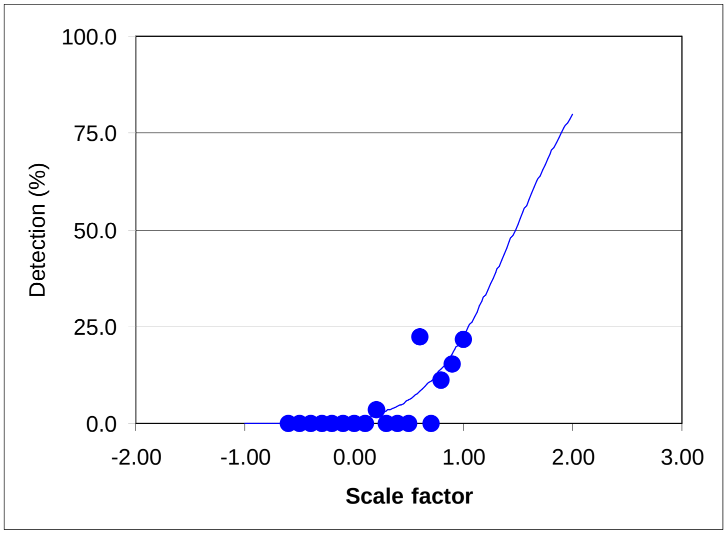

Figure 3-6. An example psychometric function from the up/faster condition with a 50% detection threshold higher than 1.0.

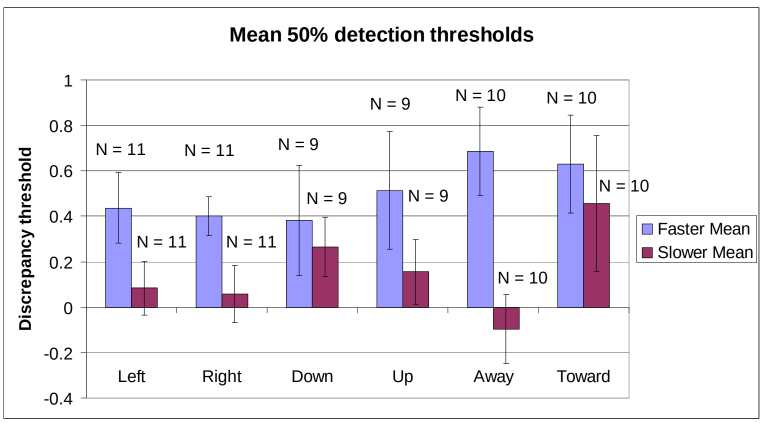

Figure 3-7. Mean 50% detection thresholds for visual/proprioceptive discrepancy. Bars represent 95% confidence intervals for the mean.

3.3.3 Testing the assumption that the detection threshold follows Weber's Law

If the discrepancy detection threshold follows Weber's Law, the threshold should be a constant fraction of the user's real-hand velocity. In other words, the slope of the function relating threshold to real hand speed should be 0. To test whether the slope is indeed 0, for every trial, I measured the user's mean hand speed from the beginning of the hand motion to the end (the *trial hand speed*). I divided each user's set of trials for a condition into the half whose trial hand speeds were above the median trial hand speed, and the half whose trial hand speeds were below it. If Weber's Law holds, these two half-sets should indicate the same discrepancy threshold (since I measured it as a fraction of the base stimulus). I constructed a psychometric function for both of these half-sets, plotted their detection thresholds against their mean trial hand speeds, and found the slope of the line connecting the two points. Figure 3-8 shows a histogram of these slopes.

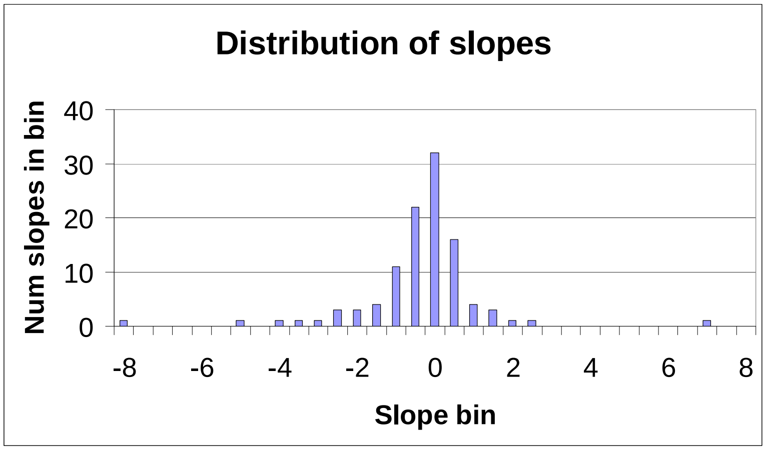

Figure 3-8. A histogram showing the distribution of slopes of the function relating velocity detection threshold to real-hand velocity for all conditions.

Three sets of data yielded ogive fits whose correlation coefficient with the actual data was not statistically significant. These datasets, therefore, could not be used. In the remaining 234 half-sets of data (117 pairs), there were 15 detection thresholds greater than 1.0 that, as described in section 3.3.1, I replaced with the value of 1.0, yielding a conservative estimate.

I performed a mixed-model ANOVA with study condition as a fixed factor, participant number as a random factor, and slope as the outcome variable. Specifying the participant as a random factor adjusted for multiple observations within subjects. I tested the null hypothesis that the mean slopes in every condition were simultaneously equal to zero. I could not reject this null hypothesis with $F_{12,\ 78} = 1.22$, $p = 0.285$. Though this test does not prove that the detection threshold does not vary with hand speed, I was unable to prove that it does vary. I will continue to assume that the Weber's Law assumption holds.

3.4 Discussion

Study 2 yielded 12 velocity discrepancy threshold values for the hand avatar moving faster and slower in six directions of real-hand motion. These values are necessary to implement MACBETH.

Chapter 4: Design of MACBETH

MACBETH removes position discrepancy by introducing velocity discrepancy that is *equally detectable*. The ideal methodology for choosing this equally detectable velocity discrepancy is based on the following four-step algorithm (Figure 4-1):

1) Find the existing position discrepancy.
2) Find the probability of detecting that discrepancy.
3) Find the point on the velocity discrepancy psychometric function with an equal detection probability.
4) Find the velocity discrepancy that corresponds to that rate of detection.

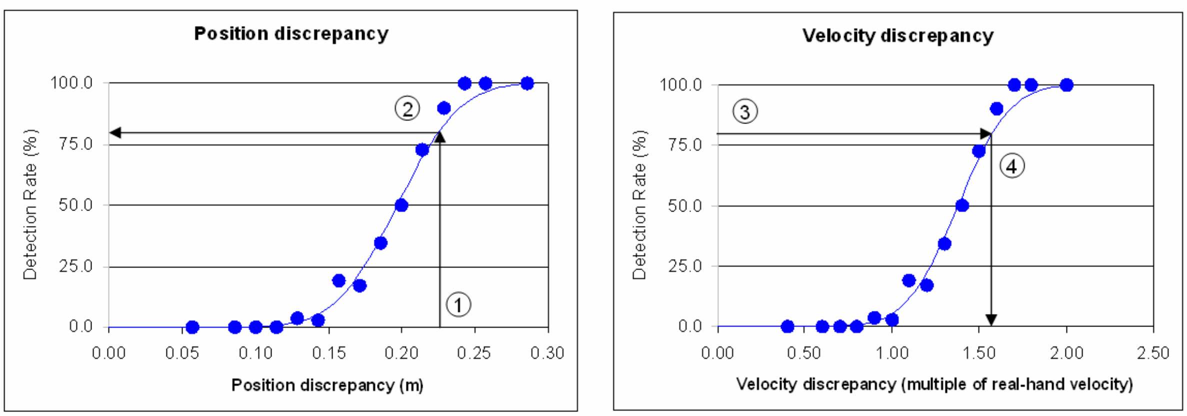

Figure 4-1. The idea behind MACBETH : 1) Find the existing position discrepancy 2) Find the probability of detecting that discrepancy 3) Find the point on the velocity discrepancy psychometric function with an equal detection probability. 4) Find the velocity discrepancy that corresponds to that rate of detection.

However, this algorithm requires psychometric functions for both position and velocity discrepancy. Since each person's psychometric functions are different and obtaining the data necessary to construct a psychometric function requires hours, this algorithm would have a lot of calibration overhead for each user. It is therefore impractical. Some simplifying assumptions are necessary to make MACBETH practical.

4.1 Assumptions to make MACBETH practical

Assumption 1: A user's psychometric functions for position and velocity discrepancy are the same shape, such that when their *x* axes are normalized by dividing stimulus values by the 50% detection threshold, the two functions are identical.

