

Estimation and Comparison of Rotation Gain Thresholds for Redirected Walking



Niall Williams

A thesis submitted to the Faculty of Davidson College
in partial fulfillment of the requirements for
Departmental Honors in Mathematics & Computer Science

Advisor: Tabitha Peck

Second Reader: Raghuram Ramanujan

May 2019

Acknowledgements

First and foremost, I would like to thank Dr. Peck for her unmatched support and enthusiasm. This thesis would not exist if you had not introduced me to wonderful world of computer graphics research. Thank you for pushing me to do my best work. Thank you for supporting and encouraging me at every step of this thesis. You have taught me how to remain positive during the low points of research, and you have helped me remain focused on what truly matters when doing research. Thank you for your patience and dedication during our countless meetings about research and life in general. Gathering outside opinions from you about my ideas and plans has been so valuable. I am truly grateful for all the guidance, honesty, and perspective you have given me during my time at Davidson.

I also thank Dr. Ramanujan for his support throughout this thesis. Your keen insights have been incredibly valuable and have taught me more than I could have hoped to learn about conducting good research and asking the right questions. I am also grateful to you for your advice on scientific writing.

I thank the entire Davidson Mathematics and Computer Science department for the welcoming community you have built. This department is incredibly friendly to *everyone*, and I admire your dedication to creating an inviting learning environment. I will miss talking to all of you about math, computer science, Seinfeld, and everything in between at the weekly Math/CS lunches. For any students who may be reading this, I encourage you to take advantage of this opportunity to build relationships with your professors outside of the classroom.

Finally, I thank my friends for their company during my time at Davidson. Going through college with you all has made the experience much more interesting and entertaining. I am especially grateful to Arthur, George, Jimmy, My, and Ryan. Arthur, Jimmy, and My, I am thankful for all the time we have spent together trying to debug code and prove theorems. George and Ryan, I will sorely miss our car rides full of thought-provoking, philosophical, and completely nonsensical discussions. More importantly, I will miss our almost weekly trips to Bojangles! Thank you all for putting up with all my bad jokes and (sometimes one-sided) conversations about computer science, Tetris, and paintings for two years. I look forward to seeing which directions our lives take us.

“Inspirational quote goes here.”

– Someone more successful than me

Abstract

Redirection techniques allow users to explore virtual environments (VEs) larger than the tracked work space by imperceptibly manipulating the VE according to rotation gain thresholds. Previous work has estimated these thresholds, but this previous work was conducted on now-outdated hardware with a 40° field of view (FOV) and without distractors—objects in the VE that aim to capture the user’s attention. We present a within-participant user study in which we estimate and compare detection threshold gains for 40° and 110° FOVs with and without distractors. Our results show that users tolerate more redirection with a wider FOV. Significant differences were found between female and male thresholds, as well as between men with 40° and 110° FOVs. We also found strong correlations between simulator sickness and threshold gains.

Psychophysical experiments often require that participants complete hundreds of trials in order to generate enough data for analysis. We investigate the viability of using fewer trials to yield similar results. Using participants’ confidence in the correctness of their responses, we implement a novel statistical method that claims to estimate psychometric curve parameters in fewer trials. Our results suggest that an alternative study design may be required to make use of this model, and future directions for research in this field are discussed.

Table of contents

List of figures	xi
List of tables	xv
1 Introduction	1
1.1 Walking in Virtual Reality	2
1.2 Thesis Statement	3
2 Perception	5
2.1 Non-visual Perception	5
2.2 Visual Perception	6
2.2.1 Optical Flow	6
2.2.2 Vection	7
2.2.3 Simulator Sickness	7
2.3 Gender Differences	8
3 Virtual Locomotion	11
3.1 Locomotion Interfaces	11
3.2 Real Walking in Virtual Reality	13
3.2.1 Redirected Walking	14

4 Estimation of Rotation Gain Thresholds	21
4.1 Equipment	21
4.2 Experiment Design	22
4.3 Virtual Environment	26
4.4 Participants	27
4.5 Procedure	28
4.6 Results	30
4.6.1 Analysis of $\psi(greater g_i)$	33
4.6.2 Gain analysis	34
4.6.3 Distractors	35
4.6.4 Simulator sickness	35
4.7 Discussion	36
5 Fitting Psychometric Functions	45
5.1 Introduction	45
5.2 Confidence Model	46
5.3 Methods	48
5.4 Results	48
5.5 Discussion	51
6 Conclusions and Future Work	61
6.1 Comparison of Rotation Gain Thresholds	61
6.2 Efficient Estimation of Detection Thresholds	62
6.3 Future Work	64
6.3.1 Redirected Walking	64
6.3.2 Confidence Model	65
References	67

List of figures

3.1	Two locomotion interfaces, powered shoes [31] (left) and treadmills [29] (right), that do not support real walking.	12
3.2	The blue rectangle on the floor represents the tracked space. The user can walk freely within this blue rectangle and the VR system will be able to track their movements. If they step outside this region, the VR system will not be able to track the user's movements. The lighthouses (circled in red) are responsible for tracking the position of the HMD that the user wears. Original image via [14].	13
3.3	Diagrams that illustrate how different RDW gains can be used to increase the size of the explorable VE. The green borders represent the real-world tracked space borders, and the purple borders represent the borders of the VE that correspond to the size of the tracked space. Arrows indicate the user (green) or VE (purple) movement.	16
4.1	(a) View of the VE in the left eye with no FOV modification (110° FOV). (b) View of the VE in the left eye with a 40° FOV restrictor.	22
4.2	Breakdown of the 2 (FOV: 40° , 110°) \times 2 (Gender: female, male) \times 2 (Distractor: without, with) study we conducted. Each box is one of the 2 values of each of the independent variables (FOV, gender, distractor status) in our study.	23

4.3	Sample view of the scene and the corridor of trees during a trial.	24
4.4	Experiment breakdown for one participant. Each subtree of the blue nodes represents one block.	25
4.5	Screen shot of the virtual forest scene with a distractor moving across the user's field of view (left) and a participant wearing the headset (right). . . .	26
4.6	The gamer-type distribution by gender for the experiment participants. Female participants are in red, and males are in blue.	27
4.7	The average probability of responding " <i>greater</i> " to the question, "Was the virtual movement <i>smaller or greater</i> than the physical movement?", with standard error, for each rotation gain, and the fitted psychometric curve. Responses for 40° FOV are in red, and 110° FOV are in blue. The vertical edges of the red and blue regions indicate the 25% (left edge) and 75% (right edge) thresholds. In each condition, gains that fall inside the colored regions are undetectable to users. (Top) Perceptual thresholds for all participants, (Middle) perceptual thresholds for female participants, (Bottom) perceptual thresholds for male participants.	40
4.8	(Top) The simulator sickness scores of each participant after the 40° FOV block (x-axis) and the 110° FOV block (y-axis). (Bottom) The simulator sickness scores of each participant after the 110° FOV block and the 110° FOV rotation gain at the 75% detection threshold. A male outlier is seen on the far right. Female participants are denoted with red circles, and male participants with blue triangles.	43
5.1	The effect of the confidence scaling factor on the cumulative Gaussian curve. When $k > 1$, the resulting curve (dashed) represents underconfidence. When $k < 1$, the resulting curve (dotted) represents overconfidence.	47

5.2	The psychometric curves calculated using the traditional method with eight trials (black), the traditional method with six trials (purple), and the CSD model with six trials (green). The CSD model comes close to approximating the “ground truth” represented by the black curve. The data used to generate these curves comes from participant 14.	51
5.3	The average confidence rating for each gain for participant 7 (110° FOV, with distractors). The curve we want to fit, a cumulative Gaussian, is shown as a black curve on the same graph. Fitting a cumulative Gaussian to the participant’s data in this case is infeasible.	52
5.4	Psychometric curves of the pilot participant, calculated using the method used in Section 4.	56
5.5	Distribution of confidence ratings in the pilot study for trials with a response of “ <i>greater</i> ”. The plots do not form a sigmoid shape, so the CSD model cannot fit a μ and $k\sigma$	58
5.6	Distribution of confidence ratings in the pilot study for the inverted pilot study data. The plots are closer to the sigmoid shape, so the CSD model does a better job at fitting a μ and $k\sigma$	59

List of tables

4.1	The PSE of the psychometric curve fit to each participant's $\psi(greater g_i)$. The - indicates participant data that could not be fit to a psychometric curve and was excluded from analysis.	32
4.2	The 25%, PSE, and 75% threshold gains derived from the psychometric curves (goodness-of-fit data (Deviance (D) and <i>p</i> -value)) calculated using maximum likelihood estimation, using a cumulative normal distribution function, of the pooled $\psi(greater g_i)$ data. Gains are presented by FOV and gender (female - ♀, male - ♂) for all trials regardless of distractors, trials without distractors, and trials with distractors.	41
4.3	Analysis of the 25%, PSE, and 75% gains calculated from each participant's psychometric curve with a 2 (distractors (DIS): present, not present) x 3 (Threshold (THR): 25%, PSE, 75%) x 2 (FOV: 40°, 110°) x 2 (Gender (GEN): female, male) ANOVA. Significance codes: ‘***’ $p < 0.001$, ‘**’ $p < 0.01$, ‘*’ $p < 0.05$, ‘+’ $p < 0.1$	42
4.4	Winsorized correlations with 10% trimmed means comparing threshold gains to FOV for all participants.	42

5.1 μ, σ and $k\sigma$ for each participant with 40° FOV, no distractors, and six trials per gain. The CSD model was unable to converge to any μ and $k\sigma$ for participant 7. There was no fit for participants 2, 5, and 12 using the traditional psychometric curve fitting process, so we are unable to use the CSD model to calculate μ and $k\sigma$ for them.	50
5.2 μ, σ and $k\sigma$ for the pilot participant. In this experiment, we tested gains ranging from 0.3 to 1.7. The CSD model is unable to converge with just “ <i>greater</i> ” trial data, but it is able to converge after inverting the data. This behavior was seen in the original CSD model experiments for some participants. It should be noted that the μ and $k\sigma$ values in the pilot study are still not good estimations.	55

Chapter 1

Introduction

Virtual reality (VR) is a computer-generated experience in which the user is fully immersed in a 3D virtual environment (VE). In VR, the user sees the environment from a first-person perspective and the computer tracks the position of the user's head. Head tracking is commonly done with a *head-mounted display* (HMD) which doubles as the display through which the user observes the VE. When the user moves their head, the view of the environment changes according to that movement. Additionally, VR is an interactive experience. This means the user does not passively observe the VE, but instead the environment changes in response to the user's actions and movements.

VR differs from a regular desktop computer or mobile phone interface due to the level of *immersion* it affords. Immersion is concerned with the technical specifications that describe the system's ability to deliver an inclusive, extensive, surrounding and vivid illusion of reality [61]. Closely related to immersion is the feeling of *presence*, which is defined as the user's psychological, subjective feeling of actually being in the VE [61].

It is important to design VR systems and environments that aim to achieve maximal levels of presence within the user. Facets of immersion have been shown to influence feelings of presence. Among these facets are display refresh rate [3], environment realism and visual

quality [75], and perceptual stimuli matching¹ [22, 69]. This thesis mainly focuses on users' limits of perceptual stimuli matching when users explore VEs via real walking, and to what extent different factors influence these limits.

1.1 Walking in Virtual Reality

Travel is essential for exploring VEs. Thus, it is important to provide users with intuitive, easy-to-understand travel interfaces to enable natural and usable VR experiences. Travel interfaces that allow users to actively move are referred to as *locomotion interfaces*. Travel in VR has been supported in numerous ways, including joystick controls, omnidirectional treadmills [29], and powered shoes [31]. However, these techniques are undesirable for immersive VEs because they often lack appropriate sensory feedback or involve unwieldy hardware. It has been shown that natural walking² is the most intuitive and beneficial locomotion technique in VR, as it improves users' sense of presence [69], memory, and performance [25, 53, 56].

One common locomotion interface that enables natural walking in VR is *redirected walking* (RDW) [55]. RDW involves imperceptibly manipulating the VE via rotations and translations so that a user subconsciously adjusts their real-world position to remain on the intended virtual path. Using RDW, we can steer users away from the edges of the tracked space while still giving them the benefits of real walking in the VE. This reduces the frequency of breaks in presence³ that occur when a user reaches the bounds of the tracked space which creates a more enjoyable and effective VR experience for the user.

¹“Perceptual stimuli matching” refers to providing the user with perceptual information that matches their actions, e.g. the viewing perspective updates as the user moves their head.

²A device such as a treadmill would not be classified as “natural/real walking” because it simulates the feeling of walking.

³A break in presence occurs when “the participant stops responding to the virtual stream and instead responds to the real sensory stream” [59].

RDW relies on estimated threshold gains, which define how much the VE can be transformed without a user noticing. Previous work by Steinicke et al. [63] estimated thresholds for rotation, translation, and curvature gains; however, that study was conducted on VR hardware with a 40° field of view (FOV), which is no longer representative of modern VR systems.

Current RDW implementation has focused on imperceptibility. However, when considering usability, factors including user gender, susceptibility to simulator sickness, and gaming frequency could influence not only imperceptible threshold gains, but also usable threshold gains that do not induce simulator sickness. Aside from individual differences, threshold gains may be influenced by characteristics of the VR system including HMD FOV and tracking latency. Improving our understanding of additional factors that influence threshold gains will enable customizable redirection gains according to the user and the VR system. Gains that are more suited for a particular user will increase the effectiveness and usability of RDW.

1.2 Thesis Statement

Redirected walking thresholds are influenced by many different system and user characteristics including field of view, user gender, and distractor presence. Furthermore, different threshold calculation methods allow for accurate threshold estimation in fewer user trials.

In particular, this thesis outlines our experiments that

- Estimate rotation gain thresholds considering FOV, user gender, and distractors.
- Highlight a correlation between simulator sickness and rotation gain thresholds.
- Employ a novel statistical model to estimate thresholds with fewer experimental trials using users' confidence ratings.

Chapters 2 and 3 present background on perception and locomotion in VEs. Chapter 4 presents our user study evaluating how rotation gain thresholds change according to different factors. Chapter 5 presents our experiments implementing a novel statistical model that estimates thresholds in fewer experimental trials. Chapter 6 reviews the final results and discusses future areas of research that may deepen our understanding of redirected walking thresholds.

Chapter 2

Perception

Virtual reality experiences are highly influenced by the user's perceptual system. When in VR, the user perceives stimuli that mostly come from the VR system. How the user reacts to these stimuli is a question that is core to virtual reality research. These stimuli often conflict with stimuli the user perceives from the real world, which can cause the user to react in an unexpected manner. We can leverage this atypical behavior to force a user to act in a way that benefits the VR system designer. Indeed, this is exactly why RDW works. Because this phenomenon is so useful for creating usable VR experiences, it is crucial that we understand how human perception works.

2.1 Non-visual Perception

Non-visual stimuli include haptic, auditory, proprioceptive, and vestibular signals. Haptic perception refers to the sense of touch and physical feeling. Auditory perception refers to hearing sounds. While both these stimuli play some part in RDW, they are out of the scope of this thesis and will not be discussed further. Proprioception relates to the internal sense of movement and position of the body. It refers to the perceptual system that, for example, tells us when our hands are behind our back, or when our legs are moving. Vestibular perception

refers to the interpretation of head movement and balance. Proprioceptive and vestibular stimuli have been shown to improve users' navigation through VEs [56]. Furthermore, locomotion interfaces that support these perceptual systems have been shown to decrease the chance that a user experiences simulator sickness [12]. Thus, a locomotion interface should aim to support these perceptual systems in order to deliver a good user experience.

2.2 Visual Perception

Visual perception refers to the brain's interpretation of an environment through the eyes. It is an important part of how observers understand their surroundings. Within VR, the visual stimuli a user perceives come from the head-mounted display (HMD). The quality of the stimuli will depend on HMD factors including refresh rate, display resolution, and field of view (FOV). FOV is the observable space an observer can see through their eyes or viewing device. FOV is of particular interest to us in this thesis, since differences in FOV have been shown to influence observers' locomotion patterns. Visual perception is crucial to virtual experiences, so we will now discuss some of the important facets of visual perception and how they interact with locomotion.

2.2.1 Optical Flow

Optical flow refers to the pattern of perceived motion of the surrounding environment that is projected onto the human observer's retina. Optical flow patterns serve as a visual signal of self-motion for the human observer. Numerous studies have shown that optical flow influences the observer's locomotion control depending on the speed and direction of optical flow [4, 50, 73]. When the observer's non-visual movement signals conflict with their visual movement signals (namely optical flow), the brain prioritizes the visual signals. That is, when the observer determines their current motion, they are more likely to believe visual

information than non-visual information if the two provide conflicting cues of self-motion [5, 38].

2.2.2 Vection

Vection is the illusory impression of self-movement provided by visual stimulation [23, 65]. It is typically felt when the observer visually perceives a moving environment, but their body moves in a manner that would not produce the perceived optical flow patterns. Becausevection is most often induced by visual stimuli, it is closely tied to the perceived optical flow. A common example ofvection is the feeling of movement when an observer sits stationary in a train and watches a neighboring train move.

It is known that peripheral stimulation plays an important role in perceiving optical flow patterns [50]. Thus, we can infer that peripheral stimulation plays an important role in the degree ofvection felt in the observer. In fact, many studies have demonstrated that optical flow perceived in the periphery increases feelings ofvection [6, 27, 74]. However, it should be noted that there is evidence of feelings ofvection when foveal, and not peripheral, stimulation is present [72].

2.2.3 Simulator Sickness

Simulator sickness is the feeling of motion sickness experienced when using a simulator (or VR system). When they experiencevection, it is common for users to also experience simulator sickness. It is also possible for users to experience simulator sickness when using VR applications. Simulator sickness decreases the usability of VR and can potentially deter people from wanting to experience VR more than once. The exact cause of simulator sickness is not known, but the main theory argues that conflict between visual, proprioceptive, and vestibular stimuli is the culprit [36]. Hettinger et al. [23] strengthened this theory when they provided data suggesting that simulator sickness is a product ofvection.

It has been noted that FOV influences simulator sickness—specifically, a smaller FOV has been shown to reduce the amount of simulator sickness users experience [15, 39]. A recent study by Fernandes et al. [16] further explored how FOV influences simulator sickness. In their study, they dynamically changed the FOV in VR using what they refer to as FOV restrictors. They concluded that changing the FOV based on visually perceived motion makes users feel more comfortable during their VR experiences [16].

2.3 Gender Differences

It is important to consider gender differences in visual perception. Halpern [21] highlighted gender comparative studies that show that compared to females, males generally have better dynamic visual acuity under the age of 40. Halpern also noted that, compared to females, males tend to perform better in spatiotemporal tasks involving judgments about and responses to moving visual displays [21]. A 5-year study by Burg [11] collected data on 17,479 people (62.8% male) and found that females have a slightly wider field of view (reported FOVs differed by roughly 1° - 2°). Furthermore, it has also been noted that, in general, females are more susceptible to motion sickness than males [34]. A study by Stanney et al. [62] found that females reported higher sickness scores, but it could not be determined if this was due to anatomical or hormonal differences. The authors of that study noted that females tend to report simulator sickness symptoms more readily than males, which may contribute to the higher intensity of simulator sickness seen in females. Gender role expectations, such as males not wanting to appear weak, may also explain the differences in simulator sickness scores between genders [77].

Few studies have specifically looked at interactions between RDW and gender. A study by Bruder et al. [9] investigated differences in perceptual thresholds of RDW between genders with a 40° FOV but did not find statistically significant differences. Hildebrandt et al. [24] studied human factors that influence simulator sickness after exposure to RDW applications.

The results from that study support the claim that females are more susceptible to simulator sickness when experiencing RDW. The authors of that study concluded that human factors such as gender should be accounted for when implementing RDW. They also showed that users tolerated more simulator sickness if the VR application was exciting or practical.

Chapter 3

Virtual Locomotion

Locomotion in VR is essential for exploring VEs and delivering an interactive experience. Without locomotion, VR is more akin to passively watching a movie. A lack of support for locomotion within the VE may reduce feelings of presence and, in turn, make VR less effective [60].

Human gait features a wide range of movements like walking, running, skipping, and waddling. A good locomotion interface must support these motions, while also accounting for a variety of tracked space shapes and user dimensions. Supporting such a variety of movements is a challenge for VR systems. In this chapter, we will discuss the advantages and disadvantages of different locomotion interfaces.

3.1 Locomotion Interfaces

A locomotion interface is a device and/or software that allows a user to travel in a virtual environment. Ideally, a locomotion interface should allow the user to really walk¹ (or *perfectly* mimic the sensations felt when one really walks), be easy to understand, and require minimal extra hardware or setup. A number of different locomotion interfaces have been

¹“Really walk” in this context refers to locomoting in the tracked space using the same kind of walking we use in everyday life.

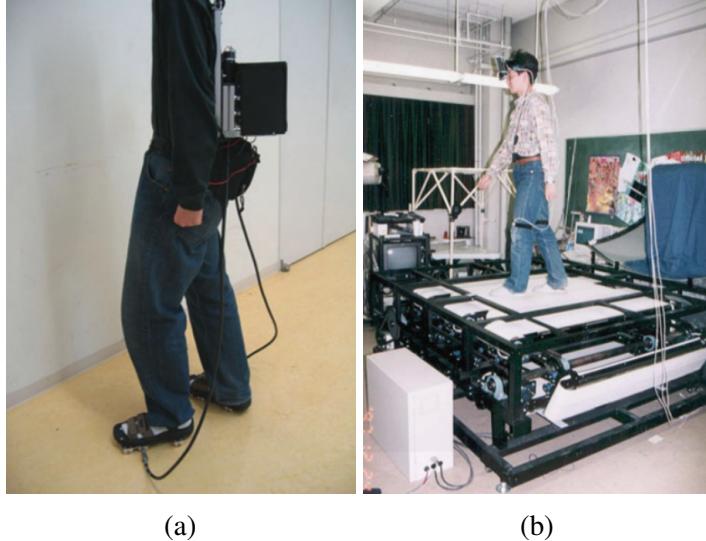


Fig. 3.1 Two locomotion interfaces, powered shoes [31] (left) and treadmills [29] (right), that do not support real walking.

proposed, prototyped, and evaluated. Some well-known interfaces include joystick controls, omnidirectional treadmills [29], powered shoes [31], moveable tiles [30], and *redirected walking* [55]. See Figure 3.1 for examples of some of these interfaces. However, many of these locomotion techniques are undesirable because they do not meet all the criteria of an ideal locomotion interface. Suboptimal locomotion interfaces are usually unsatisfactory because they involve unwieldy hardware or lack vestibular or proprioceptive feedback that is present during real walking, e.g. a treadmill.

Of the locomotion interfaces that have been studied, interfaces that utilize *redirection techniques* (RTs) appear to be the most promising because they allow users to really walk. RTs allow users to explore VEs that are larger than the tracked workspace by manipulating the user's path in the virtual environment [44]. It has been shown that natural walking is the most intuitive and beneficial locomotion technique in VR, as it improves users' sense of presence [69], memory, and performance [25, 53, 56]. As a result of the numerous benefits real walking offers, researchers have invested considerable effort into developing and understanding locomotion interfaces that support real walking.

3.2 Real Walking in Virtual Reality

Standard VR systems do allow users to walk around during a virtual experience, but users are only able to walk within the tracked space. Movement outside the workspace borders will not be tracked by the system's sensors, so the visual scene displayed on the HMD will not update according to the user's movements. Thus, the size of the VE that a user can explore is limited to the size of the tracked space. See Figure 3.2 for a diagram of a typical VR system tracked space.



Fig. 3.2 The blue rectangle on the floor represents the tracked space. The user can walk freely within this blue rectangle and the VR system will be able to track their movements. If they step outside this region, the VR system will not be able to track the user's movements. The lighthouses (circled in red) are responsible for tracking the position of the HMD that the user wears. Original image via [14].

To support real walking and increase the size of the explorable VE, we can employ RTs. A multitude of redirection techniques have been developed [8, 28, 55, 67], which has prompted

researchers to classify RTs based on their implementation-specific characteristics. Suma et al. distinguished between redirection techniques based on the conspicuousness (overt or subtle) and continuity (discrete or continuous) of their implementations [66]. Subtle and continuous techniques are preferred because they have been reported to create fewer breaks in presence. However, depending on the user’s projected path and position in the workspace, we cannot always rely on such techniques to keep users in the tracked workspace. In these situations, redirection systems may sometimes be required to fall back on more overt techniques to ensure the user’s safety [44, 66].

3.2.1 Redirected Walking

One popular subtle and continuous RT that enables natural walking in VR is *redirected walking* (RDW) [55]. RDW involves imperceptibly manipulating the VE via rotations and translations so that a user subconsciously adjusts their real-world position to remain on their intended virtual path. Using this technique, we can steer users away from the tracked-space edges while still giving users the benefits of real walking in the VE. This reduces the amount of breaks in presence caused by reaching the bounds of the tracked space.

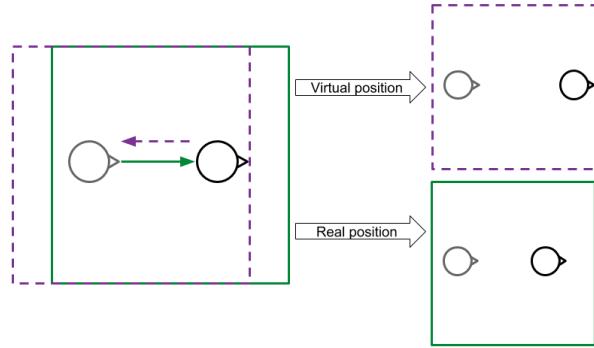
For example, a user will physically rotate by 180° when he or she wants to turn 180° in the VE, if no redirection is applied. If redirection is applied such that some real-world rotation results in a larger rotation in the VE, the user will turn until their position in the VE has rotated 180° , but the physical rotation will be less than 180° . We can also redirect such that a physical rotation results in a smaller virtual rotation. When implemented carefully, this discrepancy between the physical and virtual movements is imperceptible to the user if it is small enough. Similar transformations can be applied to a user’s walking path. When walking on a straight path, we can translate the VE in the direction opposite to the user’s walking direction, which results in a virtual displacement that is larger than the user’s physical displacement. We can also rotate the VE while the user walks to force the user to follow a

curved path in the real world. Depending on the strength and direction of the rotation, this will force the user's real path to steer away from the edges of the tracked space. See Figure 3.3 for a diagram that explains how RDW manipulates the VE. This thesis is only concerned with rotations of the VE when the user is standing in place.

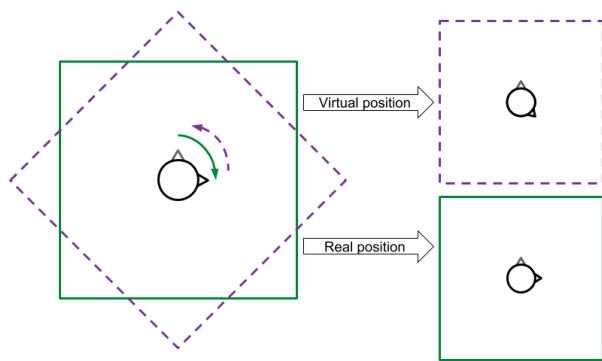
Limits of Redirection

By applying RDW, users are able to walk naturally and explore VEs larger than the tracked workspace. However, we cannot simply amplify users' movements by a large, constant factor to maximize the size of the explorable VE without incurring negative repercussions such as disorientation or increased simulator sickness. The scaling of a user's movements must be small enough to maintain the VR application's usability and ensure the user's comfort. Thus, there exists a trade-off between redirection intensity and user experience [55]. Ideally, enough redirection is applied to maximize the explorable size of the VE and minimize discomfort and breaks in presence caused by manipulating the VE.

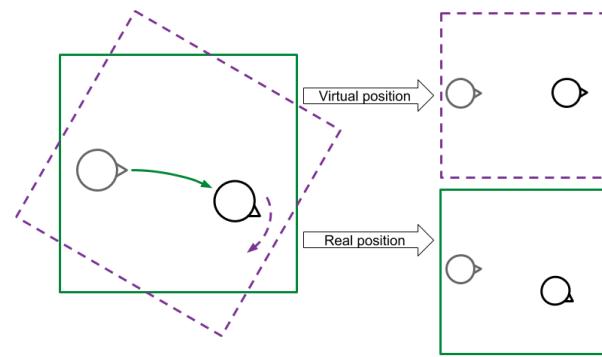
The intensity of scaling applied to the VE is controlled by parameters called *gains*. Rotation gains increase or decrease a user's rotation in the VE relative to their real-world rotation, while translation gains increase or decrease a user's displacement in the VE relative to their real-world displacement. Curvature gains, on the other hand, cause users to walk along a curved physical path while walking on a straight virtual path. Both rotation and translation gains are expressed as a ratio of virtual motion to physical motion. A gain of 1 is applied when virtual motion to physical motion is mapped 1:1. When a gain is greater than 1, the virtual movement (rotation or translation) is increased, and the resulting real-world movement is smaller than the virtual movement. Similarly, when a gain is less than 1, the virtual movement is decreased, and the resulting real-world movement is larger than the virtual movement. A *threshold* refers to the point at which the applied gain becomes noticeable to the user, and each threshold has an associated gain. A threshold t corresponds



(a) A translation gain allows the user to walk distances in the VE that are greater than the distance walked in the real world.



(b) A rotation gain allows the user to turn a greater virtual distance compared to their physical rotation.



(c) A curvature gain forces the user to walk on a curved physical path in order to walk in a straight path in the VE.

Fig. 3.3 Diagrams that illustrate how different RDW gains can be used to increase the size of the explorable VE. The green borders represent the real-world tracked space borders, and the purple borders represent the borders of the VE that correspond to the size of the tracked space. Arrows indicate the user (green) or VE (purple) movement.

to a gain g . A t threshold of g means that $t\%$ of the population will believe that their virtual movements are larger than their physical movements when the gain g is applied. For example, if the 50% threshold has a gain of 1.02, then half the population will believe that their physical and virtual movements are the same when we apply a gain of 1.02 while the other half will believe that their virtual movements are larger than their physical movements. In previous work by Steinicke et al. the threshold values of interest are users' 25% and 75% thresholds, which correspond to *decreased* and *increased* virtual rotations respectively [63].

VE rotation is often discussed in relation to the user's physical rotation. VE rotation *with* the user's physical rotation direction corresponds to a real-world rotation that is larger than the virtual rotation, and VE rotation *against* the user's physical rotation direction corresponds to a real-world rotation that is smaller than the virtual rotation.

Estimation of Detection Thresholds

Many studies have estimated threshold gains in VR [7, 10, 20, 32, 33, 45, 48, 64]. The most comprehensive study was conducted by Steinicke et al. [63], which estimated threshold gains for rotation, translation, and curvature gains. However, that study was conducted on VR hardware with a 40° field of view (FOV), which is no longer representative of modern VR systems. Since the present study only focuses on rotation gains, we will limit the discussion to previous work related to estimated rotation gains. See Langbehn et al. for a full review of redirection thresholds [37].

As defined by Steinicke et al. [63], a rotation gain $g_R = \frac{R_{virtual}}{R_{real}}$ where $R_{virtual}$ is the virtual-world rotation and R_{real} is the real-world rotation. Steinicke et al. reported 25% and 75% rotation threshold gains at 0.67 and 1.24 respectively [63]. Since then, others have replicated or conducted studies similar to [63] and reported different gains. Bruder et al. [7] reported very similar gains at 0.68 and 1.26, while Paludan et al. [48] reported gains at 0.93 and 1.27. Nilsson et al. [45] estimated threshold gains to be at 0.77 and 1.1. Additionally, Jerald et al.

[33], who studied perceptual thresholds during head rotations, reported that scenes can be rotated up to 11.2% with the direction of the user’s rotation and 5.2% against the direction of the user’s rotation (threshold gains estimated in [63] correspond to 49% rotation with and 20% against the user’s rotation direction).

Rotation threshold gains have also been studied in different experimental conditions, again yielding different values. In addition to replicating [63], Nilsson et al. [45] studied threshold gains in the presence of static and moving audio and found values at 0.8 - 1.11 and 0.79 - 1.08 respectively. In work by Bruder et al. [7], threshold gains were evaluated for users traveling in electric wheelchairs. The gains in that study were reported at 0.77 and 1.26. Serafin et al. [58] conducted a study that evaluated threshold gains using only auditory stimuli, and reported values at 0.82 and 1.2.

Incorporating Distractors

VEs often include components that are not directly related to the purpose of the VR application. For example, an outdoor VE may include butterflies even though the application has nothing to do with butterflies. These extra components of the VE can momentarily capture the user’s attention as they complete some virtual task.

These extra components are known as distractors—objects or sounds (or a combination of both) in the VE that aim to capture the user’s focus to allow larger redirection amounts to be applied without the user noticing [51]. Distractors are an overt redirection technique [66]. When a user is distracted by something in the environment, we can leverage the user’s focus on a single object to induce stronger redirection that would normally be noticed if the user was attending to the environment’s movements. The effect of distractors on users’ navigational ability and awareness of RDW has been previously studied [13, 52, 53]. Results from these studies indicate that, in general, redirection using distractors is effective, and users perform tasks at least no worse than when redirected by RDW without distractors. Peck et al.

[51] demonstrated that rotation gains can be increased while users are distracted, but we are currently unaware of any studies that formally estimate threshold gains with distractors.

Chapter 4

Estimation of Rotation Gain Thresholds

This section presents our experiment design and results.

4.1 Equipment

We used an HTC Vive Pro virtual reality headset with 6 degrees of freedom (6DoF)¹ position and orientation tracking in a $5m \times 4.2m$ tracking space. The system had about 110° diagonal FOV, a 90Hz refresh rate, and a 1440×1600 resolution per eye. The experiments were run on the Unity 2018.1.6f1 engine (with the SteamVR library) on a computer with an Intel i7-7820X processor (3.6 GHz), 32GB RAM, and NVIDIA GeForce GTX 1080 Ti GPU running on Windows 10 Pro edition. The experiments ran at 90 frames per second.