Justification: Assumption 1 proceeds from the idea that the same set of factors determines the relation between vision and proprioception for both position and velocity judgments. Therefore, the mean position and velocity detection thresholds will be correlated to their variances in the same way. Since the mean and variance of the data determine a psychometric function's shape, the two psychometric functions will be the same shape. I have no evidence for this claim, but am willing to assume it is reasonably accurate.

Implications: The entire psychometric function is no longer necessary: equal normalized stimulus levels are equally detectable. All that is necessary is the 50% detection thresholds with which to normalize the stimulus levels (Figure 4-2).

Though with Assumption 1 it is only necessary to have the 50% detection threshold for position and velocity discrepancies, rather than the whole psychometric function for each, the method described in Study 2 to find the 50% detection threshold requires first finding a psychometric function. Therefore, another assumption is necessary to avoid having to find psychometric functions for each user.

Assumption 2: Each individual's position and velocity discrepancy thresholds vary from the population means in the same proportion.

Justification: As with Assumption 1, Assumption 2 proceeds from the idea that the same set of factors determines the relation between vision and proprioception for both position and velocity judgments. Therefore, a good observer of position discrepancy will also be a good observer of velocity discrepancy, so the two detection thresholds will vary together.

Implications: The population mean 50% position and velocity discrepancy thresholds represent stimuli levels of equal detectability for every individual and can still be used as anchors for normalizing stimulus levels. Therefore, a single user study could measure an estimated population-mean detection threshold which could be used for every user.

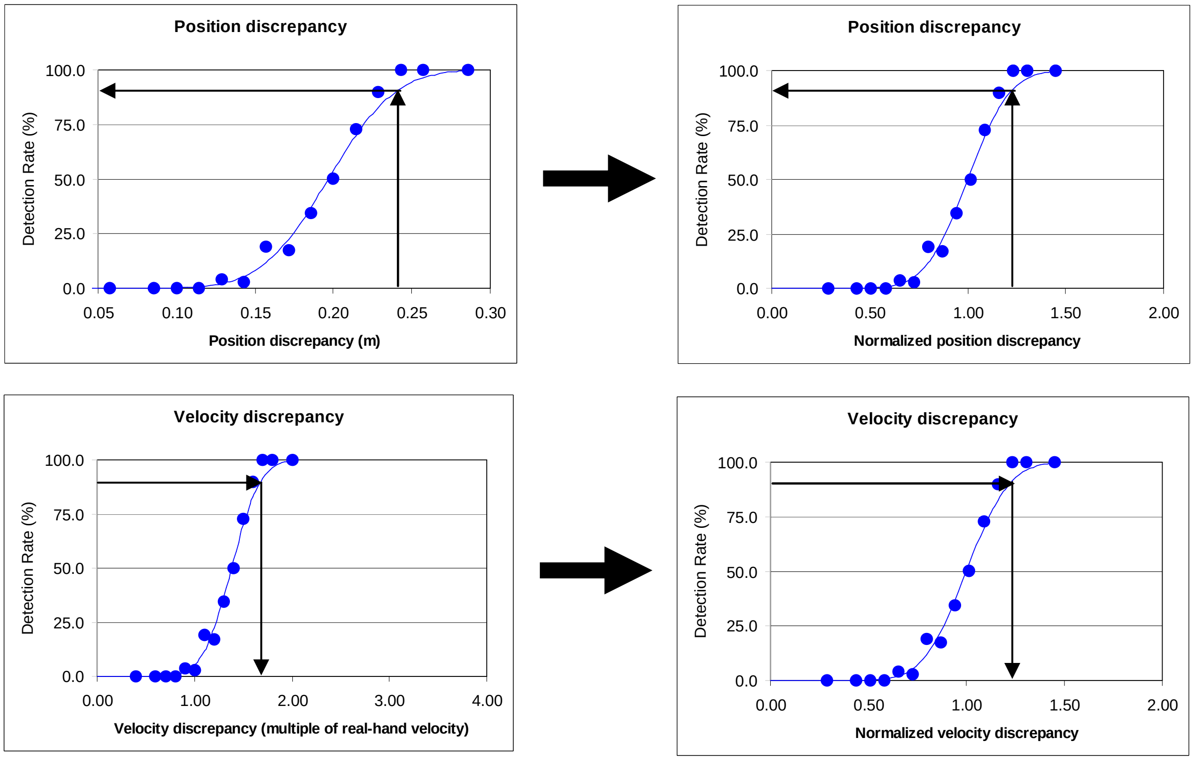

Figure 4-2. If the psychometric functions for position and velocity discrepancy are similarly shaped (left), such that when they are normalized by dividing the stimulus levels by the 50% detection threshold, the functions become identical (right), the appropriate velocity discrepancy will be the normalized velocity discrepancy of the same value as the normalized position discrepancy.

These assumptions allow for the implementation of MACBETH without requiring information about each individual user. But one difficulty still remains. Whereas the velocity discrepancy thresholds from Study 2 are clearly 50% detection thresholds, it is not clear that the position discrepancy thresholds from Study 1 are also 50% detection thresholds. Method-of-limits designs do approximate 50% detection thresholds, but in Study 1 only the ascending series half of such a design was used. Since it was not balanced with a descending series, Study 1 might have overestimated the 50% position discrepancy detection threshold. Based on the studies performed, one final assumption is necessary for the implementation of MACBETH.

Assumption 3: Thresholds from study 1 were 50% thresholds.

Justification: Assumption 3 might be true or it might not. However, the measured thresholds are plausible approximations for 50% thresholds, and making this assumption is necessary to proceed with the implementation of MACBETH.

Implications: MACBETH can be implemented using the data from Studies 1 and 2, though a potential overestimation of the position discrepancy threshold might result in position discrepancy being removed more slowly than necessary.

4.2 MACBETH algorithm

//Part 1:
//Based on the movement of the real hand, determine the desired avatar hand position if it
//were not to collide with any objects

//Start like the incremental motion method

```
idealMovement = realHandPosition - previousRealHandPosition
goalPosition  = avatarPositionLastFrame + idealMovement

Convert positions to camera-centered spherical coordinates

For each spherical coordinate component
     multipleOfPositionDiscrepThreshold =
       abs(offsetFromAvatarToRealHand) /
       positionDiscrepancyThreshold
```

//Set appropriate velocity discrepancy threshold based on the direction of real
//hand movement, and whether the hand avatar needs to be moved faster or slower
//to move it closer to the user's real hand (this referred to as *MACBETH's*
//*equation*)

```
     velocityDiscrepThreshold =
       appropriateVelocityDiscrepThreshold
     velocityDiscrepancy =
       multipleOfPositionDiscrepThreshold *
       velocityDiscrepThreshold * realHandMovement
```

//Take the extra movement created by the velocity discrepancy in each component, and
//add to the goal hand avatar position

```
goalPosition = goalPosition +
                movementDueToVelocityDiscrepancy

If movement due to velocity discrepancy causes hand avatar
to overshoot position of real hand in any component
     Set the hand avatar position component to be that of
     the real hand

Convert positions back to Cartesian coordinates
```

//Remove anomalies due to switching to spherical coordinates

```
If position discrepancy component is larger than it started
     Set component to its previous value
```

//Part 2:
//Perform collision detection and correction

//To approximate continuous collision detection, move hand avatar to goal position in
//several steps

```
While no collisions have been detected and the hand avatar
has not reached the goal position
     Move the current hand avatar position one step
     Test for collisions
     If collision
          Move the avatar hand out of the object in the
          direction perpendicular to the face it penetrated
```

4.3 Threshold values used to implement MACBETH

The threshold values measured in Studies 1 and 2 were used in the implementation of MACBETH (Table 4-1).

Table 4-1. Detection threshold values measured in Studies 1 and 2, used in the implementation of MACBETH.