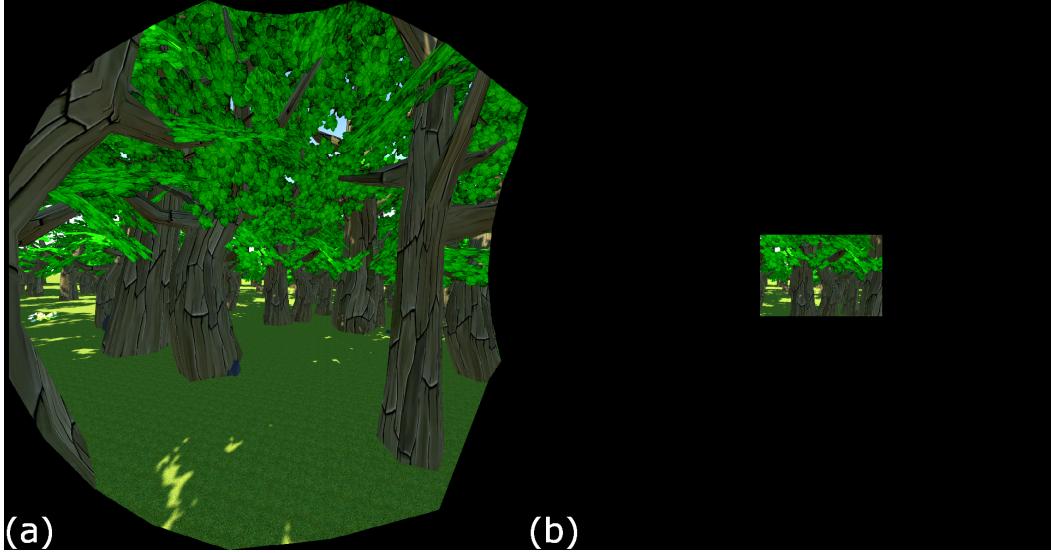


Fig. 4.1 (a) View of the VE in the left eye with no FOV modification (110° FOV). (b) View of the VE in the left eye with a 40° FOV restrictor.

4.2 Experiment Design

We limited our study to rotation gains since these gains enable larger redirection amounts compared to translation and curvature gains [63]. Our experiment tested the following hypotheses:

H1 Participant discrimination between rotation gains is different in a 110° FOV compared to a 40° FOV.

H2 Participant discrimination between rotation gains is different for females compared to males.

H3 Participant discrimination between rotation gains is different when distractors are present compared to when they are not present.

At the end of the experiment, participants filled out a demographics questionnaire including age, gender, colorblind status, and VR and video game experience.

¹The headset's yaw, pitch, roll, and movement along the x , y , and z axes are tracked by the VR system. Yaw, pitch and roll refer to head movements about the axes through a person that extend from head to toe, ear to ear, and chest to back respectively.

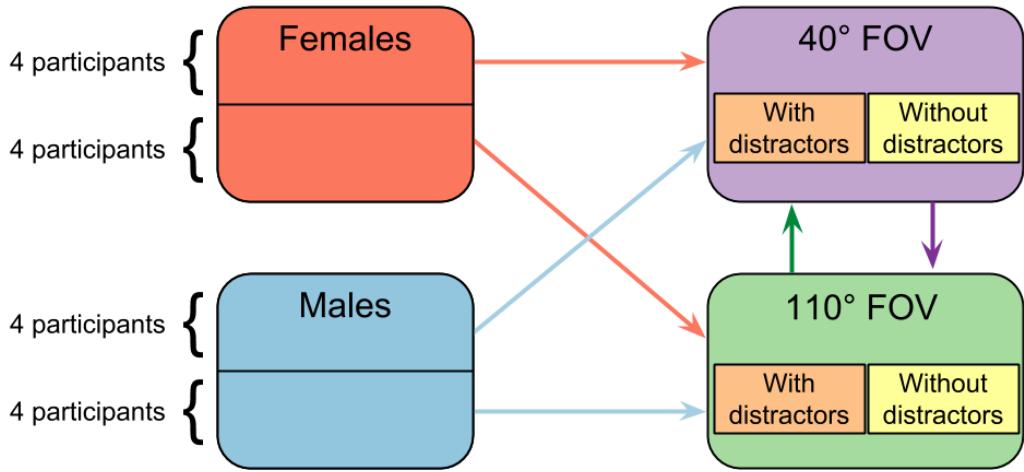


Fig. 4.2 Breakdown of the 2 (FOV: 40° , 110°) $\times 2$ (Gender: female, male) $\times 2$ (Distractor: without, with) study we conducted. Each box is one of the 2 values of each of the independent variables (FOV, gender, distractor status) in our study.

We conducted a 2 (FOV: 40° , 110°) $\times 2$ (Gender: female, male) $\times 2$ (Distractor: without, with) user study with FOV and Distractor as within-participant variables, and Gender as a between-participant variable. See Figure 4.2 for a breakdown of the experiment design. The experiment was broken into two blocks counterbalanced by FOV. Half of our participants experienced the 40° FOV block first, and then experienced the 110° FOV block. The other half of the participants experienced the 110° FOV block and then 40° FOV block. That is, we divided the list of participants into two groups, and one group started with a 40° FOV and the other with a 110° FOV. We then shuffled the list so the order in which we tested participants was random. This is done to account for order effect that may potentially confound our results. We used the Vive Pro default diagonal FOV of 110° for the 110° block, and implemented an FOV restrictor similar to Fernandes et al. [16] to create the 40° FOV viewport used in the 40° FOV block. To more closely replicate Steinicke et al. [63], our FOV restrictor had a rectangular, hard-edge border instead of the circular, soft-edge border used by Fernandes et al. [16]. A comparison of the FOVs is shown in Figure 4.1.



Fig. 4.3 Sample view of the scene and the corridor of trees during a trial.

Each block consisted of 144 trials with the same FOV. Within each block, half of the trials (randomly distributed) featured a distractor to estimate threshold gains with a distractor present. Each trial had one of three distractor conditions: no distractor present (replication of [63]), distractor present and moves in the same direction that the user turns, or distractor present and moves in the opposite direction that the user turns. The entire experiment structure for one participant is shown in Figure 4.4. We used a running deer as the distractor which can be seen in Figure 4.5. The deer moved along a 180° arc around the participant at a speed of about 6.2 m/s.

In each FOV block, each gain was tested 8 times without distractors and 8 times with distractors. For the 8 trials with distractors, 4 distractors moved with the user, and 4 moved against the user. Excluding practice trials, this totaled 2 FOVs × 2 Distractors × 9 gains × 8 times totaling 288 trials per participant. The trial order per block was randomized for each participant.

The gains we applied ranged from 0.6 (150° physical rotation resulted in a 90° virtual rotation) to 1.4 (64.3° physical rotation resulted in a 90° virtual rotation), incremented in

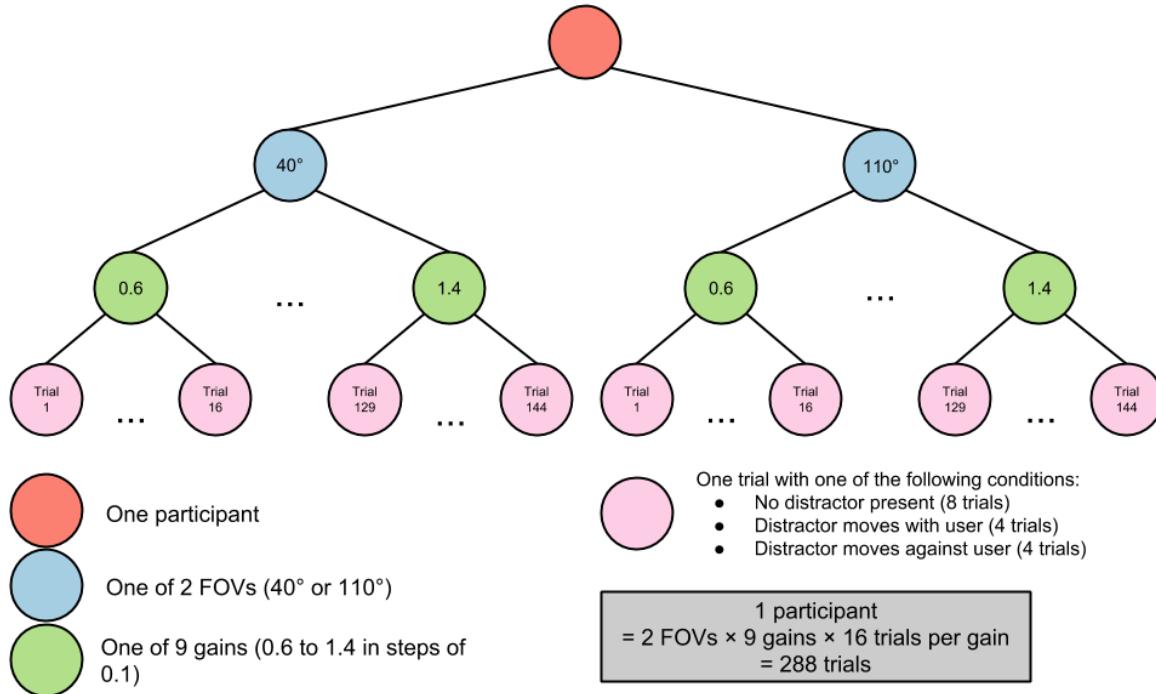


Fig. 4.4 Experiment breakdown for one participant. Each subtree of the blue nodes represents one block.

steps of 0.1. The VE rotated based on participant movement about their yaw axis. Note that, unlike Steinicke et al. [63], we did not test gains of 0.5 and 1.5. These gains were removed based on the gains reported in Steinicke et al. [63] and to keep the experiment duration under two hours. The original experiment by Steinicke et al. [63] tested all three gains — rotation, curvature, and translation — in one three-hour experiment.

The task trials were a constant stimuli two-alternative forced choice (2AFC) task. In the method of constant stimuli, stimulus intensities above, below, and at the threshold intensity are presented to the user in a random order. With a 2AFC task, the participant is given two alternative options. Except for the case when no stimulus is applied, one of these options correctly describes the stimulus. The participant is forced to select one of these two options for each completed trial. The two options given to our participants were “*greater*” and “*smaller*,” in response to the question “Was the virtual movement *smaller* or *greater* than the physical movement?” As explained in [63], 2AFC tasks avoid participant response bias as

participants are forced to guess even when they are unsure of VE rotation magnitude. In our experiment, when no stimulus is applied (physical and virtual movements are 1:1), the gain applied is 1. In this situation, participants should have no insight into which answer is correct (*neither* is correct), and they will be forced to guess and will be correct 50% of the time on average. Each participant took about two hours to complete the entire experiment, including preliminaries, debriefing, trials, breaks, and questionnaires.

4.3 Virtual Environment

The virtual scene was designed to emulate optic flow conditions for both indoor and outdoor environments. The experiment scene was a virtual outdoor forest with trees, flowers and rocks and participants were able to see the horizon through the trees. The user was positioned in a corridor-like clearing in the forest where trees were positioned such that the corridor was at least 1 meter wide. The corridor width was chosen to emulate average hallway width [47, 70]. Ambient background sounds were played through user worn headphones to mask real world sounds that could provide positional cues to the user. See Figure 4.3 for an image of the virtual scene used in the experiment. See Figure 4.5 for an image of the distractor and a participant during the experiment.

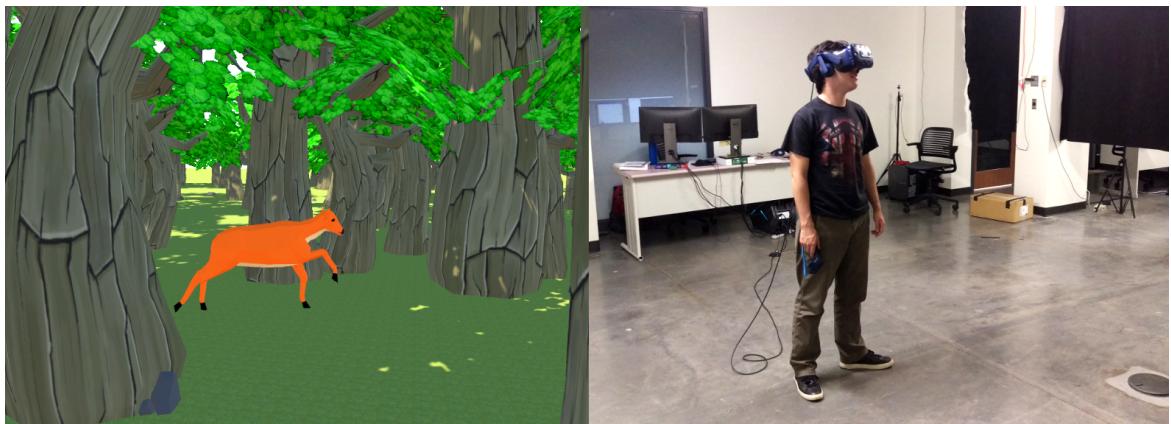


Fig. 4.5 Screen shot of the virtual forest scene with a distractor moving across the user's field of view (left) and a participant wearing the headset (right).

4.4 Participants

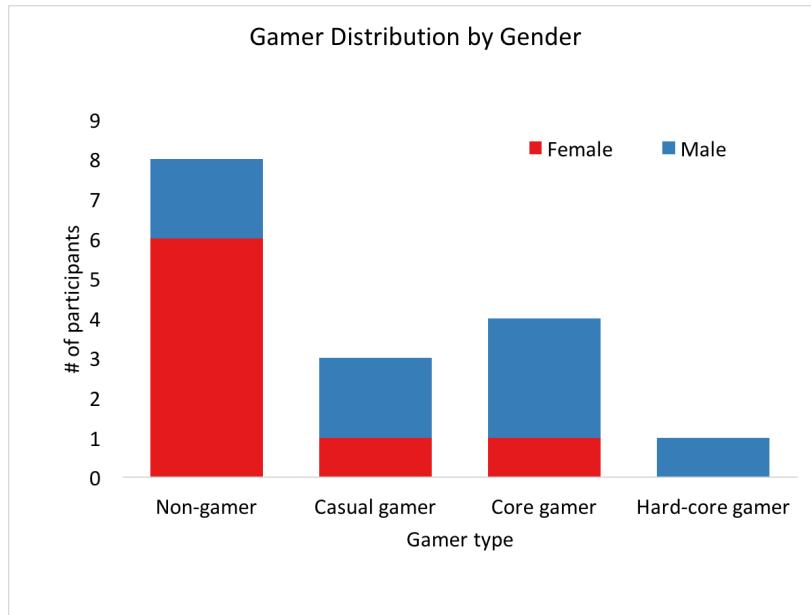


Fig. 4.6 The gamer-type distribution by gender for the experiment participants. Female participants are in red, and males are in blue.

All participants were at least 18 years of age, had normal or corrected to normal vision, and were not knowingly pregnant. Additionally, all participants had normal or corrected to normal hearing and had no history of epilepsy, seizures, or strong susceptibility to motion sickness. Participants were proficient in written and spoken English.

Participants included students, faculty, and staff of Davidson College. Nineteen people participated in the experiment. Two participants were not naïve to the purpose of the study including one of the study designers. One participant's session was terminated early due to technical difficulties. An additional two participants' data were discarded as these participants appeared to have misunderstood the experiment task including one participant who replied "greater" to all but one trial.

Sixteen participants, age 19 - 48 (8 female ($M = 22$, $SD = 5$) and 8 male ($M = 26$, $SD = 11$)) successfully completed the experiment. One participant chose not to disclose his age. The HMD was adjusted to each participant's interpupillary distance (IPD), except for

six participants (4 female) whose IPD was below the minimum setting of the HMD. In these cases, the HMD was set to the minimum IPD of *61.3mm*, and participants did not report display blurriness when asked.

Three participants had had multiple VR experiences before, nine had briefly experienced VR before, and four had never experienced VR before. Five users self-reported themselves as either core or hard-core gamers (high experience), and eleven reported themselves as non- or casual gamers (low experience). See Figure 4.6 for a breakdown of the game experience distribution among participants.

4.5 Procedure

Upon arriving at the lab, participants completed a participant eligibility checklist and consent form. Participants were offered compensation in the form of a \$10 gift card. The experiment and procedure were approved by the Davidson College institutional review board (Protocol #: 2018-067). Each participant's IPD was measured and the HMD was adjusted to match the measurement as closely as possible. Participants were briefed on the experiment task and were asked to repeat the task procedure to ensure comprehension.

Participants started in either the 40° or 110° FOV block with each block consisting of 144 trials. A trial consisted of rotating the whole body in place, not just the head, in the direction of an arrow at the center of the participant's vision. To encourage participants to focus on the VE, the direction arrow disappeared after the user began turning in the correct direction. The VE rotation was 90° , randomly ordered with half clockwise and half counter-clockwise rotations.

Participants rotated until a blue dot appeared at the center of their vision, signaling that they should maintain their position until they heard a bell sound indicating successful trial completion. If the participant rotated past the 90° virtual rotation, the blue dot's color changed to red and participants had to correct and maintain their orientation such that the

dot changed to blue again. This orientation maintenance requirement was implemented to prohibit rapid turning and overshooting the 90° virtual rotation and to encourage participants to turn slowly enough to avoid overshooting the target orientation. At the beginning of each block, participants completed three practice trials with a random gain applied (the same gain for all three practice trials) to familiarize themselves with the task and environment. Users had an opportunity after the practice trials to ask questions or adjust the HMD if needed. Response accuracy feedback was not given on any trial.

Once a trial was completed, the HMD display faded to black and post-trial questions were displayed. Participants answered questions with the Vive Pro controller by using the left and right directions on the trackpad to select an answer from options displayed below the question. Answers were submitted using the controller's trigger button. Participants were first asked the same question as that used in [63]: “Was the virtual movement *smaller* or *greater* than the physical movement?” (smaller, greater). Participants were also asked, regardless of distractor presence, “Did you see an animal in the scene?” (yes, no) to determine if participants saw the distractor.

Before the next trial began, participants were reoriented to the starting orientation of their most recently completed trial. This was done to prevent participants from getting tangled in the HMD cable. To accomplish this reorientation, participants rotated in the direction indicated by an arrow on a black screen until a red dot at the center of their vision turned blue. Like the trials, participants had to maintain their orientation when the dot turned blue to proceed to the next trial.

As the experiment progressed participants slowly strayed away from their physical starting position. After finishing a trial, if a participant had strayed too close to the edges of the tracking space, the experiment was paused and the experimenter guided the participant back to the center of the tracking area. This repositioning prevented the HTC Vive chaperone system from appearing in the VE, and the HMD wire from becoming taut. One participant

mentioned feeling the taut wire during a portion of the experiment. Eight participants were repositioned at least once in their experiment session. Those who were repositioned were repositioned on average 1.72 times, and no participant was repositioned more than 3 times per block.

Participants were allowed to take a break after any trial. If the user wanted a break, they communicated this verbally with the experimenter and the break started after completing the reorientation process. Nine participants took at least one break during their session. Three of these participants took two breaks. All breaks lasted no more than 2 minutes, except for two breaks (for two different participants) which lasted about 7.5 minutes each.

After completing all trials for the first FOV block, all participants completed the Kennedy-Lane Simulator Sickness Questionnaire (SSQ) [35] and had a mandatory break of at least five minutes. After completing the second FOV block, another SSQ was completed, as well as a questionnaire about their demographics (age, gender, colorblind status, amount of experience playing video games, and amount of previous exposure to VR).

4.6 Results

The probability, $\psi(greater|g_i)$, of responding “*greater*” at gain g_i to the question, “Was the virtual movement *smaller* or *greater* than the physical movement?” was calculated for each participant, for each gain both with and without distractors. No significant effect of clockwise versus counter-clockwise rotations was found, and the rotation direction data were pooled for analysis. We fit a Gaussian cumulative distribution function (CDF) to participants’ data using maximum likelihood estimation. This is known as a psychometric curve and is shown Figure 4.7. The point of subjective equality (PSE), σ , 25% and 75% threshold gains, and deviance were calculated for each participant. The PSE for a participant is the gain for which he or she replies “*greater*” to half of the trials. Therefore, at the PSE, the participant has a 50% chance to reply “*greater*.⁷ This can be interpreted as being the point at which the user

has no insight into the direction of the VE rotation, and is forced to guess (because we used a 2AFC design).

The psychometric function in Equation 4.1 was fit to our data using the Quickpsy package in R.

$$\psi(greater|g_i; \theta) = \psi(greater|g_i; \mu, \sigma, \gamma, \lambda) = \gamma + (1 - \gamma - \lambda)\Phi(greater|g_i; \mu, \sigma) \quad (4.1)$$

Thus, $\psi(greater|g_i; \theta)$ is the probability assigned by our model to a user responding “greater” when presented with the gain g_i . The vector $\theta = (\mu, \sigma, \gamma, \lambda)$, where μ and σ are the mean and standard deviation of the fitted function, and γ and λ represent the left and right asymptotes of ψ , and Φ is the Gaussian CDF. When the gain $g_i = 1$ this implies that virtual rotation is neither increased nor decreased. Full details on the fitting procedure of $\psi(greater|g_i; \theta)$ can be found in [40].

For each experimental condition, participants with data that could not be fit to a psychometric curve with a PSE centered within ± 100 from 1,² or with a probability of fit less-than .05 were considered to have a bad fit and were removed from the analysis. See Table 4.1 for each participant’s PSE and removed data.

To determine if our data replicated the results presented in Steinicke et al. [63], an estimated Bayes factor was calculated comparing the PSEs from our experiment (40° FOV, without distractors condition, pooled participant data) to the PSEs presented in Steinicke et al. [63]. We used Bayesian Information Criteria to compare the fit of the data under the null hypothesis (that there is no statistical significance between the two experiments) to the fit of the data under the alternative hypothesis (that there is a statistical significance between the two experiments) [71]. The estimated Bayes factor is 0.76. The data was only 1.31 times more likely to occur in our experiment than in the experiment from Steinicke et al.

²By using this cutoff, participants with completely unusable fits were removed but participants with less severe but still bad fits (such as participant 2) were not removed.

ID	No Distractors		Distractors	
	40°		110°	
	PSE	PSE	PSE	PSE
1	1.0071	1.0453	1.1315	1.0215
2	0.5560	-	-	-0.6501
3	0.9818	1.0375	1.1315	1.1893
4	1.0237	1.2419	1.0143	1.3850
5	-	0.4711	-	0.7330
7	0.9393	-	0.3522	-0.0458
8	0.9508	0.7485	1.0117	0.5300
11	0.7961	0.9586	0.7152	0.3361
12	-	1.3403	-	1.6478
13	0.8911	1.3779	1.0652	1.6428
14	0.7762	1.0947	0.7175	0.9403
15	1.0859	1.2343	1.1324	1.2473
16	1.0105	0.9526	1.0041	1.1778
17	0.5642	-	0.5390	-
18	0.9873	1.2513	0.9034	0.8571
19	1.0236	0.9807	1.0037	0.9494
μ	.8995	1.0565	.9015	.8641

Table 4.1 The PSE of the psychometric curve fit to each participant's $\psi(\text{greater}|g_i)$. The - indicates participant data that could not be fit to a psychometric curve and was excluded from analysis.

[63]. That is, there is weak evidence in support of the alternative hypothesis [54] and, thus, our data successfully replicated previous work. Additionally, using maximum likelihood estimation, a psychometric curve was fit to the pooled results of our participant $\psi(\text{greater}|g_i)$ data, ($\mu = .95$, $\sigma = .40$, deviance= 4.70, $p = .89$) and the gains at the 25%, PSE, and 75% thresholds were calculated (see Table 4.2. All calculated gains were within .03 of the original gains calculated in Steinicke et al. [63].

Psychometric curves were also fit to the pooled results of participant $\psi(\text{greater}|g_i)$ data by gender, (Female: $\mu = .9286$, $\sigma = .5253$, deviance= 7.3365, $p = .85$), (Male: $\mu = .9586$, $\sigma = .3267$, deviance= 8.2436, $p = .75$), and the gains at the 25%, PSE, and 75% thresholds were calculated. Comparison of gains at the 25%, PSE, and 75% detection thresholds to the gains presented by Bruder et al. [9] evaluating gender reveal that our gains are similar and

within .02 at the 75% threshold and within .09 at the 25% threshold. Individual participant data was not provided in [9], and additional analysis could not be performed.

To determine if there were gaming experience differences between detection threshold gains, we grouped self reported non- and casual gamers (low experience) together and core and hard-core gamers (high experience) together, and analyzed $\psi(greater|g_i)$ with game experience (low to high) as a between participant variable. No significant game experience or game experience interactions were found regardless of simulator sickness, nor were game experience or game experience interactions found between gains. As a result, the rest of the results we present in this section will not consider gaming experience. In the following sections, we will first present the results of our analyses and then discuss their meanings in Section 4.7.

4.6.1 Analysis of $\psi(greater|g_i)$

To test our hypotheses (see section 4.2) we compared the raw $\psi(greater|g_i)$ data, the probability of responding “*greater*” at gain g_i , with a 2 (Distractors: present, not present) \times 9 (Gain: 0.6:1.4:0.1) \times 2 (FOV: 40°, 110°) \times 2 (Gender: female, male) ANOVA with Distractor, Gain, and FOV as within-participant variables and Gender as a between participant variable. Mauchley’s test indicated that the assumptions of sphericity had not been violated, which is a prerequisite for performing an ANOVA. See Figure 4.7.

There was a significant main effect of Gain, $F(8, 72) = 46.98, p < .0001, \eta^2 = .52$. There was also a significant Gender \times FOV interaction, $F(1, 9) = 8.97, p = .02, \eta^2 = .06$, and a Gain \times FOV interaction, $F(8, 72) = 2.62, p = .01, \eta^2 = .01$. Finally, there was a significant 3-way Gender \times Gain \times FOV interaction, $F(8, 72) = 2.96, p = .0064, \eta^2 = .05$, and a trending 3-way Gender \times Gain \times Distractor interaction, $F(8, 72) = 2.00, p = .0581, \eta^2 = .02$. The Gender \times Gain \times FOV interaction indicates that the 2-way Gain \times FOV interaction is different between male and female participants.

4.6.2 Gain analysis

To breakdown the interactions found in the raw $\psi(greater|g_i)$ data, each participant's gains at the 25%, PSE, and 75% thresholds were estimated from their fitted psychometric curve. These thresholds were chosen based on previous threshold estimation. To determine if there were differences in gains at the 25%, PSE, and 75% thresholds we performed a 2 (Distractors: present, not present) \times 3 (Threshold: 25%, PSE, 75%) \times 2 (FOV: 40°, 110°) \times 2 (Gender: female, male) ANOVA with Distractor, Threshold, and FOV as within-participant variables and Gender as a between-participant variable. Mauchly's test [42] was used to verify that sphericity was not violated. See Table 4.3 and Figure 4.7.

Significant main effects of Gender, $F(1, 9) = 5.22, p = .0482, \eta^2 = .05$, and Threshold, $F(2, 18) = 17.11, p < .0001, \eta^2 = .39$ were found. A significant Threshold \times FOV interaction was found, $F(2, 18) = 4.28, p = .0302, \eta^2 = .09$. Post-hoc analysis of pairwise comparisons was performed using estimated marginal means with Bonferroni adjustments. Results support significant differences at the 25% threshold, $t(22.85) = 2.10, p = .0473$, between the 40° FOV ($M = .73, SE = .12$) and the 110° FOV ($M = .44, SE = .12$), and at the 75% threshold, $t(22.85) = -2.62, p = .0152$, between the 40° FOV ($M = 1.25, SE = .12$) and the 110° FOV ($M = 1.62, SE = .12$). This supports H1, that participant discrimination between rotation gains is different in a 110° FOV compared to a 40° FOV. Rotation gain thresholds are significantly wider when using a 110° FOV compared to a 40° FOV.

Additionally, a significant Gender \times FOV interaction was found, $F(1, 9) = 11.19, p = .0086, \eta^2 = .04$. Post-hoc analysis found a significant difference in gains between male and female participants at the 110° FOV, $t(14.66) = -3.67, p = .0023$, and a significant difference between the 40° FOV and 110° FOV for males, $t(9) = -3.37, p = .0082$. This supports H2, that participant discrimination between rotation gains is different for females compared to males.

4.6.3 Distractors

Gain analysis from subsection 4.6.2 revealed four trending interactions involving distractors. See Table 4.3. Exploratory post-hoc analysis of the highest-order trending effect using estimated marginal means with Bonferroni adjustments, comparing gender and presence of distractor, pairwise, at each threshold and FOV was performed. Two significant comparisons were found in the 110° FOV at the 25% threshold. There was a significant difference in gains for females in the presence of distractors ($M = -.0822$, $SE = .2108$) versus no distractors present ($M = .6305$, $SE = .2108$), $t(48.87) = 3.488$, $p = .0062$. Additionally, there was a significant difference in gains in the presence of distractors between males ($M = .7120$, $SE = .1640$) and females, $t(70.61) = -2.968$, $p = .0245$. This suggests that distractors may affect males and females differently, however only in the 110° FOV at the 25% threshold. This weakly supports H3, that participant discrimination between rotation gains is different when distractors are present compared to when they are not, though further studies should be performed to investigate this result.

4.6.4 Simulator sickness

Simulator sickness scores for each participant can be seen in Figure 4.8a. The simulator sickness data including nausea, oculomotor, disorientation, and total scores were independently evaluated in a 2 (SSQ order: first, second) \times 2 (FOV order: 40° first, 110° first) \times 2 (Gender: Female, Male) ANOVA with SSQ order as a within-participant variable and FOV order and Gender as between-participant variables. No significant effects or interactions were found. Additionally, six participants had IPDs smaller than the HTC Vive IPD and completed the experiment with too large an IPD ($M = 2mm$ too large). Note that Steinicke et al. [63] did not adjust for IPD. The averaged simulator sickness scores over the two blocks for each participant were calculated, and the simulator sickness scores of participants wearing incorrect IPDs ($M = 44.88$, $SE = 6.18$) were compared to participants wearing correct IPDs

($M = 36.84$, $SE = 7.46$). See [35] for a reference on how to interpret these scores. No significant difference between groups was found, $t(13.85) = 0.83$, $p = 0.42$. Additionally, the estimated Bayes factor = 0.34 provides weak support of a difference in simulator sickness scores between participants wearing an HMD with the correct IPD compared to wearing an HMD with a slightly too large IPD.

Exploratory analysis was performed to determine if threshold gains were correlated with simulator sickness scores. Robust Winsorized correlations with 10% trimmed means were calculated comparing gains at the 25%, PSE, and 75% thresholds for the 40° FOV and the 110° FOV, with the corresponding FOV block simulator sickness score. See Table 4.4. At the 40° FOV significant negative correlations were found between participant simulator sickness scores at both the 25% and PSE threshold gains, $\rho_w = -.56$, $p = .04$ and $\rho_w = -.64$, $p = .02$ respectively. The significant correlation shows that at the 25% threshold, participants with a greater decrease in rotation gains had higher simulator sickness scores. Additionally, participants whose PSE was further decreased from 1 had higher simulator sickness scores. At the 110° FOV, a significant positive correlation was found at the 75% threshold, $\rho_w = .64$, $p = .01$. When removing the male outlier (see Figure 4.8b) the correlation strengthened to $\rho_w = .74$, $p = .003$. Participants with a greater increase in rotation gains had higher simulator sickness scores.

Using confidence interval tests for correlation coefficients, we evaluated the equality of correlation coefficients for males and females at each threshold and FOV. No significant differences were found between genders.

4.7 Discussion

We tested our hypotheses (see Section 4.2) with a 2 (FOV: 40° , 110°) \times 2 (Gender: female, male) \times 2 (Distractor: without, with) user study with FOV and Distractor as within-participant variables, and Gender as a between-participant variable. We successfully repli-

cated previous threshold estimations for rotation gains at the 40° FOV, and compared results to threshold gain estimations at the 110° FOV. Our results strongly supported H1 in that rotation gains are wider in a 110° FOV compared to a 40° FOV. Gains at the 25% and 75% thresholds were significantly different in the 40° FOV compared to the 110° FOV. Additionally, our results supported H2 in that females and males have different threshold gains in the 110° FOV. Finally, H3 was partially supported. Males did not have a significant difference in threshold gains comparing the presence of distractors, however females did have a significant difference in threshold gains in the presence of distractors at the 25% threshold in the 110° FOV.

In a 40° FOV participants were unable to discriminate between 90° virtual rotations and real rotations ranging from 73° to 132°. In a 110° FOV participants were unable to discriminate between 90° virtual rotations and real rotations ranging from 60° to 145°, equating to a 33% decrease and 61% increase in rotations. This supports H1, that participant discrimination between rotation gains is different in a 110° FOV compared to a 40° FOV. Rotation gain thresholds are significantly wider when using a 110° FOV compared to a 40° FOV.

The difference in threshold gains between FOVs may be attributed to the increased visual information received with a 110° FOV viewport. Human perception literature makes two important observations about the effects of an increased FOV. First, it is noted that peripheral stimulation afforded by a wider FOV increases the observer's feelings of self-motion. Based on that observation, it is possible that participants' perceived self-motion increased between a 40° FOV viewport and a 110° FOV viewport. It is possible that this different sense of self-motion altered their sensitivity to the gains. However, no effects of FOV block order were found.

Second, the literature notes that when they conflict, visual information more strongly influences locomotion than vestibular and proprioceptive information do. To successfully

complete the experiment task, participants needed to accurately compare perceived visual information (which signaled the magnitude of the virtual rotation) with the perceived extraretinal information (which signaled the magnitude of the real rotation). Compared to a 40° FOV, more visual information is generated in a 110° FOV viewport. However, since the rotations are the same, the amount of extraretinal information received remains the same across both FOVs. It is possible that compared to the 40° FOV viewport, the increased visual information received with a 110° FOV viewport diminishes the participants' ability to differentiate between visual and extraretinal information, thus making it harder for participants to successfully distinguish between rotation gains.