Position discrepancy threshold		
All directions treated the same	19.09º (0.20 m)	
Velocity discrepancy threshold		
Real-Hand Motion Direction	Faster scale factor	Slower scale factor
Left	+0.44	-0.08
Right	+0.40	-0.06
Up	+0.51	-0.16
Down	+0.38	-0.27
Toward	+0.63	-0.46
Away	+0.69	0.00

4.4 Motion profiles

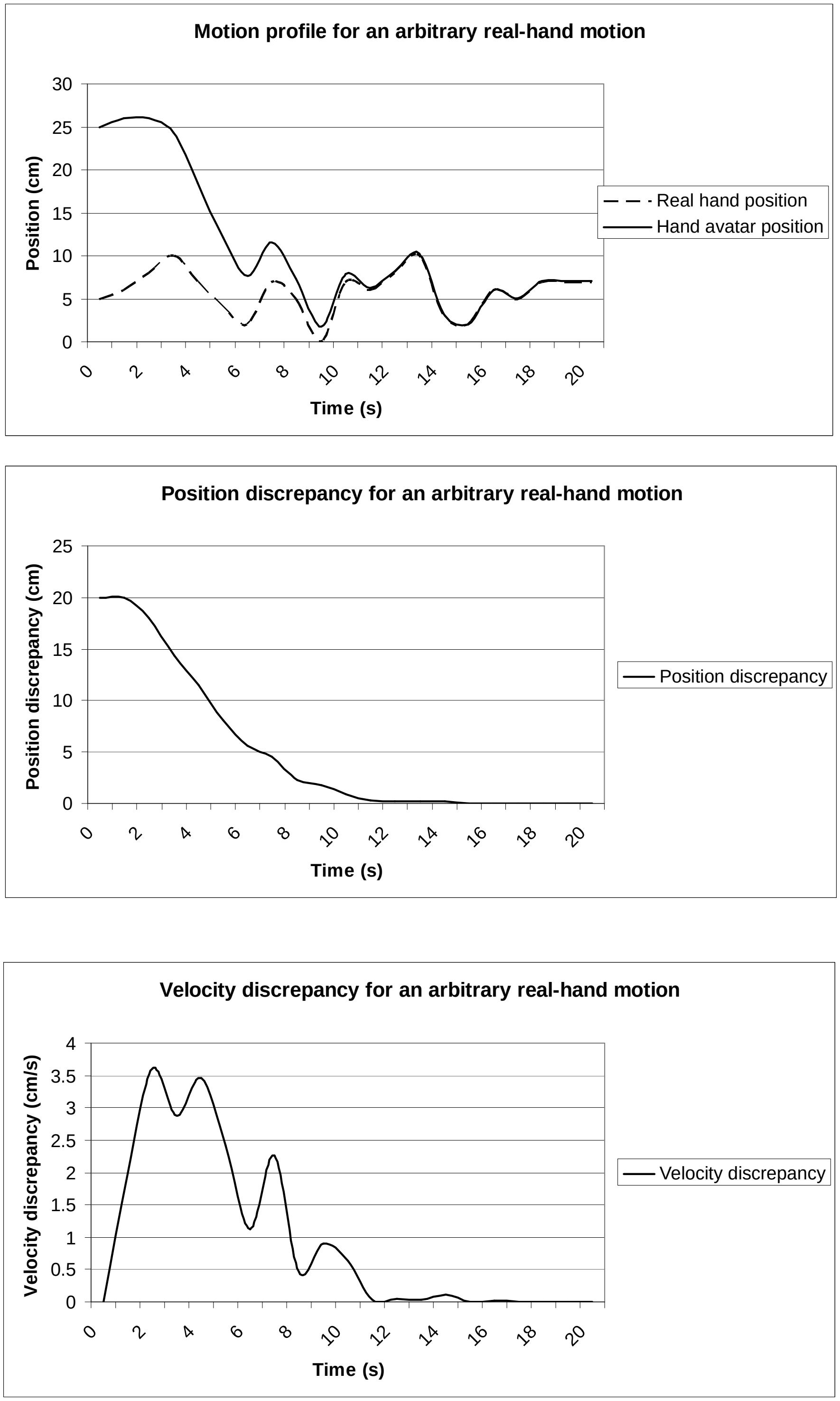

Figure 4-3. Motion profile for an arbitrary real-hand motion using MACBETH

An example MACBETH motion profile for an arbitrary hand motion in one dimension is shown in the graphs in Figure 4-3. The position discrepancy decreases steadily, but the velocity discrepancy varies as a function of both position discrepancy and real-hand velocity.

Figure 4-4 shows the position and velocity discrepancy over time for each of the three techniques in a case like the one shown in Figure 1-3 where a user penetrates a virtual object and then moves his hand away from the object at a steady velocity. The rubber-band technique decreases the position discrepancy most quickly, but does so by exhibiting its characteristic "sticking" problem. The hand avatar does not move until the real hand meets it, leading to a velocity discrepancy equal to the user's real-hand velocity.

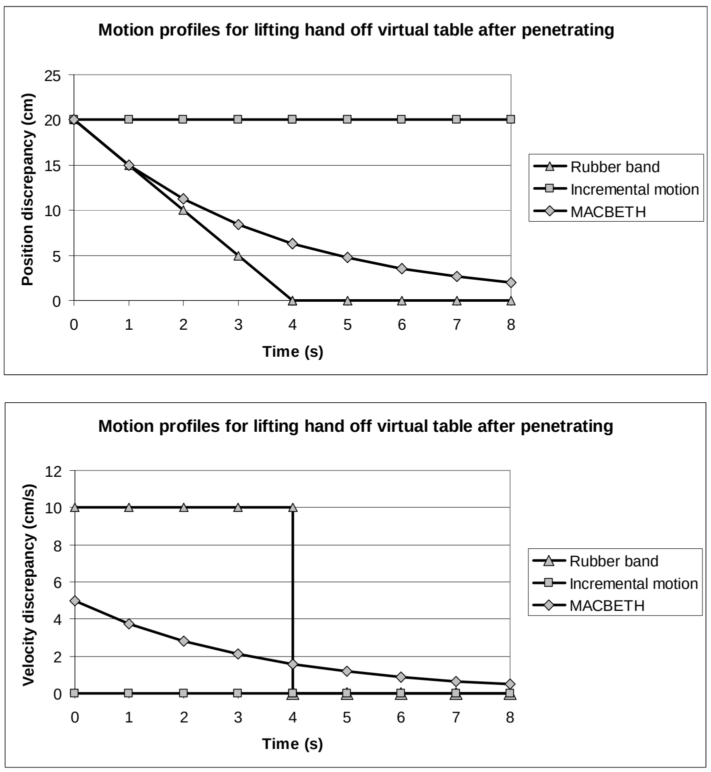

Figure 4-4. Position and velocity discrepancies for each technique when a user penetrates a virtual object and then removes his real hand at a constant velocity.

4.5 All three techniques are instances of virtual coupling

Virtual coupling is a method used to calculate forces for haptic displays when a user penetrates a simulated object [Colgate, Stanley, and Brown, 1995]. The user's real hand on the haptic display handle is connected virtually to the simulated hand by a damped spring. The force displayed by the haptic device is then calculated by the damped spring equation:

$$F = -kx - Bv$$

Where k is the spring constant and B is the damping coefficient.

The rubber-band technique, the incremental-motion technique, and MACBETH can all be viewed as versions of this model with different values for the constants. In the absence of a haptic display, calculating a force is not useful; however, calculating the movement of the avatar hand is. The force in the above equation can be replaced using Newton's second law of motion:

$$F = ma$$

which yields:

$$ma = -kx - Bv$$

In this equation, the three constants do not have any physical meaning. Since the avatar hand is not real, it does not have mass and there is no real spring to have a spring constant and damping coefficient. To simplify the equation, the mass can be set to 1, leaving:

$$a = -kx - Bv$$

By manipulating k and B, all three hand avatar management techniques can be represented.

The rubber-band technique corresponds to a spring-damper system with an infinite spring constant and a finite damping coefficient, meaning the acceleration of the hand avatar with respect to the real-hand position is potentially infinite. The incremental-motion technique corresponds to a spring-damper system with a spring constant of 0 or an infinite damping coefficient with a finite spring constant, meaning the acceleration of the hand avatar with respect to the user's real hand position is always 0. MACBETH corresponds to a spring-damper system in which the damping coefficient is a function of the user's real-hand speed. This can be shown by combining the spring-damper equation above with MACBETH's equation:

$$v(i) = -\frac{velocity\_threshold * real\_hand\_velocity * x(i)}{position\_threshold}$$

to determine the appropriate amount of velocity discrepancy. First, acceleration in the spring-damper equation can be approximated by:

$$a = \frac{v(i) - v(i-1)}{\Delta t}$$

to get:

$$\frac{v(i) - v(i-1)}{\Delta t} = -kx - Bv(i-1)$$

Solving for *v(i)* yields:

$$v(i) = -\Delta tkx(i-1) + (1 - B\Delta t)v(i-1)$$

For simplicity of notation, a change of variable can be applied to the MACBETH equation as follows:

$$v(i) = -\frac{velocity\_threshold * real\_hand\_velocity * x(i)}{position\_threshold} = -Qx(i)$$

Substituting this value of *v(i)* into the spring-damper equation yields:

$$-Qx(i) = -\Delta tkx(i-1) + (1 - B\Delta t)[-Qx(i-1)]$$

Since the current position of the hand avatar is the previous position plus the amount of change, the left side of the equation can be expanded:

$$-Q[x(i-1) + v(i-1)\Delta t] = -\Delta tkx(i-1) + (1 - B\Delta t)[-Qx(i-1)]$$

Substituting for *v(i-1)*, once again, yields:

$$-Q[x(i-1) - Qx(i-1)\Delta t] = -\Delta tkx(i-1) + (1 - B\Delta t)[-Qx(i-1)]$$

Simplifying this equation yields:

$$B = Q + \frac{k}{Q}$$

Expanding Q yields:

$$B = \frac{velocity\_threshold * real\_hand\_velocity(i)}{position\_threshold} + \frac{k}{\frac{velocity\_threshold * real\_hand\_velocity(i)}{position\_threshold}}$$

The velocity and position thresholds for a given directional component of discrepancy are constant. If *k* is chosen as an arbitrary constant, the value of *B* for a given direction of discrepancy varies as a function of the user's real hand velocity.