Previous research by Bruder et al. did not find gender differences with a 40° FOV [9]. Our results also did not find gender differences with 40° FOV. However, the 110° FOV condition found significant gender differences between threshold gains. Females were unable to discriminate between 90° virtual rotations and real rotations ranging from 65° to 171° equating to a 28% decrease and 90% increase in rotations. Males were unable to discriminate between 90° virtual rotations and real rotations ranging from 56° to 126°, equating to a 37% decrease and 40% increase in rotations. Additionally, male thresholds were significantly larger in the 110° FOV compared to the 40° FOV where males were only unable to discriminate between 90° virtual rotations and real rotations ranging from 75° to 120°. This supports H2, that participant discrimination between rotation gains is different for females compared to males, however this difference is only seen in the 110° FOV. In modern HMDs, designers should use different threshold gains for males and females.

When considering distractors, significant differences were found in the 110° FOV condition at the 25% threshold. In the presence of a distractor, female participants were unable to discriminate between 90° virtual rotations and real rotations ranging up to 244°, compared to only 139° when no distractor was present. Due to this significant difference, we recommend for a general RDW implementation to use the female 25% threshold without distractors

present. Males, in the presence of a distractor, were significantly different from females in the presence of a distractor and were unable to discriminate between 90° virtual rotations and real rotations ranging up to 124°. The results provide some support of H3, that participant discrimination between rotation gains is different when distractors are present compared to when they are not.

Perceptual studies have shown that the observer's attention and cognitive load can either increase or decrease one's sense ofvection, but those studies have provided mixed results [57, 68]. It is difficult to directly connect results found in these studies to the present study as participants in those studies were instructed to focus on a particular stimulus, or were asked to perform a memory exercise. In our experiment, we did not specifically instruct participants to focus on the deer as it ran across the VE, and therefore cannot know how much users attended to the distractor. We are confident that users at least partially attended to the distractor since they correctly reported seeing the distractor for 97% of the trials when it was present.

Work studying distractors in VEs by Peck et al. [53] and Chen et al. [13] either specifically instructed users to look at the distractor, or implemented the distractor such that users attended to it in order to complete a different goal. We designed the distractor in this study to emulate natural distractors that will not always capture users' full attention, such as a tour guide in a virtual house tour.

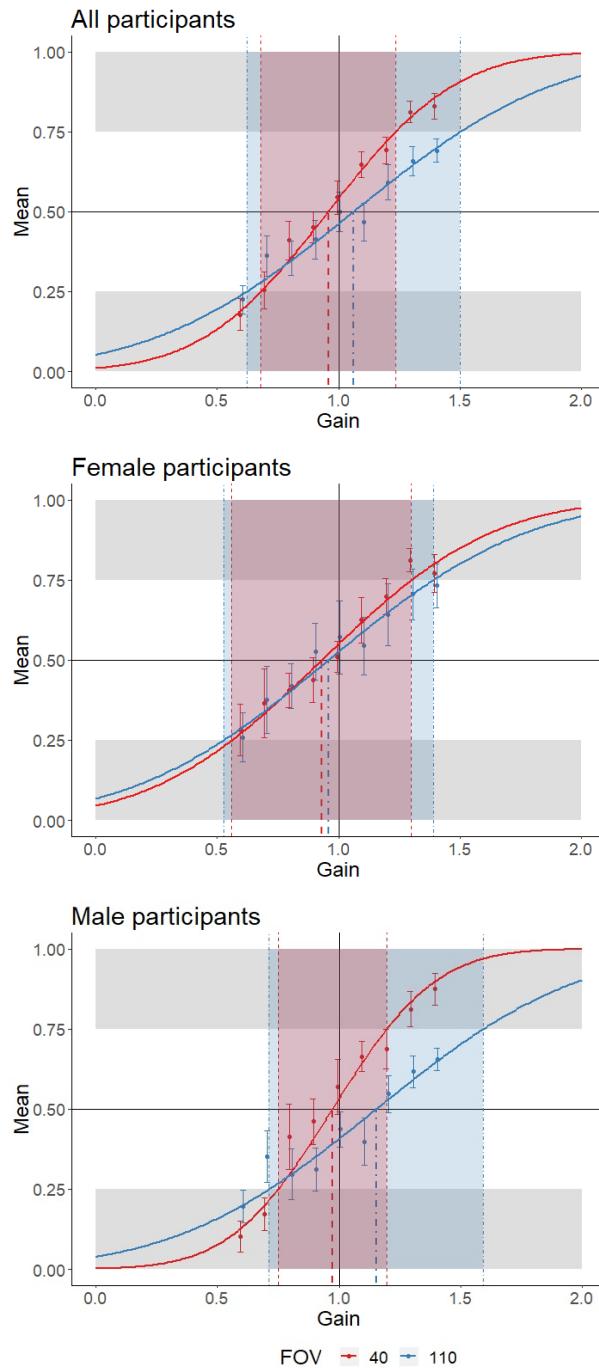


Fig. 4.7 The average probability of responding “*greater*” to the question, “Was the virtual movement *smaller* or *greater* than the physical movement?”, with standard error, for each rotation gain, and the fitted psychometric curve. Responses for 40° FOV are in red, and 110° FOV are in blue. The vertical edges of the red and blue regions indicate the 25% (left edge) and 75% (right edge) thresholds. In each condition, gains that fall inside the colored regions are undetectable to users. (Top) Perceptual thresholds for all participants, (Middle) perceptual thresholds for female participants, (Bottom) perceptual thresholds for male participants.

With and Without Distractors						
FOV	25%	PSE	75%	σ	D	p
40	0.6800	0.9572	1.2343	0.44	5.33	0.88
40 ♀	0.5604	0.9305	1.3006	0.55	4.43	0.92
40 ♂	0.7490	0.9726	1.1962	0.33	11.86	0.32
110	0.6207	1.0613	1.5019	0.65	8.69	0.70
110 ♀	0.5261	0.9579	1.3896	0.64	3.62	0.97
110 ♂	0.7122	1.1527	1.5932	0.65	11.05	0.54

Without Distractors						
FOV	25%	PSE	75%	σ	D	p
40 [63]	0.67	0.96	1.24	-	-	-
40 ♀[9]	0.66	0.96	1.29	-	-	-
40 ♂[9]	0.69	0.94	1.19	-	-	-
40	0.6800	0.9481	1.2170	0.40	4.70	0.93
40 ♀	0.5742	0.9286	1.2829	0.53	7.34	0.85
40 ♂	0.7382	0.9586	1.1790	0.33	8.24	0.75
110	0.6702	1.0555	1.4407	0.57	4.04	0.95
110 ♀	0.6459	0.9839	1.3218	0.50	3.69	0.98
110 ♂	0.6999	1.1307	1.5616	0.64	4.58	0.88

With Distractors						
FOV	25%	PSE	75%	σ	D	p
40	0.6810	0.9675	1.2541	0.42	4.31	0.91
40 ♀	0.5455	0.9327	1.3198	0.57	8.50	0.52
40 ♂	0.7619	0.9888	1.2156	0.34	6.95	0.78
110	0.5676	1.0678	1.5680	0.74	6.43	0.90
110 ♀	0.3692	0.9232	1.4772	0.82	3.74	0.96
110 ♂	0.7242	1.1727	1.6211	0.66	8.09	0.56

Table 4.2 The 25%, PSE, and 75% threshold gains derived from the psychometric curves (goodness-of-fit data (Deviance (D) and p-value)) calculated using maximum likelihood estimation, using a cumulative normal distribution function, of the pooled $\psi(greater|g_i)$ data. Gains are presented by FOV and gender (female - ♀, male - ♂) for all trials regardless of distractors, trials without distractors, and trials with distractors.

Effect	df	F	η^2	p
GEN	1, 9	5.22 *	.05	.05
THR	2, 18	17.11 ***	.39	<.0001
GEN×THR	2, 18	0.06	.002	.94
FOV	1, 9	0.52	.002	.49
GEN×FOV	1, 9	11.19 **	.04	.009
DIS	1, 9	0.45	.002	.52
GEN×DIS	1, 9	1.72	.006	.22
THR×FOV	2, 18	4.28 *	.09	.03
GEN×THR×FOV	2, 18	0.27	.006	.77
THR×DIS	2, 18	1.02	.010	.38
GEN×THR×DIS	2, 18	2.78 +	.03	.09
FOV×DIS	1, 9	3.82 +	.005	.08
GEN×FOV×DIS	1, 9	4.17 +	.005	.07
THR×FOV×DIS	2, 18	0.32	.002	.73
GEN×THR×FOV×DIS	2, 18	2.96 +	.02	.08

Table 4.3 Analysis of the 25%, PSE, and 75% gains calculated from each participant's psychometric curve with a 2 (distractors (DIS): present, not present) x 3 (Threshold (THR): 25%, PSE, 75%) x 2 (FOV: 40°, 110°) x 2 (Gender (GEN): female, male) ANOVA. Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Threshold	40° FOV		110° FOV	
	ρ_w	p	ρ_w	p
25%	-0.5579*	0.04	-0.2852	0.31
50%	-0.6416*	0.02	0.1031	0.72
75%	0.3732	0.19	0.6373*	0.01

Table 4.4 Winsorized correlations with 10% trimmed means comparing threshold gains to FOV for all participants.

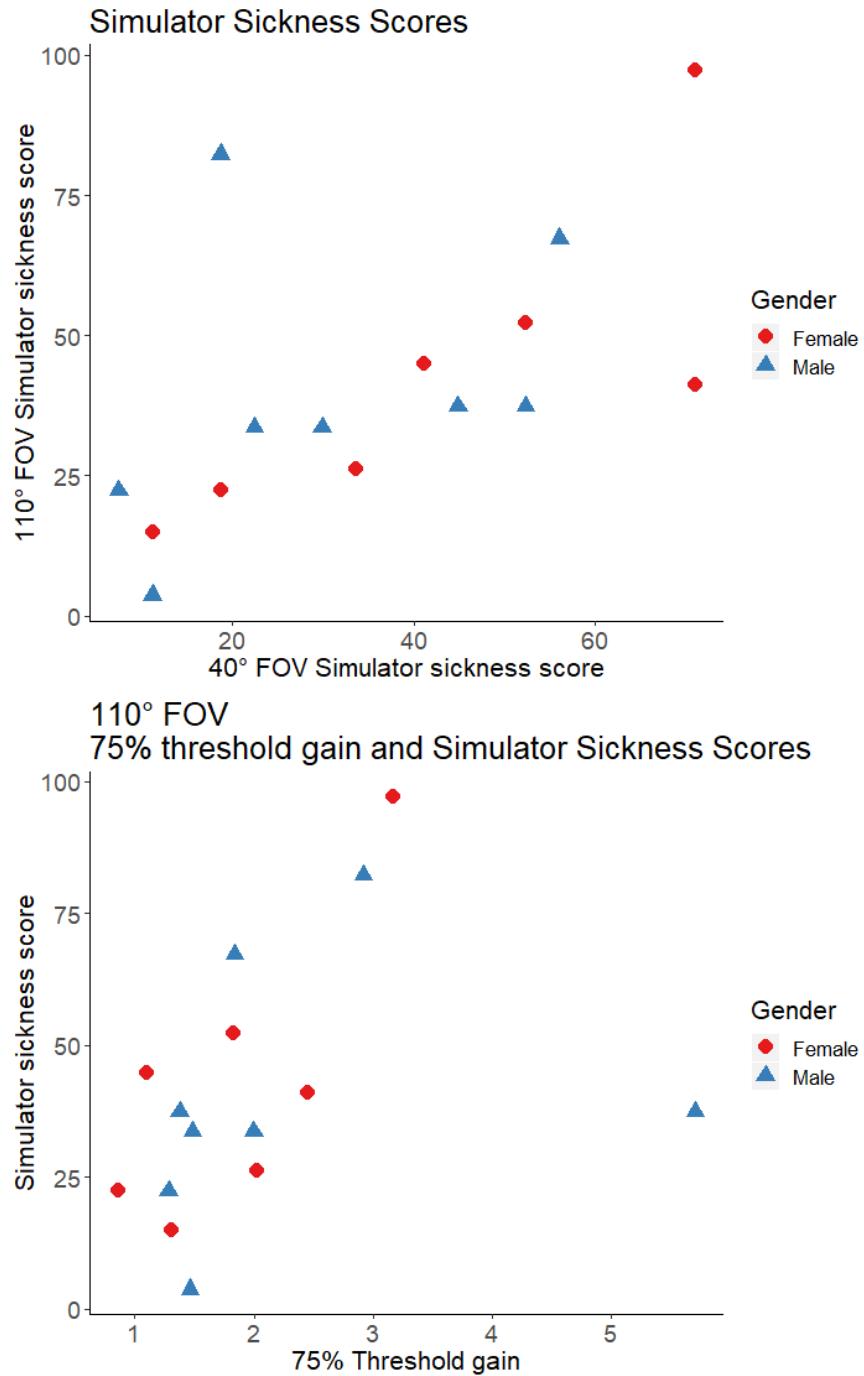


Fig. 4.8 (Top) The simulator sickness scores of each participant after the 40° FOV block (x-axis) and the 110° FOV block (y-axis). (Bottom) The simulator sickness scores of each participant after the 110° FOV block and the 110° FOV rotation gain at the 75% detection threshold. A male outlier is seen on the far right. Female participants are denoted with red circles, and male participants with blue triangles.

Chapter 5

Fitting Psychometric Functions

Psychophysical experiments involve quantifying the relationship between physical stimuli and perception of these stimuli. VR relies heavily on the user’s perception of stimuli to create a particular experience. Thus, the study of psychophysics is relevant to VR research. New statistical methods that quantify how humans interact with VR can provide necessary insight into the psychophysics of VR.

5.1 Introduction

Visual psychophysics studies the relationship between physical stimuli and visual perception [41]. Psychophysical experiments typically require participants to experience some stimulus and then report on the perceived intensity of the stimulus. In the case of this thesis, the stimulus is the VE rotation and the participant’s response is “*greater*” or “*smaller*”.

Psychophysical experiments typically require participants to undergo anywhere from one hundred to three hundred trials. This is done to generate enough data to carry out a rigorous analysis. Thus, psychophysical experiments can often last over one hour. Our experiments lasted two to two and a half hours.

Since these experiments take so long, participants can succumb to fatigue and boredom, which can negatively affect performance. This chapter will discuss our investigation of the viability of using a novel psychometric model to reduce the total number of trials required per participant. We implement a psychometric model introduced by Yi et al. that incorporates participants' confidence in their trial responses [78].

5.2 Confidence Model

Yi et al. proposed a psychometric model that can produce accurate results using less data given access to participants' confidence ratings [78]. The intuition behind this model is that confidence provides insight into a user's self-assessment of the conviction in a decision [78]. Yi et al. suggest that the additional information provided by confidence ratings allows us to calculate thresholds with fewer user trials. The traditional method used in Section 4 only uses the participants' “*greater*” and “*smaller*” response data to estimate psychometric curves. With this new method, we also have information about the participants' confidence in their response. This effectively makes each trial more “valuable”—from a single trial we gain more information that we can use in our analysis to estimate psychometric curves more quickly. In the original paper, this new model is referred to as the confidence signal detection (CSD) model.

The psychometric curves generated in Section 4 describe the relationship between a gain and the probability that a user replies “*greater*” when that gain is applied. With the CSD model, the curves we calculate describe the relationship between a gain and how confident the user is that their answer of “*greater*” was correct at that gain.

The model uses a confidence-scaling factor, k , to adjust the psychometric curve to account for underconfidence or overconfidence. Using the confidence model, the psychometric function becomes:

$$\chi = \Phi(x; \mu, k\sigma) \tag{5.1}$$

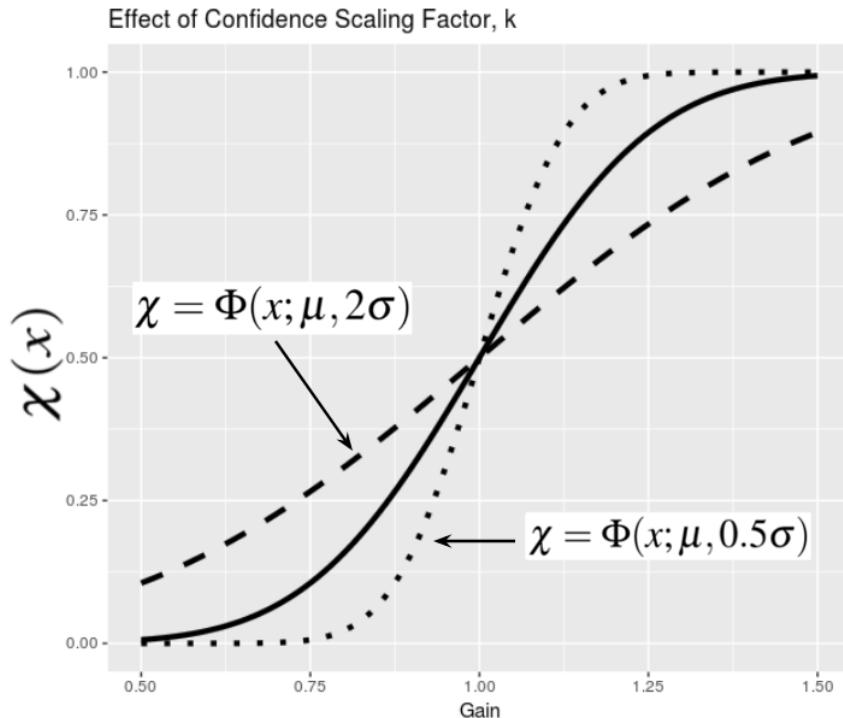


Fig. 5.1 The effect of the confidence scaling factor on the cumulative Gaussian curve. When $k > 1$, the resulting curve (dashed) represents underconfidence. When $k < 1$, the resulting curve (dotted) represents overconfidence.

where Φ is a Gaussian cumulative distribution function, μ is the mean of the Gaussian, and σ is the standard deviation of the Gaussian. Refer to Figure 5.1 for an illustration of how k influences a psychometric curve. It should be noted that because k is multiplied by σ in χ , the confidence model still only fits two parameters. However, it is helpful to separate k from σ for the purpose of understanding how confidence affects our fit.

The function fitting process is an optimization problem: we want to find the parameters μ , σ , and k that approximate a psychometric curve that best fits the user response data. To do this, we choose some μ , σ , and k and compare the CSD curve to the traditional psychometric curve. We use maximum likelihood estimation to calculate how closely the CSD curve fits to the original psychometric curve, and we iteratively vary μ , σ , and k to optimize the fit of the CSD model.

5.3 Methods

Our experiments tested the following hypothesis:

H4 We can use the CSD model to estimate threshold values that are not statistically different than values estimated using traditional methods.

During the experiment discussed in Section 4, we asked users to report how confident they were that their trial response was correct. They rated confidence on a Likert scale ranging from 1 to 5 in increments of 1. These are the confidence ratings we used in our implementation of [78].

The CSD model is interesting because Yi et al. claim that by using the model, we are able to reduce the number of trials needed to converge to an accurate μ and σ . To test if this was the case, we systematically reduced the number of trials used in our analysis. Our original experiment features eight trials per gain. If we remove the last two trials for each gain from the participant’s data, this is effectively the same as running the experiment with only six trials per gain. We can repeat this process to acquire data for an experiment with only four and two trials per gain, as well. We decrement the number of trials in groups of two to keep the number of trials even, which helps prevent the data from being skewed to one side of the curve and allows us to more easily calculate a PSE. The model was implemented using the R programming language.

5.4 Results

In general, we were unable to reliably fit psychometric functions using the CSD model proposed in [78]. If the CSD model works as described by Yi et al. we would expect the μ and $k\sigma$ to be similar to the μ and σ for the regular psychometric curves (which we would verify with an ANOVA). This was the case for some participants, but there were many participants for whom we could not converge on a similar μ and $k\sigma$ for each experiment

condition. A sample of the μ and $k\sigma$ compared to the μ and σ calculated with the traditional method is shown in Table 5.1. The calculated parameters showed similar trends for all other conditions (all FOVs, distractor conditions, and numbers of trials) and therefore are not included in this text.

Although the CSD model did not work for all participants, it converged to reasonable μ and $k\sigma$ values for some participants in some experimental conditions. Note that convergence for one participant in one condition did not mean the model converged for their data in another condition. In the 40° FOV, no distractor, six trials condition shown in Table 5.1, the CSD model converged to reasonable values for twelve of the sixteen participants. If we only consider these twelve participants, the mean $\mu = 1.9763$ and the mean $k\sigma = 1.7467$. An example of the normal psychometric curve from Section 4 compared to the curve calculated in Table 5.1 for one participant is shown in Figure 5.2. An example of a failed fit for the CSD model is participant 7 in Table 5.1. Assuming our data were distributed normally, we would compare the mean μ and σ from the traditional analysis with all trials to the μ and $k\sigma$ from the CSD model with fewer trials. Non-significant differences would suggest that the μ and $k\sigma$ from the CSD model are statistically no worse than the μ and σ calculated using the original method. We would also perform equivalence testing using Bayesian statistics. This analysis could not be conducted because our results violate the assumptions of the ANOVA test (specifically, our results were not normally distributed).

40° FOV Without Distractors (6 trials per gain)

ID	Original fit		CSD fit	
	μ	σ	μ	$k\sigma$
1	1.0274	0.3422	0.8988	0.4685
2	-	-	-	-
3	0.9592	0.8738	12.2302	13.1579
4	0.9782	0.3131	0.7398	0.3902
5	-	-	-	-
7	0.6664	2.3759	-2094.7892	9282.2374
8	0.9731	0.3009	0.8571	0.4347
11	0.8013	0.8129	1.3481	1.5719
12	-	-	-	-
13	0.9182	0.1417	1.2024	0.6667
14	0.7635	0.1817	1.0582	0.3443
15	1.1168	0.1568	1.1886	0.4349
16	1.0309	0.2365	0.7844	0.4100
17	0.4024	1.0834	1.8069	1.7646
18	1.0020	0.3510	0.7962	0.6336
19	1.0628	0.4205	0.8055	0.6837

Table 5.1 μ , σ and $k\sigma$ for each participant with 40° FOV, no distractors, and six trials per gain. The CSD model was unable to converge to any μ and $k\sigma$ for participant 7. There was no fit for participants 2, 5, and 12 using the traditional psychometric curve fitting process, so we are unable to use the CSD model to calculate μ and $k\sigma$ for them.

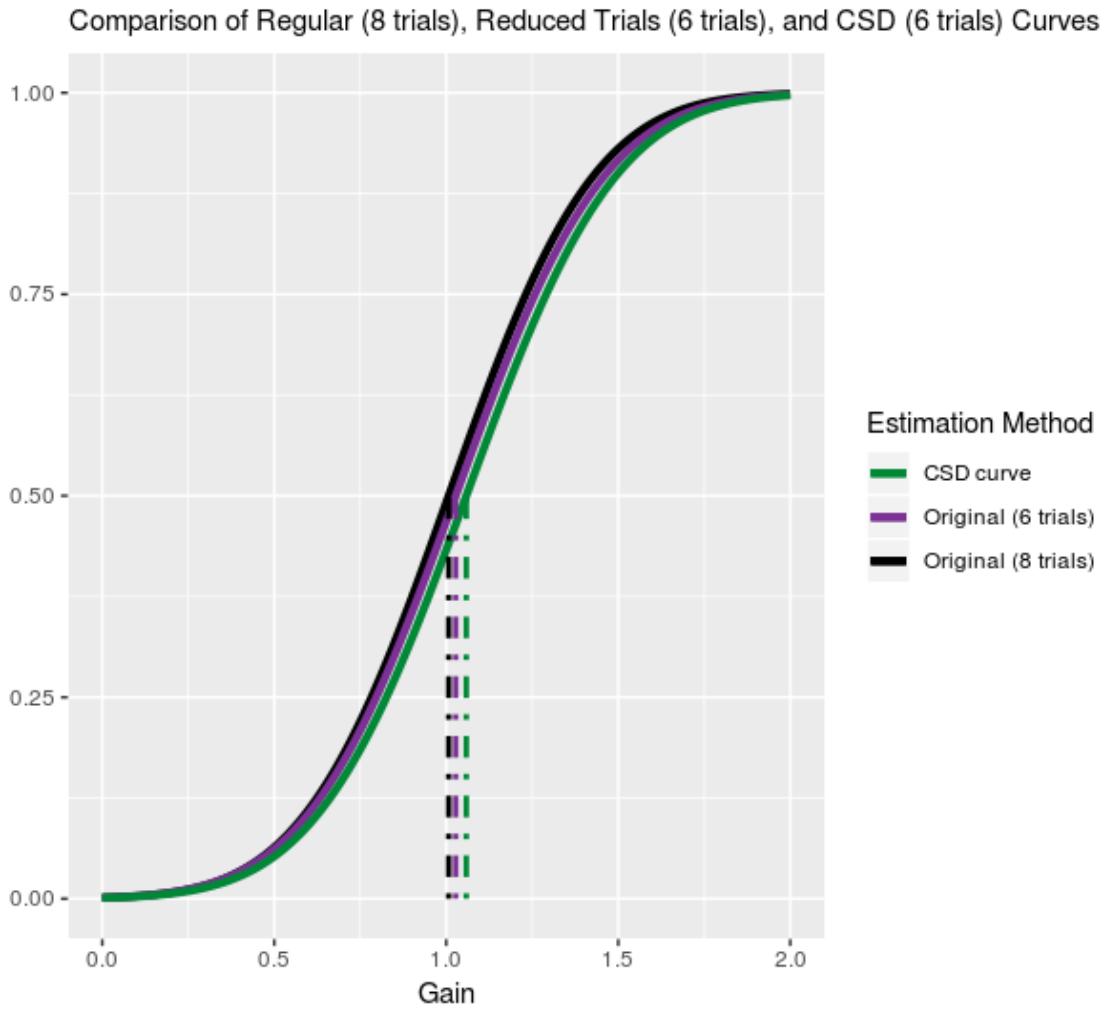


Fig. 5.2 The psychometric curves calculated using the traditional method with eight trials (black), the traditional method with six trials (purple), and the CSD model with six trials (green). The CSD model comes close to approximating the “ground truth” represented by the black curve. The data used to generate these curves comes from participant 14.

5.5 Discussion

There are many different reasons why the CSD model may have failed on our data. One reason is that the participants’ confidence ratings may simply be distributed in a shape such that it is unusable in the CSD model. In particular, the CSD model assumes that participant confidence ratings are normally distributed. This is not the case for some participants. See

Figure 5.3 for an example of a poor confidence rating distribution. No matter how hard we try, fitting a reasonable cumulative Gaussian to the data in Figure 5.3 will not be possible.

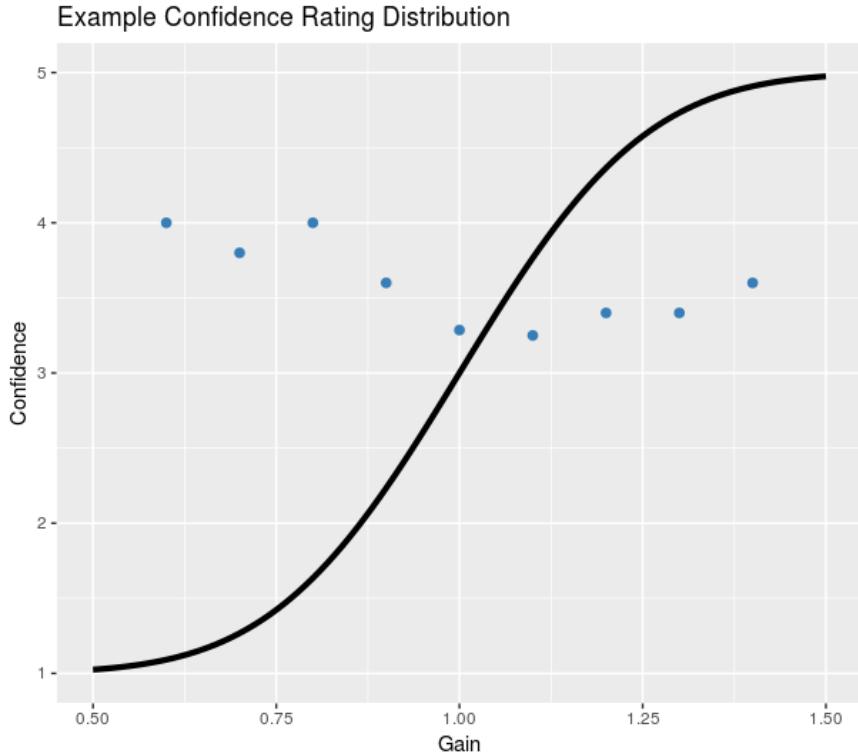


Fig. 5.3 The average confidence rating for each gain for participant 7 (110° FOV, with distractors). The curve we want to fit, a cumulative Gaussian, is shown as a black curve on the same graph. Fitting a cumulative Gaussian to the participant's data in this case is infeasible.

In an attempt to mitigate this problem, we converted trials where users replied “*smaller*” into trials with the response “*greater*” and the opposite confidence rating via the following formula

$$c' = 6 - c \quad (5.2)$$

where c is the confidence rating and c' is the converted confidence rating. Suppose a participant is highly confident that the virtual rotation was smaller than the physical rotation (answered with “*smaller*” with a confidence rating of 5). If, for some reason, on this same trial the participant instead believed that the virtual rotation was *greater* than the physical

rotation, he or she should have low confidence in this belief. If a participant answered a trial with “*smaller*” and had a confidence rating of 4, we replaced this with a trial in which they replied “*greater*” with a confidence of 2. This process of converting trials to “*greater*” will be referred to as “inverting” the data.

We invert the data so that we can augment our usable dataset to include trials that were originally ignored by the model [78]. Inverting the data helps us model the participant’s confidence ratings at gains for which they mostly replied with “*smaller*” (since the analysis requires that we only consider trials with a response of “*greater*”). By inverting these potentially well-distributed “*smaller*” trials to “*greater*” trials, we hope to get more useful data to use in our CSD model. Unfortunately, this did not work. For participants whose confidence ratings were not distributed in a sigmoid-like shape, converting their “*smaller*” trials to “*greater*” trials, in general, did not improve the distribution shape. As a result, the μ and $k\sigma$ calculated with this converted data did not improve in any meaningful way. Some participants’ μ and $k\sigma$ did improve after inverting the data, but there were still participants for which the model could not converge on a μ and $k\sigma$.

We believe that inverting the data does not work because it changes the implied meaning of a response given the way we designed our study. If a user replies with high confidence, it is reasonable to assume that he or she would have low confidence when forced to reply with the other option. However, if a user replies with low confidence, it is not necessarily the case that he or she would feel very confident in the correctness of the alternative answer. A response with low confidence implies that the user does not know if their answer is correct. If they cannot answer with conviction one way, it is probable that they will not be able to answer confidently the other way, too. A good example of this interaction is at the PSE: at the PSE, users should be unconfident in all of their answers regardless of which option they choose. Furthermore, if it *were* true that inverting an answer of “*smaller*” with low confidence to an answer of “*greater*” with high confidence were a legitimate method, there is

no reason why the user would not have replied with “*greater*” and high confidence in the first place, since that option was available to them during the experiment.

In the model described by Yi et al. the experiment asked participants to state if they rotated left or right which is similar to our question of “*greater or smaller*.” However, in [78], the authors used a half-range confidence scale; the confidence ratings only ranged from 50% to 100%. In order to model one half of the sigmoid shape with the study design used in [78], inverting the data is a requirement. Because we used a full-range confidence scale, inverting the data no longer implies the same thing as in [78] which may confound the analysis.

It is also clear to us that participants’ confidence ratings are not distributed in a way that we expected them to be. One reason this may be the case is because each participant has their own subjective interpretation of how the confidence scale options relate to their own feelings of confidence. For example, it is possible that two participants can feel the same level of confidence but will report different scores on the confidence scale [17, 18].

Furthermore, it is known that users exhibit central tendency when using rating scales. Central tendency refers to the reluctance to use the extreme ends of any scale [1]. There is evidence of cognitive load affecting the prominence of central tendency, which could also be a factor in our experiments [2]. It is not clear if our experiments cause participants to experience high or low cognitive load, but future research in this area should be aware of this potential confounding factor.

Another factor that may have contributed to the poor distribution of confidence ratings is the range of gains we tested. We believe it is possible that the range of gains we tested was too small for us to see the full gamut of confidence ratings. If most participants are not highly confident in their replies until the gain reaches something outside of the gains we tested, then it is unlikely that we will see ratings that correspond to high confidence in our data.