4.6 Computational time

Without any claim that the rubber-band and incremental-motion implementations are optimal, I measured the per-frame time required for each calculation. The median computation time required for MACBETH was slightly larger than that required for either of the other techniques. However, the computation time was still a mere fraction of the 16.7 ms frame time on a 60 Hz display (Figure 4-5).

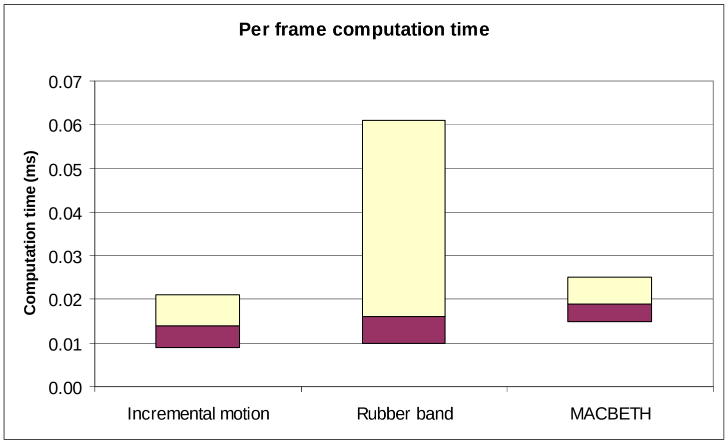

Figure 4-5. Per-frame computation time – the bottom of each box represents the 25th percentile mark, the mid-line is the median, and the top of the box represents the 75th percentile.

Chapter 5: **Study 3 – Evaluating MACBETH**

5.1 Hypothesis

MACBETH offers an improvement over the rubber band or incremental motion methods in user-rated naturalness, user preference of VE experience, or user performance on a hand navigation task.

5.2 Study Design

Testing the avatar management techniques required a task that would ensure that the user's hand would collide repeatedly with virtual objects. Having the user navigate the hand avatar through a tight maze would make it nearly impossible to avoid frequent collisions. The participant's time to complete the maze then measured performance.

Since the maze design might give a particular avatar management technique a performance advantage, I intended to design three mazes: two that tipped the scales to the advantage of MACBETH's two competitors and one average case.

The maze designed to favor the rubber-band method had a staircase shape, which required repeated up and left motions (Figure 5-1). As the avatar hand slid off the edge of a surface, it would snap to the user's real-hand position, covering the distance more quickly than with either the incremental-motion method or MACBETH. Furthermore, this method put the incremental-motion model at a disadvantage because these repeated collisions in the same direction would result in a position discrepancy that grew with every collision.

I could not find a maze to favor the incremental-motion method. I thought a maze that required repeated back and forth movements, with the same number of collisions in each direction would favor the incremental-motion method because the increments of discrepancy would tend to cancel each other out, whereas with the rubber-band method the user would have problems with the ball sticking to surfaces as he moved his hand from one surface to an opposite surface and back, repeatedly. A pilot study showed that such a maze did not actually favor the incremental-motion method. It proved difficult to design any maze that did give the advantage to the incremental-motion method. So, I was forced to abandon that goal.

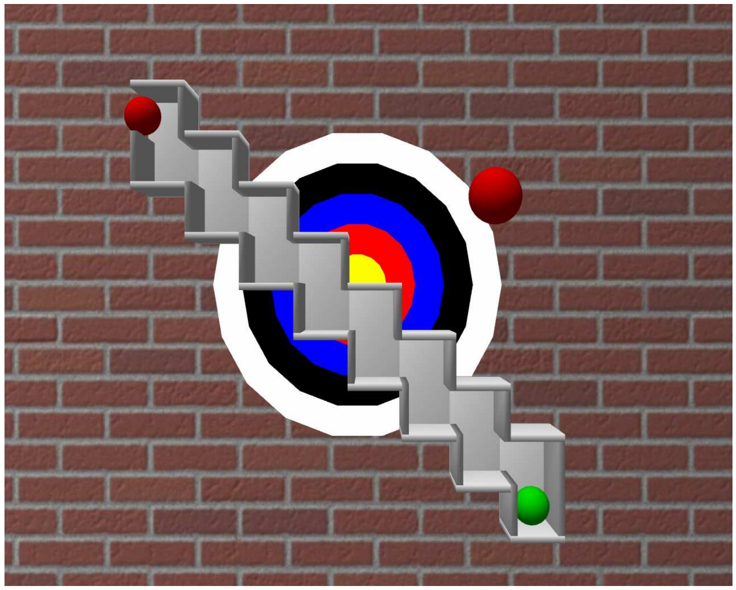

Figure 5-1. Staircase Maze – The participant maneuvered the hand avatar (the red ball in the upper right) through the maze from the green ball in the lower right to the red ball in the upper left.

Therefore, the only other maze was the “average case.” I randomly generated 10 mazes by starting from the entrance and using a random number generator to decide in which direction the maze would go next. The mazes were constrained to fit in an 8x8 grid so that they did not become too wide or tall for users to navigate them. The maze with the longest path length of the ten was then selected for use in the study (Figure 5-2).

I wanted to test the hypothesis proposed in Chapter 2 that the potentially large position discrepancies that can arise with the incremental-motion technique lead to a misperception of hand position. Therefore, I added a shooting task to the end of each trial to see if users’ accuracy varied with avatar management technique. I measured the time to shoot after completing the maze and the distance from the target center to where the ball landed.

I was most interested in which avatar management technique participants would feel was most natural. Therefore, on each trial, users were asked to rate the naturalness of the avatar management technique on a scale from 1 to 9.

I paired trials so that each would use a different avatar management technique. At the end of the two trials, participants were asked which of the two they preferred. Statistical tests for binary data such as I recorded from this question have low power (probability of rejecting the null hypothesis if the null hypothesis is indeed false), but I decided to try anyway.

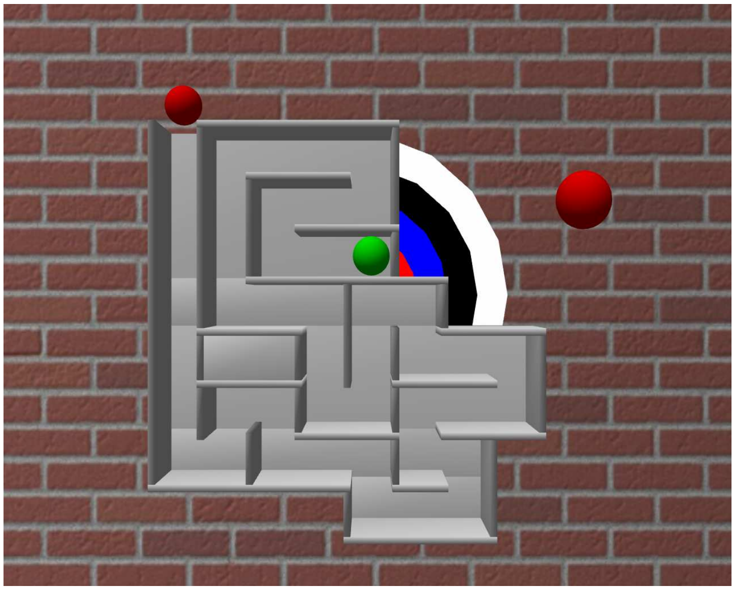

Figure 5-2. Randomly-generated maze – The second maze was generated to be an "average" case.

Each user was to see all possible pairings of the three avatar management techniques, in both possible orders, with both mazes. This yielded 12 sets of 2 trials. These pairings were balanced using a balanced Latin square matrix.

A pilot study suggested that there might be a significant learning effect during the first few trials of the study. Every participant, therefore, ran through the 12 pairs of trials once as a training period and then did so again for real.

5.3 Study Execution

5.3.1 Participants

Twelve right-handed introductory psychology students (3 males and 9 females) participated in this study. All gave consent and were given class credit for their participation.

5.3.2 Equipment

Each participant wore a Virtual Research Systems V8 head-mounted display and held a joystick in the right hand. Both the head and hand were tracked using a 3rdTech Hiball 3000. Participants sat in a chair and were visually immersed in the VE from Study 2 (Figure 3-1).

5.3.3 *The sequence of a pair of trials*

At the beginning of each trial, the user's right hand avatar was displayed as a red sphere capable of passing through virtual objects. A green sphere indicated the desired starting position of the user's hand for the trial. When the participant moved his hand to the green sphere, the red sphere turned green and could no longer pass through virtual objects.

The participant then moved his hand through the maze, which was suspended in air in front of the user. If the user pulled his hand toward himself, out of the maze, the ball that followed the user's hand became red again, and a green sphere appeared, indicating the last position of the hand that was inside the maze. When the red sphere again intersected the green sphere, the user continued moving the hand through the maze until it reached a red sphere at the exit of the maze.

When the user reached the red sphere at the exit of the maze, the maze disappeared and the user aimed at a target on the wall behind the maze (Figure 5-1) and clicked the button on the joystick to fire the ball at the target. When the button was clicked, the ball traveled in the direction of the vector connecting the user's dominant eye with the center of the ball. The ball traveled until its center hit the wall, where it stayed embedded in the wall. The experimenter then asked him to rate the naturalness of the avatar management technique.

The user then performed another trial, identical except for the use of a different hand avatar management technique. At the end of the second trial, the experimenter also asked the participant which of the two trials he preferred. To give participants incentive to do as well as possible, a feedback screen appeared on the wall after the user answered. The feedback screen displayed all three performance measures for each of the two trials: maze time, distance from the target center, and shooting time.

5.3.4 *Data*

Each trial yielded four pieces of data:

1) Maze traversal time
2) Time to pull the trigger, measured from the end of the maneuvering task
3) Distance from the center of the target
4) A rating on a scale from 1 to 9 of the naturalness of the avatar management technique, plus an explanation of what seemed unnatural, if anything

Each *pair* of trials yielded the following additional piece of data:

5) Participant's preference between the two avatar management methods

5.4 Results

For each measurement, I analyzed the data from all trials together and then analyzed each set of trials for each maze separately. This is true for all values except the performance time values, which are really not comparable across maze type. The mazes were different lengths, requiring different navigation times, and their end points were at different positions, so moving to aim at the target would take different amounts of time.

5.4.1 User rating of naturalness

Table 5-1. Results of an ordered multinomial regression of naturalness rating on avatar management technique – the results of the overall test of all values equal are presented along with unadjusted pairwise comparisons.

Test	Mean	d.f	χ^2	*p*
Both mazes				
Overall		2	10.00	0.007
Incremental-motion vs. rubber-band	5 5.68	1	1.41	0.24
Incremental-motion vs. MACBETH	5 6.75	1	49.79	< 0.001
Rubber-band vs. MACBETH	5.68 6.75	1	3.93	0.047
Staircase Maze				
Overall		2	7.66	0.022
Incremental-motion vs. rubber-band	4.23 6.08	1	6.35	0.012
Incremental-motion vs. MACBETH	4.23 6.19	1	16.93	< 0.001
Rubber-band vs. MACBETH	6.08 6.19	1	0.05	0.82
Randomly-generated Maze				
Overall		2	10.12	0.006
Incremental-motion vs. rubber-band	5.77 5.27	1	0.28	0.60
Incremental-motion vs. MACBETH	5.77 7.31	1	53.19	< 0.001
Rubber-band vs. MACBETH	5.27 7.31	1	12.15	0.001

Users rated MACBETH as more natural than the other two techniques, although on the staircase maze, user ratings were not statistically significantly higher than the rubber-band method (Figure 5-3). Because these data fall into discrete categories which have an inherent order, I tested these data with the parametric ordered multinomial regression test.

Results of the ordered multinomial regression of naturalness rating on avatar management technique, adjusting for multiple observations per participant, are given in Table 5-1. All three overall tests rejected the null hypothesis that all mean naturalness ratings were equal. Pairwise estimates showed all but three pairs to be significantly different. In all tables, results falling short of statistical significance are shown in gray.

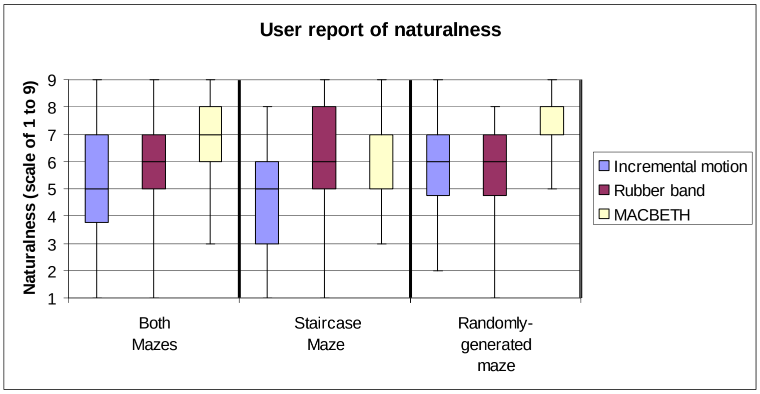

Figure 5-3. User reports of naturalness on a scale from 1 to 9 - the bottom of each box represents the 25th percentile mark, the mid-line is the median, and the top of the box represents the 75th percentile. Error bars represent the minimum and maximum responses.

5.4.2 Preference

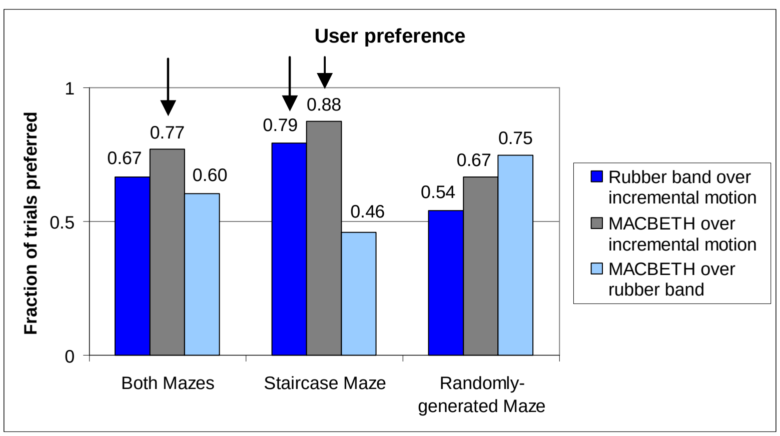

Figure 5-4. User preference - values represent the fraction of the pairs of trials in which the user chose the first technique over the second – statistically significant results are indicated with arrows

Figure 5-4 shows that in a 2-alternative forced-choice test, users preferred MACBETH to incremental motion in all cases and preferred MACBETH to the rubber-band method in the randomly-generated maze and in the overall analysis. However, users preferred the rubber-band method over MACBETH on the staircase maze. A logistic regression showed only three of the pairwise comparisons to be statistically significant (Table 5-2).

Table 5-2. Results of a logistic regression on the preference data, testing the null hypothesis that the probability of preferring one technique over another equaled 0.5 – the results of the overall test of all values equal to 0.5 are presented along with the individual unadjusted tests.

Test	**Fraction preferred**	**d.f.**	χ^2	***p***
Both mazes				
Overall		3	9.84	0.020
Rubber-band over incremental-motion	0.67	1	2.66	0.10
MACBETH over incremental-motion	0.77	1	6.25	0.012
MACBETH over rubber-band	0.60	1	0.81	0.37
Staircase Maze				
Overall		3	10.05	0.018
Rubber-band over incremental-motion	0.79	1	5.66	0.017
MACBETH over incremental-motion	0.88	1	6.15	0.013
MACBETH over rubber-band	0.46	1	0.09	0.76
Randomly-generated Maze				
Overall		3	7.87	0.049
Rubber-band over incremental-motion	0.54	1	0.11	0.74
MACBETH over incremental-motion	0.67	1	2.04	0.15
MACBETH over rubber-band	0.75	1	3.50	0.062

Because statistical tests on binary data have notoriously low power, I decided to increase the power by using all data available. Assuming that users' preference for the techniques did not change significantly over the course of the study, I added the preference data for the first 12 trial pairs and also the data from a pilot run of the study with 24 participants. The pilot study was identical to the final study except that participants were not asked to rate the naturalness of the technique, and they did not have the 12 training trial pairs before beginning the trials that counted. Adding these two extra sets of data quadrupled the amount of preference data. Running the same logistic regression on this larger data set, I found that the overall preference for MACBETH over the rubber-band method became statistically significant, as did the preference for MACBETH over both the incremental-motion and rubber-band methods for the randomly-generated maze (Table 5-3). However, both the overall preference and the preference on the randomly-generated maze for the rubber-band method over the incremental-motion method remained not statistically significant, and the preference for the rubber-band method over MACBETH on the staircase maze disappeared completely (Figure 5-5).