To test this idea, we ran a pilot study on one participant. In this pilot study, the participant experienced gains ranging from 0.3 to 1.7, incremented in steps of 0.1. To keep the experiment

40° FOV Without Distractors

ID	Original fit		CSD fit		CSD fit (inverted data)	
	μ	σ	μ	$k\sigma$	μ	$k\sigma$
20	0.9094	0.2634	4277.3395	8530.5920	6.2561	19.0077

40° FOV With Distractors

ID	Original fit		CSD fit		CSD fit (inverted data)	
	μ	σ	μ	$k\sigma$	μ	$k\sigma$
20	1.1073	0.3292	2897.5377	8501.5304	2.8598	2.6161

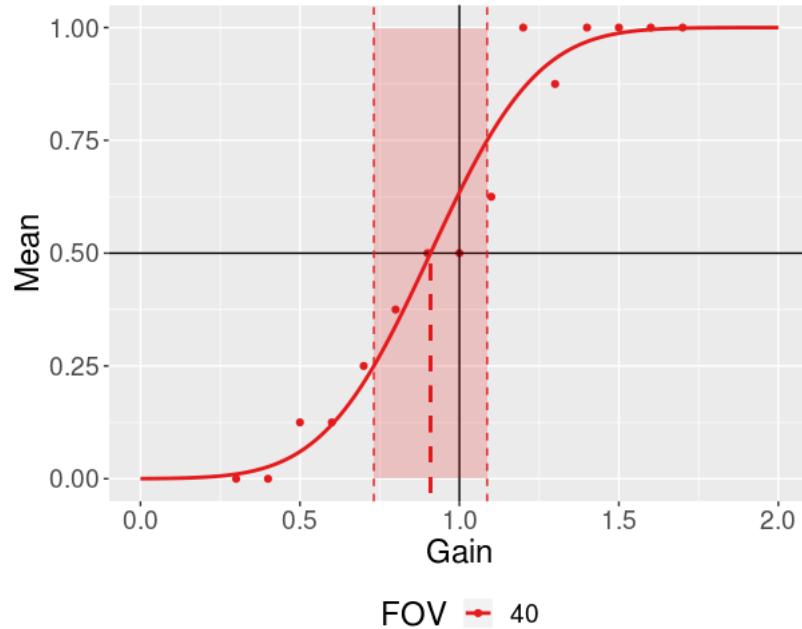
Table 5.2 μ , σ and $k\sigma$ for the pilot participant. In this experiment, we tested gains ranging from 0.3 to 1.7. The CSD model is unable to converge with just “*greater*” trial data, but it is able to converge after inverting the data. This behavior was seen in the original CSD model experiments for some participants. It should be noted that the μ and $k\sigma$ values in the pilot study are still not good estimations.

within two hours, we only tested 40° FOV both with and without distractors. Thus, the entire pilot study consisted of 240 trials. The participant had a 5 minute break after 120 trials. The participant was a 21 year old male with corrected to normal vision, identified as a hard-core gamer, and had brief experience with VR before. His SSQ score was 7.48 halfway through and after the experiment. This person did not participate in the original experiment. He was not naïve to the purposes of the original study in which we compare thresholds under different conditions, but he was naïve to the confidence analysis portion of the study.

The results of the pilot study are summarized in Table 5.2 and Figures 5.4a and 5.4b. Table 5.2 presents the participant’s μ , σ , and $k\sigma$. Figures 5.4a and 5.4b present the psychometric curves for the pilot participant, calculated by the traditional method used in Section 4. Without distractors, the participant’s thresholds are: 25% = 0.7317, PSE = 0.9094, and 75% = 1.0870. With distractors, the participant’s thresholds are: 25% = 0.8852, PSE = 1.1073, and 75% = 1.3293. These thresholds are very close to those found in Section 4 which suggests that the participant’s data is normal.

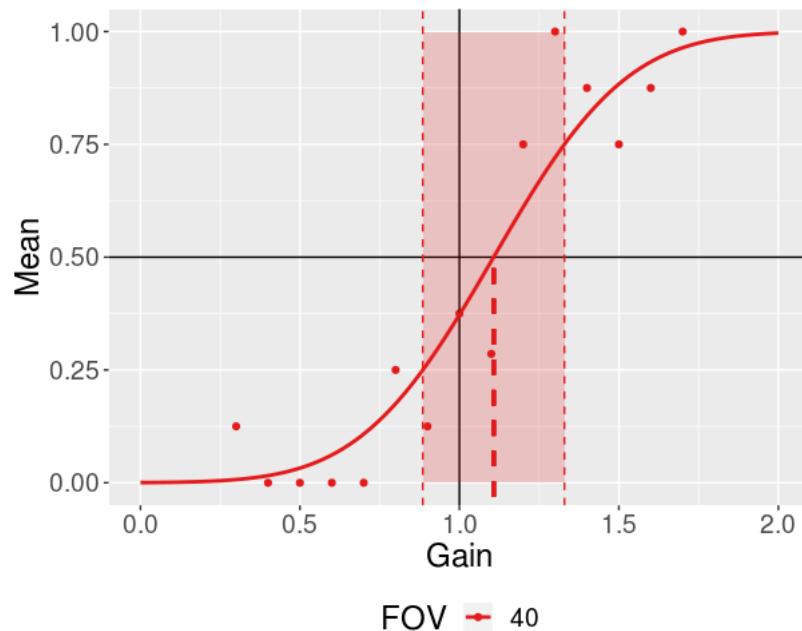
Although the CSD model did not yield good estimates in the pilot study, we did see some improvements in the data. See Figures 5.5 and 5.6 for graphs of the pilot participant’s

Pilot Study: Perception Thresholds Without Distractors



(a) The psychometric curve for the pilot participant without distractors.

Pilot Study: Perception Thresholds With Distractors



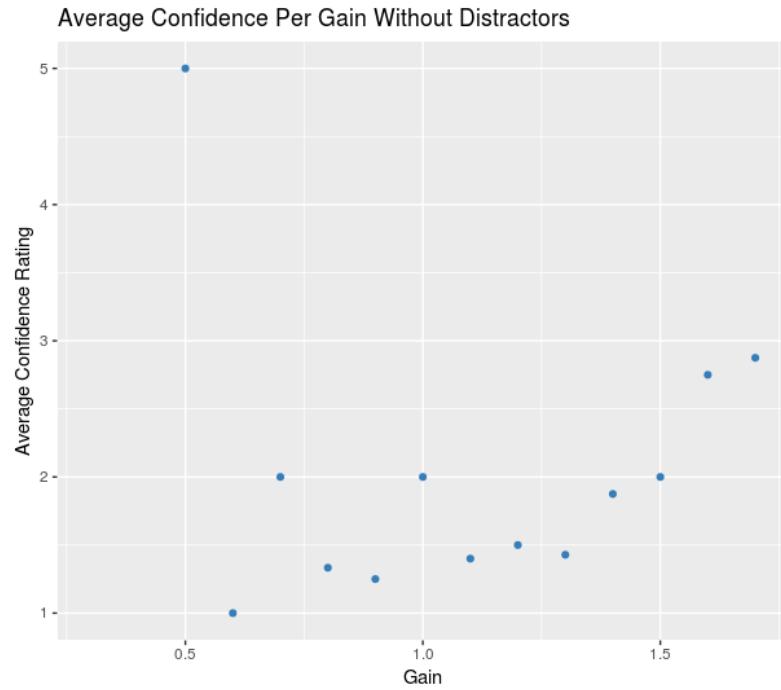
(b) The psychometric curve for the pilot participant with distractors.

Fig. 5.4 Psychometric curves of the pilot participant, calculated using the method used in Section 4.

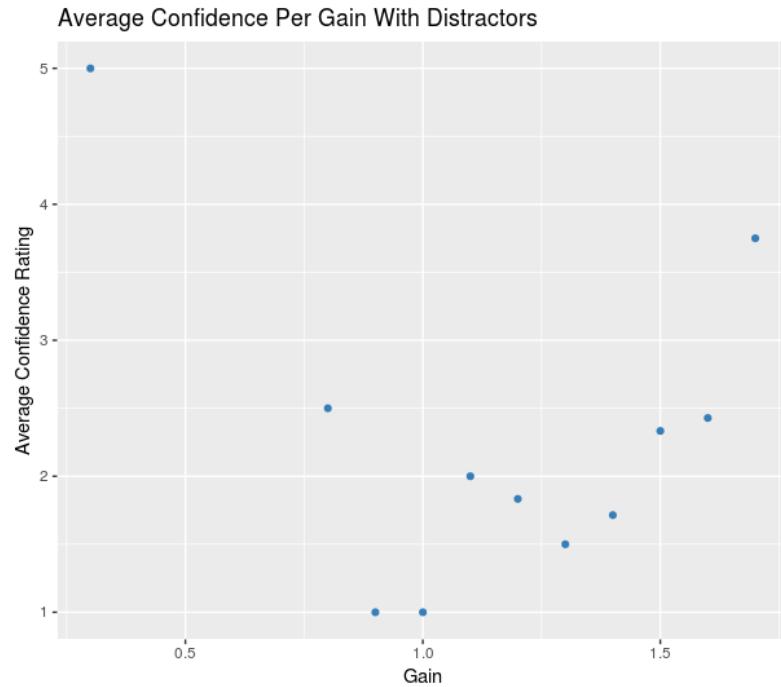
confidence rating distributions. As these figures show, the distribution of confidence values is closer to the sigmoid shape that we initially expected to see once we invert the data. This explains why the CSD fit in Table 5.2 does not converge until we invert the data.

We believe the largest contributing factor to the failed fits in our pilot study is the poor distribution of confidence scores. Despite increasing the range of gains (with the hopes of yielding a more sigmoid-like distribution), the pilot participant’s confidence ratings cannot be easily fit to a cumulative Gaussian. As discussed earlier, this is possibly due to the participant’s subjective bias and sensitivity (reliability) in confidence ratings, or due to the way in which we presented the confidence rating question. However, we believe the most likely reason why the confidence rating distribution of our pilot participant was so unexpected is because the factors influencing confidence decisions are not the same as those that influence perceptual decisions. There is evidence that a user’s performance is not directly correlated to their confidence rating distributions [79]. That is, the participant may perceive the stimulus and use this information to correctly answer the question about the stimulus, but the information used to make this decision is not used when determining their confidence in the response.

This dissociation between perceptual decisions (reporting perceived stimulus direction) and self-reflection decisions (reporting confidence ratings) can explain why we are able to calculate threshold gains using the traditional method but not the CSD model. The traditional threshold gain calculation method relies only on the correctness of the participants’ responses, while the CSD model relies on the participants’ confidence ratings. Our study was not designed with the CSD model in mind, so there are likely confounding elements of our study that affect the participants’ confidence ratings. Further studies that are designed with this phenomenon in mind should be conducted.

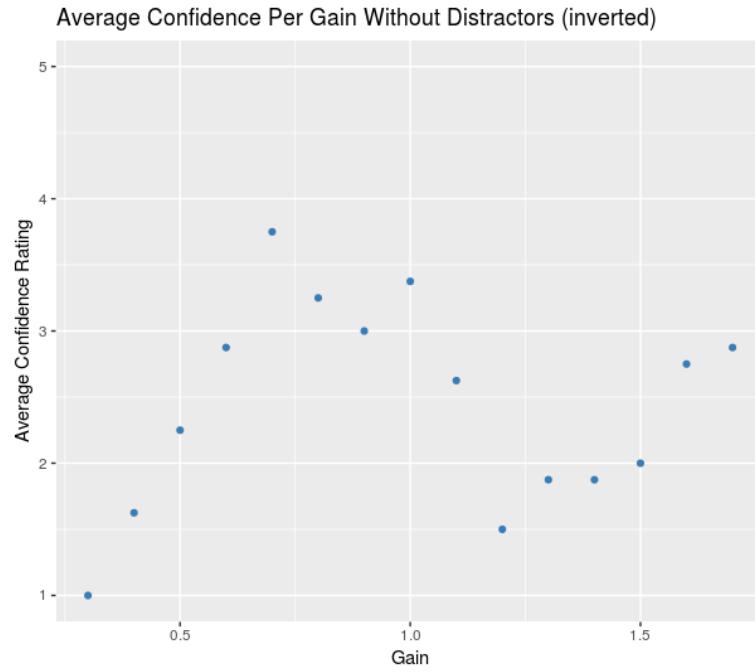


(a) Distribution of confidence ratings for the pilot study, without distractors.

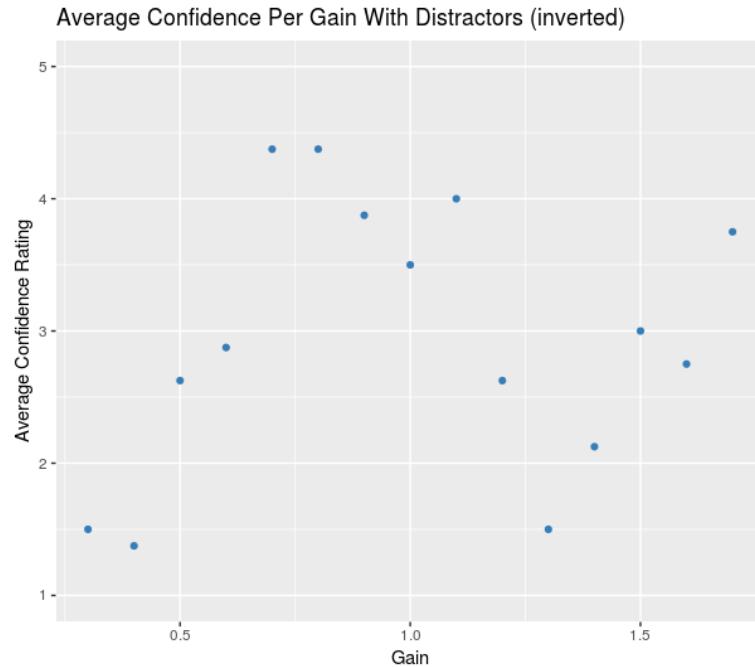


(b) Distribution of confidence ratings for the pilot study, with distractors.

Fig. 5.5 Distribution of confidence ratings in the pilot study for trials with a response of “greater”. The plots do not form a sigmoid shape, so the CSD model cannot fit a μ and $k\sigma$.



(a) Distribution of confidence ratings for the pilot study after inverting the data, without distractors.



(b) Distribution of confidence ratings for the pilot study after inverting the data, with distractors.

Fig. 5.6 Distribution of confidence ratings in the pilot study for the inverted pilot study data. The plots are closer to the sigmoid shape, so the CSD model does a better job at fitting a μ and $k\sigma$.

Chapter 6

Conclusions and Future Work

6.1 Comparison of Rotation Gain Thresholds

In the first half of this thesis we developed and conducted experiments that compare the perceptual detection thresholds of redirected walking rotation gains under different user and system conditions. Specifically, our experiments studied the differences in thresholds considering different FOVs, visual distractor conditions, and user genders. The main conclusions that can be drawn from this research are:

1. There are significant differences in detection threshold gains between FOVs.
 - Participant discrimination between rotation gains is different in a 110° FOV compared to a 40° FOV. Rotation gains are significantly wider with a 110° FOV compared to a 40° FOV.
2. Detection threshold gains show differences between genders.
 - Significant differences were found between female and male gains with a 110° FOV and males had significantly wider gains using a 110° FOV compared to a 40° FOV.

3. There are strong correlations between simulator sickness scores and detection threshold gains.
 - We did not find any significant effects or interactions between simulator sickness and FOV and gender. However, we did find a significant correlation between users' simulator sickness scores and threshold gains.
4. Participant discrimination between rotation gains is different when distractors are present compared to when they are not, under *some* conditions.
 - We found a significant difference in threshold gains for females when a distractor was present versus when a distractor was not present. We also found a significant difference in threshold gains between males and females when a distractor was present.

As discussed in Section 4, these results highlight the importance of considering user and system factors when calibrating RDW thresholds.

6.2 Efficient Estimation of Detection Thresholds

In the second part of this thesis, we investigated the viability of using a novel statistical model to more efficiently calculate users' detection thresholds. The novel model we implemented and tested involved using participants' confidence in their trial responses to calculate detection threshold gains using fewer trials [78]. Our experiments did not yield results supporting the hypothesis that this novel confidence model can be used to reliably and efficiently calculate detection thresholds.

Possible reasons why this model failed are:

1. User's confidence ratings are not distributed normally.

- The confidence model requires that we fit a cumulative Gaussian to each user's confidence response data. If their data is distributed such that it creates a shape that does not resemble a sigmoid, the confidence model will fail to find a fit to their data.
2. Users' confidence ratings are too subjective for us to reliably expect them to have similar rating distributions.
- Some users had sigmoid-shaped confidence distributions, while others did not. The confidence model is potentially not robust enough to tolerate the differences in distribution shapes that can arise from users having their own subjective confidence scoring metrics.
3. We did not test a large enough range of gains to elicit the full range of confidence values from the users.
- If the gains are never strong enough to convince users of the correctness of their trial responses, the confidence rating distributions will not form the full sigmoid shape. Similarly to item 1, this will make it difficult to fit a cumulative Gaussian to the confidence data.

As discussed in Section 5, to address the issues described above, we post-processed the participant confidence data via inversion and ran a pilot study. These efforts did not produce improved results, which suggests that there may be some mismatches between the conditions of the original study by Yi et al. and the conditions of our experiment.

6.3 Future Work

6.3.1 Redirected Walking

This thesis detailed RDW thresholds under only a few particular conditions. It is possible that alternative implementations of the conditions we studied can influence the estimated threshold gains. For example, different implementations of FOV restrictors—solid borders that lessen the user’s FOV—may yield different threshold gains. Initial tests by Fernandes et al. [16] noted that hard-edge restrictors are more distracting and easier to notice than soft-edge restrictors. One participant in our study commented on trying to use the FOV restrictor edge to determine the rotation speed, which supports the observation made in [16]. Our study used a rectangular hard-edge restrictor for the sake of replication, but we believe different FOV restrictor edge softness and shapes are worth studying since there is evidence that the FOV restrictor parameters do not go unnoticed by users.