Table 5-3. Results of a logistic regression on the preference data (including data from training trials and pilot study), testing the null hypothesis that the probability of preferring one technique over another equaled 0.5 – the results of the overall test of all values equal to 0.5 are presented along with the individual unadjusted tests.

Test	Fraction preferred	d.f.	χ^2	*p*
Both mazes				
Overall		3	20.64	< 0.001
Rubber-band over incremental-motion	0.54	1	0.66	0.42
MACBETH over incremental-motion	0.77	1	51.70	< 0.001
MACBETH over rubber-band	0.67	1	12.67	< 0.001
Staircase Maze				
Overall		3	17.79	< 0.001
Rubber-band over incremental-motion	0.68	1	7.40	0.007
MACBETH over incremental-motion	0.79	1	29.92	< 0.001
MACBETH over rubber-band	0.50	1	0	1.0
Randomly-generated Maze				
Overall		3	19.20	< 0.001
Rubber-band over incremental-motion	0.41	1	2.40	0.12
MACBETH over incremental-motion	0.75	1	30.03	< 0.001
MACBETH over rubber-band	0.83	1	26.93	< 0.001

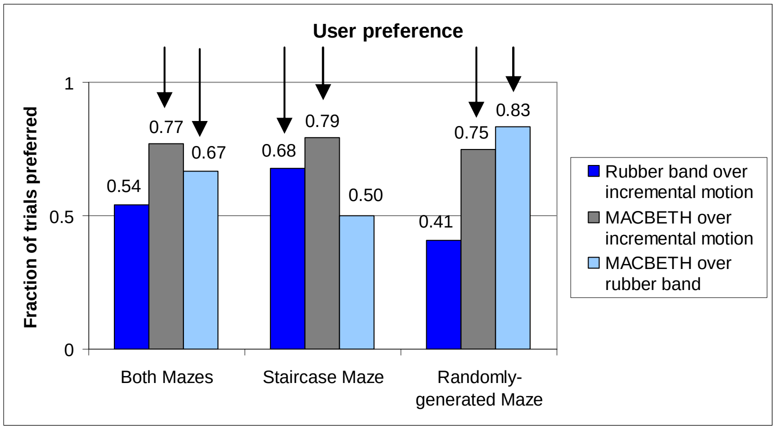

Figure 5-5. User preference (including data from training trials and pilot study) - values represent the fraction of the number of pairs of trials between two techniques in which the user chose the first technique over the second – statistically significant results are indicated with arrows

5.4.3 Time to navigate through maze

On the staircase maze, users performed worst with incremental motion, better with MACBETH, and best with the rubber-band method. On the randomly-generated maze, users performed best with MACBETH, but not statistically significantly better than with the rubber-band method (Figure 5-6). A mixed model ANOVA, adjusted for multiple observations per participant, showed that both MACBETH and the rubber-band method yielded statistically significantly better performance than the incremental-motion method (Table 5-4).

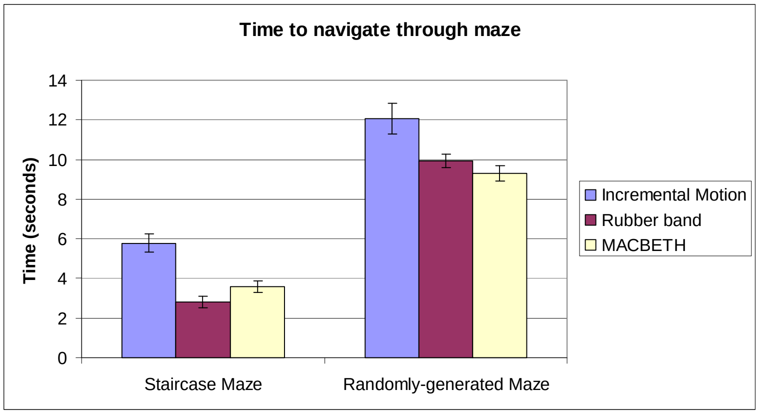

Figure 5-6. Mean times to navigate through the mazes (smaller numbers are better) - error bars represent 95% confidence intervals for the means

Table 5-4. Results of a mixed model ANOVA on time to navigate through maze, adjusting for multiple observations per participant. Results from the overall test of all values equal are presented along with unadjusted pairwise comparisons.

Staircase Maze				
Test	**Means**	**d.f.**	***F***	***p***
Overall		2, 130	39.82	< 0.001
Incremental-motion vs. rubber-band	5.77 2.82	1, 130	74.13	< 0.001
Incremental-motion vs. MACBETH	5.77 3.59	1, 130	40.32	< 0.001
Rubber-band vs. MACBETH	2.82 3.59	1, 130	5.11	0.026
Randomly-generated Maze				
Overall		2, 130	11.55	< 0.001
Incremental-motion vs. rubber-band	12.06 9.93	1, 130	12.46	< 0.001
Incremental-motion vs. MACBETH	12.06 9.29	1, 130	21.07	< 0.001
Rubber-band vs. MACBETH	9.93 9.29	1, 130	1.12	0.29

5.4.4 Shooting accuracy

I hypothesized that the large position discrepancies possible with the incremental-motion method would lead to a perceived hand position other than that of the hand avatar. Such a perception would lead to poor performance on an aiming task that required knowledge of the hand avatar position. Participants indeed performed better on the shooting task with both MACBETH and the rubber-band method than with the incremental-motion method (Figure 5-7). There was no statistically significant difference between performance with MACBETH and the rubber-band method on either maze (Table 5-5). However, observation of participants suggests that the poor shooting performance using the incremental-motion technique was not due to a misperception of hand position but rather, the increased physical difficulty of aiming the hand avatar that was in front of the body by using the real hand which was displaced laterally by a large amount.

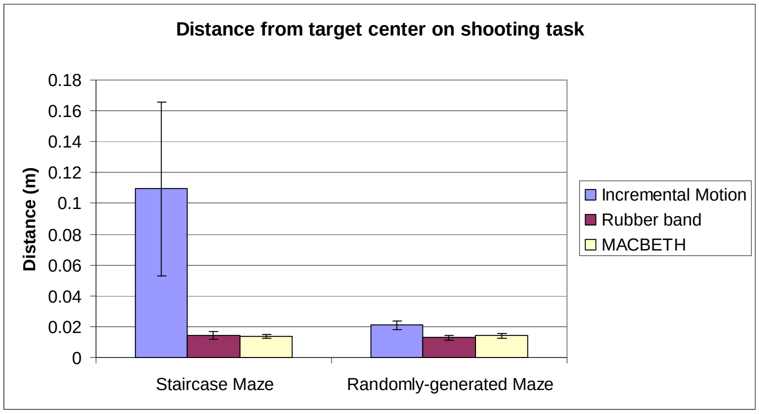

Figure 5-7. Mean distances from the target center on the shooting task - error bars represent 95% confidence intervals for the means

Table 5-5. Results of a mixed model ANOVA on distance from target center, adjusting for multiple observations per participant. Results from the overall test of all values equal are presented along with unadjusted pairwise comparisons.

Staircase Maze				
Test	**Means**	**d.f.**	***F***	***p***
Overall		2, 130	3.88	0.023
Incremental-motion vs. rubber-band	0.109 0.014	1, 130	5.81	0.018
Incremental-motion vs. MACBETH	0.109 0.014	1, 130	5.86	0.017
Rubber-band vs. MACBETH	0.014 0.014	1, 130	0.0001	0.99
Randomly-generated Maze				
Overall		2, 130	11.55	< 0.001
Incremental-motion vs. rubber-band	0.021 0.013	1, 130	12.46	0.001
Incremental-motion vs. MACBETH	0.021 0.014	1, 130	21.07	< 0.001
Rubber-band vs. MACBETH	0.013 0.014	1, 130	1.12	0.29

5.4.5 Time to shoot

Neither of the overall tests rejected the null hypothesis that all mean times to shoot (Figure 5-8) were equal (Table 5-6). Therefore, pairwise tests gave no useful information.

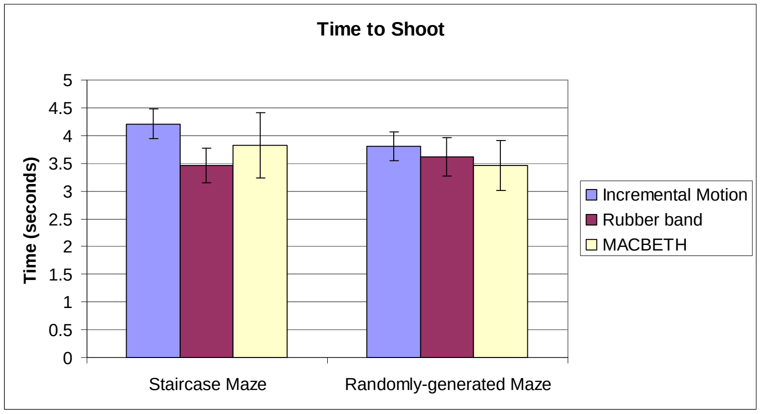

Figure 5-8. Mean times to shoot after completing the maze - error bars represent 95% confidence intervals for the means.

Table 5-6. Results of a mixed model ANOVA on time to shoot, adjusting for multiple observations per participant - Neither overall test produced statistically significant results.

Staircase Maze			
Test	**d.f.**	***F***	***p***
Overall	2, 130	1.70	0.19
Randomly-generated Maze			
Test	**d.f.**	***F***	***p***
Overall	2, 130	0.54	0.59

5.4.6 Independence of measures

It seems possible that a user's report of naturalness, preference, and performance would be all highly correlated and would be three measures of the same phenomenon, rather than three different measures. I performed an informal investigation into this possibility.

Overall, users rated MACBETH as most natural and preferred it over the other techniques, which suggests a direct correlation between naturalness and preference. However, participants preferred the technique that they rated most natural in only 53.3% of trials, and a graph of preference as a function of the difference in rated naturalness between two trials showed a slight *negative* trend (Figure 5-9). However, the data points vary from the trend line enough to suggest that naturalness and preference are not highly correlated at all.

A scatter plot of naturalness ratings versus maze times showed that naturalness ratings do not vary consistently with maze time (Figure 5-10). Therefore, any correlation between the two measures is very weak.

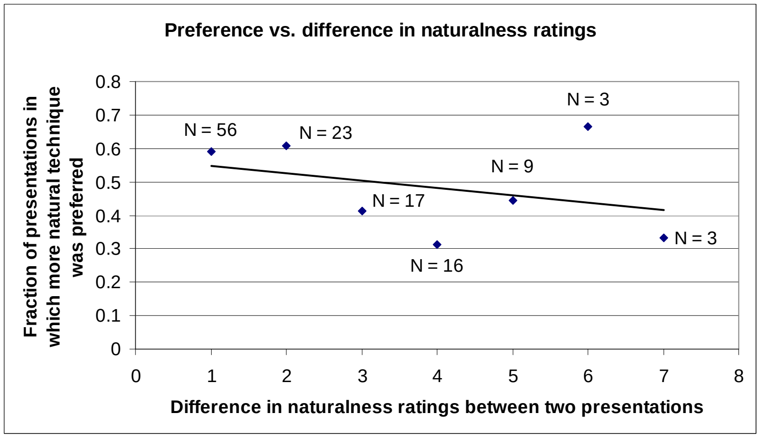

Figure 5-9. Users' preference for the technique they rated as more natural varied enough to suggest that the two measures are not highly correlated.

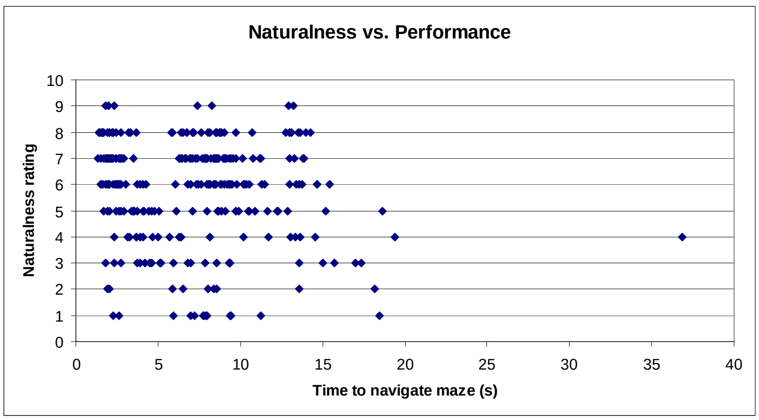

Figure 5-10. Values of maze time with respect to naturalness show a subtle downward trend.

Preference ratings also seemed to be uncorrelated with maze time. Participants preferred the trial in which they took longer to navigate the maze in 77 trials, and preferred the trial in which they took less time in only 67 trials. A chart showing which trial was preferred based on the difference in maze time between the two trials did not show any appreciable correlation between the two measures (Figure 5-11).

5.5 Discussion

The most interesting result from this study is that, overall, users rate MACBETH as a more natural technique than either the rubber-band or incremental-motion techniques and prefer MACBETH over both. Users did indeed perform better on the staircase maze with the rubber-band method, since that maze had been designed specifically to favor that method. However, it was notable that in the overall analysis of both mazes together, users performed statistically significantly better with MACBETH than with the incremental-motion method, and no worse than the rubber-band method.

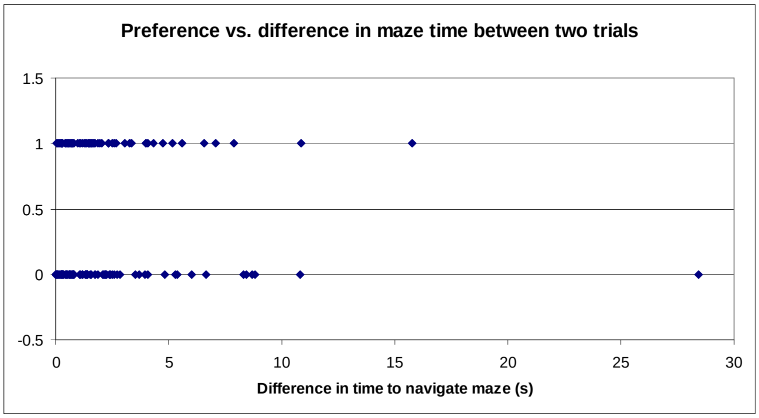

Figure 5-11. Trial preference as a function of difference in performance between the two trials. Pairs in which the trial with better performance was preferred were assigned a value of 1, and pairs in which the trial with worse performance was preferred were assigned a value of 0.

As hypothesized, users performed worse on the shooting task with the incremental-motion technique. However, from my observing the participants, it seems more likely that the physical task of aiming became more difficult as they were required to move their hand in some other direction to point the ball straight ahead, than that there was some sort of perceptual reason for the worse aiming.

Shooting time did not differentiate among the three methods.

Chapter 6: **Conclusions**

6.1 The thesis statement and the findings

The first part of the thesis is:

> Users are more likely to notice visual penetration of virtual objects by the hand avatar than the discrepancy in visual and proprioceptive hand-position cues introduced by preventing such penetration.

Study 1 showed this to be true.

The second part of this thesis is:

> If a user's hand avatar is rejoined to the real hand so that sensory discrepancy in position and velocity are equalized, one or more of the following will result:
> - The user will rate the technique as more natural.
> - The user will prefer his virtual environment experience.
> - The user will perform better on tasks in a virtual environment.

I used the data from Studies 1 and 2 to manage avatar sensory conflict such that sensory discrepancy in both displacement and velocity are minimized together. Study 3 then showed the first and second bullets to be true and the third most likely not to be true. However, users performed with MACBETH no worse than they did with the rubber-band and incremental-motion methods.

I found no reason that the incremental-motion method should ever be used. The rubber-band method is slightly easier to implement than MACBETH, but overall, it seems that MACBETH should be the method of choice.

6.2 What I would have done differently knowing what I do now and with plenty of time and money

6.2.1 Design Study 1 for direct comparison to Study 2

Study 1 was designed to compare the detectability of visual interpenetration to visual/proprioceptive discrepancy and not to compare position discrepancy thresholds to the velocity discrepancy thresholds measured in study 2. Table 6-1 shows the differences between the studies that weaken my claim that the discrepancy thresholds measured in each can be directly compared.