The significant difference in rotation threshold gains between 40° and 110° FOVs suggests that differences are likely to be found for translation and curvature threshold gains as well. Future work should estimate threshold gains for translation and curvature gains comparing a 40° FOV with a 110° FOV. Gender differences in thresholds have recently been found in curvature gains. Nguyen et al. found that females have higher curvature thresholds than men when wearing an HMD with a 100° FOV [43]. The significant differences found between genders in rotation gains, and the strong correlations with simulator sickness highlight the importance of considering gender and simulator sickness when evaluating translation and curvature gains.

Inclusivity in design is essential both in hardware and software [46]. In our study, six participants had a measured IPD below 61.3mm , which is the minimum IPD setting on the Vive Pro. Four of these participants were female. Although participants stated that they were able to clearly see the VE, a correct IPD setting will increase display sharpness. Previous

research by Willemsen et al. [76] suggests that overall performance in egocentric distance estimation was not improved using measured IPDs compared to the average *65mm* male IPD. The average American female IPD is *61.7mm* [19] and the average Asian female IPD is *63.6mm* [49]. The Vive Pro is designed such that almost 50% of the American female population and 20% of the Asian female population is unable to set the device to the correct IPD. This hardware limitation is a cause for concern and an example of non-inclusive design practices.

Finally, threshold gains were strongly correlated with simulator sickness scores. Participants with higher simulator sickness scores also had threshold gains farther from 1 compared to participants with lower simulator sickness scores. Follow-up research should explore the effect of simulator sickness on perception thresholds for RDW. In an effort to minimize the chance of making users sick, conservative practitioners may want to use threshold gains closer to 50% instead of the standard 25% and 75% thresholds until future research provides necessary insight into the relationship between simulator sickness and rotation gains.

6.3.2 Confidence Model

In their original paper, Yi et al. [78] noted that their CSD model is novel and that they invite other researchers to test the efficacy of it. Their study only included four human participants, and the rest of the experiments were conducted on synthetic data. Future research in perceptual VR studies should consider incorporating the CSD model to provide more evidence either in support of or against the model.

Future studies should also consider testing the CSD model using different study designs. Specifically, we believe there is merit in asking users to answer questions with different wording. In our study, the wording was chosen to force the user to pick one of two options (2AFC). This was a requirement to reduce bias in the traditional method of calculating detection thresholds. However, studies have shown that the wording of the trial question

and the way in which confidence ratings are presented (simultaneously or sequentially) can influence users' responses [18, 26]. Future RDW studies should explore different methods of presenting the confidence question since alternative methods may help reduce bias.

References

- [1] Albaum, G. (1997). The likert scale revisited. *Market Research Society. Journal.*, 39(2):1–21.
- [2] Allred, S. R., Crawford, L. E., Duffy, S., and Smith, J. (2016). Working memory and spatial judgments: Cognitive load increases the central tendency bias. *Psychonomic bulletin & review*, 23(6):1825–1831.
- [3] Barfield, W. and Hendrix, C. (1995). The effect of update rate on the sense of presence within virtual environments. *Virtual Reality*, 1(1):3–15.
- [4] Baumberger, B., Flückiger, M., and Roland, M. (2000). Walking in an environment of moving ground texture. *Japanese Psychological Research*, 42(4):238–250.
- [5] Berthoz, A. (2000). *The brain’s sense of movement*, volume 10. Harvard University Press.
- [6] Brandt, T., Dichgans, J., and Koenig, E. (1973). Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experimental brain research*, 16(5):476–491.
- [7] Bruder, G., Interrante, V., Phillips, L., and Steinicke, F. (2012). Redirecting walking and driving for natural navigation in immersive virtual environments. *IEEE transactions on visualization and computer graphics*, 18(4):538–545.
- [8] Bruder, G., Steinicke, F., and Hinrichs, K. H. (2009a). Arch-explore: A natural user interface for immersive architectural walkthroughs. In *2009 IEEE Symposium on 3D User Interfaces*, pages 75–82. IEEE.
- [9] Bruder, G., Steinicke, F., Hinrichs, K. H., Frenz, H., and Lappe, M. (2009b). Impact of gender on discrimination between real and virtual stimuli. In *Workshop on Perceptual Illusions in Virtual Environments*, pages 10–15. Citeseer.
- [10] Bruder, G., Steinicke, F., Hinrichs, K. H., and Lappe, M. (2009c). Reorientation during body turns. In *EGVE/ICAT/EuroVR*, pages 145–152.
- [11] Burg, A. (1968). Lateral visual field as related to age and sex. *Journal of applied psychology*, 52(1p1):10.
- [12] Chance, S. S., Gaunet, F., Beall, A. C., and Loomis, J. M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence*, 7(2):168–178.

- [13] Chen, H. and Fuchs, H. (2017). Supporting free walking in a large virtual environment: imperceptible redirected walking with an immersive distractor. In *Proceedings of the Computer Graphics International Conference*, page 22. ACM.
- [14] Corporation, V. (2019). Virtual reality tracked space diagram.
- [15] DiZio, P. and Lackner, J. R. (1997). Circumventing side effects of immersive virtual environments. In *HCI* (2), pages 893–896.
- [16] Fernandes, A. S. and Feiner, S. K. (2016). Combating vr sickness through subtle dynamic field-of-view modification. In *3D User Interfaces (3DUI), 2016 IEEE Symposium on*, pages 201–210. IEEE.
- [17] Fleming, S. M. and Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594):1338–1349.
- [18] Fleming, S. M. and Lau, H. C. (2014). How to measure metacognition. *Frontiers in human neuroscience*, 8:443.
- [19] Gordon, C. C., Blackwell, C. L., Bradtmiller, B., Parham, J. L., Barrientos, P., Paquette, S. P., Corner, B. D., Carson, J. M., Venezia, J. C., Rockwell, B. M., et al. (2014). 2012 anthropometric survey of us army personnel: Methods and summary statistics. Technical report, ARMY NATICK SOLDIER RESEARCH DEVELOPMENT AND ENGINEERING CENTER MA.
- [20] Grechkin, T., Thomas, J., Azmandian, M., Bolas, M., and Suma, E. (2016). Revisiting detection thresholds for redirected walking: Combining translation and curvature gains. In *Proceedings of the ACM Symposium on Applied Perception*, pages 113–120. ACM.
- [21] Halpern, D. F. (2000). *Sex differences in cognitive abilities*. Psychology press.
- [22] Hendrix, C. and Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators & Virtual Environments*, 5(3):274–289.
- [23] Hettinger, L. J., Berbaum, K. S., Kennedy, R. S., Dunlap, W. P., and Nolan, M. D. (1990). Vection and simulator sickness. *Military Psychology*, 2(3):171–181.
- [24] Hildebrandt, J., Schmitz, P., Valdez, A. C., Kobbelt, L., and Ziefle, M. (2018). Get well soon! human factors’ influence on cybersickness after redirected walking exposure in virtual reality. In *International Conference on Virtual, Augmented and Mixed Reality*, pages 82–101. Springer.
- [25] Hodgson, E., Bachmann, E., and Waller, D. (2011). Redirected walking to explore virtual environments: Assessing the potential for spatial interference. *ACM Transactions on Applied Perception (TAP)*, 8(4):22.
- [26] Hollard, G., Massoni, S., and Vergnaud, J.-C. (2010). Subjective beliefs formation and elicitation rules: experimental evidence.

- [27] Hulk, J. and Rempt, F. (1983). Vertical optokinetic sensations by limited stimulation of the peripheral field of vision. *Ophthalmologica*, 186(2):97–103.
- [28] Interrante, V., Ries, B., and Anderson, L. (2007). Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. In *3D User Interfaces (3DUI), 2007 IEEE Symposium on*, pages 167–170. IEEE.
- [29] Iwata, H. (1999). The torus treadmill: Realizing locomotion in ves. *IEEE Computer Graphics and Applications*, 19(6):30–35.
- [30] Iwata, H., Yano, H., Fukushima, H., and Noma, H. (2005). Circulafloor: A locomotion interface using circulation of movable tiles. In *null*, pages 223–230. IEEE.
- [31] Iwata, H., Yano, H., and Tomioka, H. (2006). Powered shoes. In *ACM SIGGRAPH 2006 Emerging technologies*, page 28. ACM.
- [32] Jaekl, P., Allison, R. S., Harris, L. R., Jasiobedzka, U., Jenkin, H., Jenkin, M., Zacher, J. E., and Zikovitz, D. C. (2002). Perceptual stability during head movement in virtual reality. In *Virtual Reality, 2002. Proceedings. IEEE*, pages 149–155. IEEE.
- [33] Jerald, J., Peck, T., Steinicke, F., and Whitton, M. (2008). Sensitivity to scene motion for phases of head yaws. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, pages 155–162. ACM.
- [34] Kennedy, R. S. and Frank, L. H. (1985). A review of motion sickness with special reference to simulator sickness. Technical report, CANYON RESEARCH GROUP INC WESTLAKE VILLAGE CA.
- [35] Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220.
- [36] Kolasinski, E. M. (1995). Simulator sickness in virtual environments. Technical report, Army research inst for the behavioral and social sciences Alexandria VA.
- [37] Langbehn, E. and Steinicke, F. (2018). Redirected walking in virtual reality. *Encyclopedia of Computer Graphics and Games. Springer International Publishing*.
- [38] Lappe, M., Bremmer, F., and Van den Berg, A. (1999). Perception of self-motion from visual flow. *Trends in cognitive sciences*, 3(9):329–336.
- [39] Lin, J.-W., Duh, H. B.-L., Parker, D. E., Abi-Rached, H., and Furness, T. A. (2002). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Virtual Reality, 2002. Proceedings. IEEE*, pages 164–171. IEEE.
- [40] Linares, D. and López i Moliner, J. (2016). quickpsy: An r package to fit psychometric functions for multiple groups. *The R Journal*, 2016, vol. 8, num. 1, p. 122-131.
- [41] Lu, Z.-L. and Dosher, B. (2013). *Visual psychophysics: From laboratory to theory*. MIT Press.
- [42] Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics*, 11(2):204–209.

- [43] Nguyen, A., Rothacher, Y., Lenggenhager, B., Brugger, P., and Kunz, A. (2018). Individual differences and impact of gender on curvature redirection thresholds. In *Proceedings of the 15th ACM Symposium on Applied Perception*, page 5. ACM.
- [44] Nilsson, N. C., Peck, T., Bruder, G., Hodgson, E., Serafin, S., Whitton, M., Steinicke, F., and Rosenberg, E. S. (2018). 15 years of research on redirected walking in immersive virtual environments. *IEEE computer graphics and applications*, 38(2):44–56.
- [45] Nilsson, N. C., Suma, E., Nordahl, R., Bolas, M., and Serafin, S. (2016). Estimation of detection thresholds for audiovisual rotation gains. In *Virtual Reality (VR), 2016 IEEE*, pages 241–242. IEEE.
- [46] Nussbaumer, L. L. (2012). *Inclusive design: A universal need*. Fairchild Books.
- [47] of Justice, D. (2010). Ada standards for accessible design. *Title III regulation*, 28.
- [48] Paludan, A., Elbaek, J., Mortensen, M., Zobbe, M., Nilsson, N. C., Nordahl, R., Reng, L., and Serafin, S. (2016). Disguising rotational gain for redirected walking in virtual reality: Effect of visual density. In *Virtual Reality (VR), 2016 IEEE*, pages 259–260. IEEE.
- [49] Park, D. H., Choi, W. S., Yoon, S. H., and Song, C. H. (2008). Anthropometry of asian eyelids by age. *Plastic and reconstructive surgery*, 121(4):1405–1413.
- [50] Patla, A. E. (1997). Understanding the roles of vision in the control of human locomotion. *Gait & Posture*, 5(1):54–69.
- [51] Peck, T. C., Fuchs, H., and Whitton, M. C. (2009). Evaluation of reorientation techniques and distractors for walking in large virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 15(3):383.
- [52] Peck, T. C., Fuchs, H., and Whitton, M. C. (2010). Improved redirection with distractors: A large-scale-real-walking locomotion interface and its effect on navigation in virtual environments. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 35–38. IEEE.
- [53] Peck, T. C., Fuchs, H., and Whitton, M. C. (2011). An evaluation of navigational ability comparing redirected free exploration with distractors to walking-in-place and joystick locomotion interfaces. In *Virtual Reality Conference (VR), 2011 IEEE*, pages 55–62. IEEE.
- [54] Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological methodology*, pages 111–163.
- [55] Razzaque, S., Kohn, Z., and Whitton, M. C. (2001). Redirected walking. In *Proceedings of EUROGRAPHICS*, volume 9, pages 105–106. Citeseer.
- [56] Ruddle, R. A. and Lessels, S. (2009). The benefits of using a walking interface to navigate virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 16(1):5.
- [57] Seno, T., Ito, H., and Sunaga, S. (2011). Attentional load inhibitsvection. *Attention, Perception, & Psychophysics*, 73(5):1467–1476.

- [58] Serafin, S., Nilsson, N. C., Sikstrom, E., De Goetzen, A., and Nordahl, R. (2013). Estimation of detection thresholds for acoustic based redirected walking techniques. In *Virtual Reality (VR), 2013 IEEE*, pages 161–162. IEEE.
- [59] Slater, M., Brogni, A., and Steed, A. (2003). Physiological responses to breaks in presence: A pilot study. In *Presence 2003: The 6th Annual International Workshop on Presence*, volume 157. Citeseer.
- [60] Slater, M., Usoh, M., and Steed, A. (1995). Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(3):201–219.
- [61] Slater, M. and Wilbur, S. (1997). A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6):603–616.
- [62] Stanney, K. M., Hale, K. S., Nahmens, I., and Kennedy, R. S. (2003). What to expect from immersive virtual environment exposure: Influences of gender, body mass index, and past experience. *Human Factors*, 45(3):504–520.
- [63] Steinicke, F., Bruder, G., Jerald, J., Frenz, H., and Lappe, M. (2010). Estimation of detection thresholds for redirected walking techniques. *IEEE transactions on visualization and computer graphics*, 16(1):17–27.
- [64] Steinicke, F., Bruder, G., Ropinski, T., and Hinrichs, K. (2008). Moving towards generally applicable redirected walking. In *Proceedings of the Virtual Reality International Conference (VRIC)*, pages 15–24. IEEE Press.
- [65] Steinicke, F., Visell, Y., Campos, J., and Lécuyer, A. (2013). *Human walking in virtual environments*. Springer.
- [66] Suma, E. A., Bruder, G., Steinicke, F., Krum, D. M., and Bolas, M. (2012). A taxonomy for deploying redirection techniques in immersive virtual environments. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, pages 43–46. IEEE.
- [67] Suma, E. A., Clark, S., Krum, D., Finkelstein, S., Bolas, M., and Warte, Z. (2011). Leveraging change blindness for redirection in virtual environments. In *Virtual Reality Conference (VR), 2011 IEEE*, pages 159–166. IEEE.
- [68] Trutoiu, L. C., Streuber, S., Mohler, B. J., Schulte-Pelkum, J., and Bülthoff, H. H. (2008). Tricking people into feeling like they are moving when they are not paying attention. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, pages 190–190. ACM.
- [69] Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., and Brooks Jr, F. P. (1999). Walking > walking-in-place > flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 359–364. ACM Press/Addison-Wesley Publishing Co.
- [70] Vasylevska, K., Kaufmann, H., Bolas, M., and Suma, E. A. (2013). Flexible spaces: Dynamic layout generation for infinite walking in virtual environments. In *3D user interfaces (3DUI), 2013 IEEE Symposium on*, pages 39–42. IEEE.

- [71] Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic bulletin & review*, 14(5):779–804.
- [72] Warren, W. H. and Kurtz, K. J. (1992). The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics*, 51(5):443–454.
- [73] Warren Jr, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., and Sahuc, S. (2001). Optic flow is used to control human walking. *Nature neuroscience*, 4(2):213.
- [74] Webb, N. A. and Griffin, M. J. (2003). Eye movement,vection, and motion sickness with foveal and peripheral vision. *Aviation, space, and environmental medicine*, 74(6):622–625.
- [75] Welch, R. B., Blackmon, T. T., Liu, A., Mellers, B. A., and Stark, L. W. (1996). The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence. *Presence: Teleoperators & Virtual Environments*, 5(3):263–273.
- [76] Willemsen, P., Gooch, A. A., Thompson, W. B., and Creem-Regehr, S. H. (2008). Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(1):91–101.
- [77] Wise, E. A., Price, D. D., Myers, C. D., Heft, M. W., and Robinson, M. E. (2002). Gender role expectations of pain: relationship to experimental pain perception. *Pain*, 96(3):335–342.
- [78] Yi, Y. and Merfeld, D. M. (2016). A quantitative confidence signal detection model: 1. fitting psychometric functions. *American Journal of Physiology-Heart and Circulatory Physiology*.
- [79] Zylberberg, A., Barttfeld, P., and Sigman, M. (2012). The construction of confidence in a perceptual decision. *Frontiers in integrative neuroscience*, 6:79.