Table 6-1. Differences in study design

First study	Second study
Measured detection threshold using only an ascending series, which overestimates the 50% detection threshold	Measured a 50% detection threshold directly from a psychometric function
Users most likely underwent sensorimotor adaptation, which continually recalibrated how vision and proprioception relate to one another, increasing the threshold for detecting a difference between them	The randomness of the stimuli worked against sensory adaptation
Users distracted from detection task by playing the Simon® game	No secondary task to distract users

6.2.2 Run more participants for both Study 1 and 2

With more time and money I would have run many, many more participants through the redesigned Study 1 and Study 2. The 95% confidence intervals for the means in Study 2 were quite large. For example, the confidence interval for the mean detection threshold in the toward/slower condition ranged from 0.244 times the real hand speed to 0.843 times the real hand speed. With more participants, these confidence intervals would shrink and would remove uncertainty from the values that are used for the balancing of visual-proprioceptive discrepancy.

6.3 Future work

6.3.1 Packaging this up and making it publicly available

I should like to create a package that would take a tracked hand position and scene geometry as inputs and would return an avatar hand position as determined by MACBETH. However, this requires some sort of generalized collision-response library. I currently use SWIFT++ [Ehmann & Lin, 2001] to detect collisions, but preventing the hand from penetrating virtual objects I must do myself. It was simple to do so in Study 3 since the hand avatar was a sphere colliding with box-shaped objects. However, providing realistic collision response for arbitrary objects needs something more.

6.3.2 Prediction

If one could predict when a user is about to penetrate a virtual object and by how much that user is likely to penetrate, the system could introduce velocity discrepancy before collision to create an opposite position discrepancy to that about to occur. Therefore, when the user would see his avatar contact the object before his real hand actually reached the same space. He would then begin the process of stopping and would penetrate the object less. If the pre-collision position discrepancy were half the total distance the user would normally penetrate, then the user would have only penetrated half as deeply. Therefore, the maximum position discrepancy and subsequent velocity discrepancy would be halved (half before the collision, and half afterward).

Such a prediction algorithm requires knowledge of user behavior when contacting virtual objects. Perhaps all users act similarly, and a simple prediction technique based on distance from objects and direction of hand movement would suffice. However, it is also possible that every user acts differently and the prediction technique would need to adapt to the user's style of interaction. A study would then need to be done to make sure that the prediction actually reduced the average amount of discrepancy rather than increasing it due to prediction errors.

6.3.3 Rotation

This entire work has dealt with discrepancies in position, but not in orientation. When a user contacts an angled surface with his hand, he might expect to see his hand rotate to lie flat on the surface. Therefore, introducing some orientation discrepancy may be more natural for the user. However, it is unclear when to introduce a rotation and when not to. Furthermore, it is unclear whether the most natural method for removing rotational discrepancy is similar to that employed in MACBETH or if rotational discrepancy should be eliminated immediately as position discrepancy is in the rubber-band method.

6.3.4 Adding an arm

In all of my studies, participants had a disembodied hand avatar in the virtual environment. It is unclear how an avatar arm would affect the use of MACBETH. A difference in position of the hand requires a difference in orientation of the arm. So, the question of rotations would need to be answered first.

References

Blanch, R., Guiard, Y. & Beaudouin-Lafon, M. (2004). Semantic pointing: improving target acquisition with control-display adaptation. *Proceedings of the SIGCHI 2004 Conference on Human Factors in Computing Systems*, Vienna, Austria, ACM Press, 519 – 526.

Burns, E. and Brooks, F.P. (2006). Perceptual sensitivity to visual/kinesthetic discrepancy in hand speed, and why we might care. *Proceedings of Virtual Reality Software and Technology 2006*, Limassol, Cyprus, ACM Press, 3 – 8.

Burns, E. & Razzaque, S. (2003). Avatar redirection: An exploration of sensory conflict. *Presence-Connect*, 7.

Burns, E., Razzaque, S., Whitton, M.C., McCallus, M.R., Panter, A.T., & Brooks, F.P. (2006). The hand is more easily fooled than the eye: Users are more sensitive to visual interpenetration than to visual-proprioceptive discrepancy. *Presence: Teleoperators and Virtual Environments, 15*, 1-15.

Colgate, J.E., Stanley, M.C., & Brown, J.M. (1995). Issues in the haptic display of tool use. *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Pittsburgh, PA, 140 – 145.

Coren, S., Ward, L.M., & Enns, J.T. (1999). *Sensation and Perception*. 5th edition, Philadelphia: Harcourt Brace College Publishers, 20 – 25.

Dominjon, L., Lecuyer, A., Burkhardt, J.M., Richard, P. & Richir, S. (2005). Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. *Proceedings of IEEE VR 2005*, Alexandia, VA, IEEE Press, 19 – 25.

Ehmann, S. & Lin, M. (2001). Accurate and fast proximity queries between polyhedra using surface decomposition. *Computer Graphics Forum, Proceedings of Eurographics 2001*. Manchester, UK, ACM Press, *20*, 500 – 510.

Fechner, G.T. (1966). *Elements of Psychophysics*. Translated by Helmut E. Adler. New York: Holt, Rinehart and Winston, Inc, 54.

Gibson, J. (1933). Adaptation, after-effect and contrast in the perception of curved lines. *Journal of Experimental Psychology*, *16*, 1-31.

Howard, I.P. (1968). Displacing the optical array. In S.J. Freedman (ed.), *The neuropsychology of spatially oriented behavior* (pp. 19 – 36). Homewood, IL: Dorsey Press.

Insko, B., Meehan, M., Whitton, M., Brooks, F. (2001). "Passive Haptics Significantly Enhances Virtual Environments," *Proceedings of 4th Annual Presence Workshop,* Philadelphia, PA. (UNC-CH CS Tech Report No. 01-010, available at http://www.cs.unc.edu/Research/tech-report.html)

Jenkins, W.L. & Connor, M.B. (1949). Some design factors in making settings on a linear scale. *Journal of Applied Psychology, 33*, 395.

Jenkins, W.L., Maas, L.O., & Rigler, D. (1950). Influence of friction in making settings on a linear scale. *Journal of Applied Psychology, 34*, 434.

Jenkins, W.L. & Karr, A.C. (1954). The use of a joystick in making settings on a simulated scope face. *Journal of Applied Psychology, 38*, 457.

Lindeman, R., Sibert, J., & Templeman, J. (2001). The effect of 3D widget representation and simulated surface constraints on interaction in virtual environments. *Proceedings of IEEE VR 2001*, Yokohama, Japan, IEEE Press, 141 – 148.

MacKenzie, I.S. & Riddersma, S. (1994). Effects of output display and control-display gain on human performance in interactive systems. *Behaviour & Information Technology*, *13*, 328-337.

McNeely, W.A., Puterbaugh, K.D., & Troy, J.J. (1999). Six degree-of-freedom haptic rendering using voxel sampling. *Proceedings of SIGGRAPH 1999*, Los Angeles, CA, ACM Press, 401-408.

“Shakespeare, W. (1606). *Macbeth*. Act 2, Scene 1, Lines 33-45.

Slater, M., & Steed, A. (2000). A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments*, 9, 413-434.

Slater, M. & Usoh, M. (1994). Body centred interaction in immersive virtual environments. In Thalmann, N.M. and Thalmann, D. (eds.), *Artificial Life and Virtual Reality* (pp. 125-148). New York: John Wiley and Sons.

Somers, D. C., & McNally, R. (2003). Kinesthetic visual capture induced by apparent motion [Abstract]. *Journal of Vision*, 3(9), 35a.

van Beers, R., Sittig, A., & Denier van der Gon, J. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology*, *81*, 1355-1364.

van Beers, R., Wolpert, D., & Haggard, P. (2002). When feeling is more important than seeing in sensorimotor adaptation. *Current Biology*, *12*, 834-837.

Warren, D. H., & Cleaves, W.T. (1971). Visual-proprioceptive interaction under large amounts of conflict. *Journal of Experimental Psychology, 90*, 206-214.

Welch, R. (1986). Adaptation of space perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (eds.), *Handbook of perception and human performance* (pp. 24-1 – 24-45). New York, NY: John Wiley and Sons.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*. 88, 638-667.

Yuan, Y. (2000). Multiple imputation for missing values: Concepts and new Development. *SAS Users Group International Proceedings*, Indianapolis, Indiana, Paper 267.

Zachmann, G. & Rettig, A. (2001). Natural and robust interaction in virtual assembly simulation. *Eighth ISPE International Conference on Concurrent Engineering: Research and Applications (ISPE/CE 2001)*, Anaheim, CA.