DEPTH PERCEPTION UNDER MOTION PARALLAX SUBJECT TO GAIN IN VIRTUAL REALITY

XUE TENG

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTERS OF APPLIED SCIENCE

GRADUATE PROGRAM IN COMPUTER ENGINEERING
YORK UNIVERSITY
TORONTO, ONTARIO

2023

Abstract

Humans use visual, vestibular, kinesthetic and other cues to navigate through the world seamlessly. The way they are integrated has potentially significant implications for human perception of geometric layout in virtual reality (VR) systems. In particular, when observers experience a virtual motion while not moving in the real world, vestibular and kinesthetic cues conflict with visual cues and these conflicts between sensory cues could cause disorientation or discomfort. Therefore, it is important to study how conflicts between sensory systems affect human perception of the VR world. This thesis investigates the impact of mismatch between physical and portrayed head motion on the perception of object shape. We varied the gain between virtual and physical head motion and measured the effect on depth, distance and shape perception. Observers viewed a 'fold' stimulus, a convex dihedral angle formed by two irregularly-textured, wall-oriented planes joined at a common vertical edge. Stimuli were rendered and presented using head-mounted displays. On each trial, observers adjusted the angle of the fold till the two joined planes appeared perpendicular while keeping the head still (stationary condition) or moving it rhythmically from side to side (motion parallax condition). For motion parallax conditions, the virtual

motion was scaled relative to observers' change in vantage point by gain factors of 1/2, 2/3, 4/5, 1, 5/4, 3/2 and 2. In following experiments, observers were also asked to match a virtual pole to the remembered position of the apex of the fold. Our results showed that under monocular viewing, both depth and distance settings decreased with increasing gain, especially at close distances. The average effect sizes of gain were up to -0.061 m/gain unit and -0.40 m/gain unit on depth and distance, respectively, when measured on a standard fold with depth of 1 m. Observers experienced less distortions than predicted from a geometric model and very little depth distortion (not statistically significant effect of gain) under binocular viewing. The distance distortion caused by gain was reduced by up to 56.6% compared to monocular viewing. Binocular cues to depth and distance and large distance (at 6 m) enhance humans' tolerance to visual and kinesthetic mismatch.

Acknowledgements

I would like to express my deepest gratitude to my supervisor and advisor, Prof. Robert S. Allison and Prof. Laurie M. Wilcox, for their invaluable guidance and support throughout my Master's program. Their expertise and encouragement helped me complete this research and write this thesis. They also inspired me in so many ways that I can not keep count of. Without them, this journey would not have been possible. I would like to thank Prof. I. Scott MacKenzie and Prof. Taylor W. Cleworth for serving on my thesis committee and providing helpful feedback and suggestions. I would also like to thank my family and friends for their love and support during this process. Finally, I would like to thank all of the participants in this study for their time and willingness to share their experiences. This work would not have been possible without their contribution.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Virtual reality (VR) and augmented reality (AR) create an immersive experience, which provides the users a full-scale virtual world of unlimited possibilities. A wide range of AR/VR applications have entered people's everyday life. It has changed our common practice in education (Kavanagh et al. 2017; Cheng, Yang, and Andersen 2017), gaming (Lugrin et al. 2018) and many other areas (Boyd and Koles 2019). When interacting with objects in VR, whether or not the intended object can be rendered at the intended spatial location is essential for providing a satisfying user experience. Perception of objects, e.g., shape and depth, in VR has been one of the most important problems considered in the literature. Swan et al. (Swan et al. 2007), Waller and Richardson (Waller and Richardson 2008), and Thompson et al. (Thompson et al. 2011) did rigorous studies of distance estimation in virtual environments. Armbrüster et

al. and many others showed that people tend to underestimate virtual distances (Eggleston, Janson, and Aldrich 1996; Philbeck and Loomis 1997; Armbrster et al. 2008).

Humans use visual, vestibular, kinesthetic and other cues to effectively navigate through the world. Therefore, conflict between these sources of information has potentially significant implications for human perception of geometric layout. Virtual environments present simulated visual and other sensory content to users and, either by design or due to technical shortcomings, create distortions within and between sensory systems (Welch and Cohen 1988; Smith, Salvendy, and Koubek 1997). These sensory distortions might impact user well being and performance in virtual environments on tasks which require precise and accurate localization of objects and surfaces. In extreme cases, conflict and inconsistency may result in VR sickness. Symptoms include nausea, pallor, sweating, headache and disorientation (Stanney, Mourant, and Kennedy 1998; Stanney and Kennedy 1997). Although observers can adapt to the sensory discord experienced in virtual environments, they may then exhibit post-use aftereffects (Anstis 1995). To ensure the safety and effectiveness of VR/AR, it is of vital importance that the impact of sensory distortions on users is fully studied.

Depth perception is the ability to perceive the three dimensional (3D) environment and particularly the 3D shape of objects (Todd and Norman 2003). A variety of depth cues contribute to depth perception, such as binocular disparity, convergence, accommodation, occlusion and motion parallax (Braunstein et al. 1986; Surdick et al. 1994). Motion parallax is the relative change in direction

between stationary objects at different distances created by observers' motion (Rogers and Graham 1979a; Rogers and Graham 1979a). Previous work has found that introducing motion parallax gain differences between physical and virtual head movement had little effect on distance perception (Cutone, Wilcox, and Allison 2020). Now I am interested to see if the same distortion will affect observers' perception of depth. I varied the gain between virtual and physical head motion and measured the effect on depth perception. My results show that gain only had an effect on depth perception with monocular viewing when distance range was within 1.3 – 1.5 m. However, when distance range extended to 1.5 – 6.0 m, no such interaction effect was observed.

1.2 Contribution

In this study, my primary focus is how the conflict between physical and head mounted display (HMD) rendered self-motion impacts object shape perception. This thesis contributes to the fields of depth perception and virtual reality by

- Evaluating the impact of conflict between kinesthetic and virtual head motion on the perception of object shape.
- Proposing a mathematical model to predict the relationship between depth, distance and gain.

1.3 Thesis Outline

The rest of the thesis is organized as follows. Chapter 2 gives a brief introduction of related work. Detailed experiment methods and theoretical modeling are given in Chapter 3. Results and discussion are given in Chapter 4. Finally, conclusions and future work are given in Chapter 5.

Chapter 2

Background and Related Work

One of the most important tasks of human perception is to identify the three-dimensional (3D) shape of objects. Depth perception plays an important role in making these judgements and hence is an important factor in designing studies of 3D object perception. This is particularly important for the development of VR systems which often require the rendering of realistic 3D structure using head-tracked, stereoscopic displays.

2.1 Depth Perception

Depth perception is the visual ability to perceive the world in three dimensions (3D), rather than the 2D images formed on our retinas. Depth perception is traditionally thought to arise from a variety of depth cues. These are typically

classified as either binocular or monocular depending on the number of eyes required (Goldstein and Brockmole 2016). A illustrative example is given in Figure 2.1. People can easily perceive the relative depth of objects in the photo. The perception could originate from occlusion, relative size, shadow, perspective and other cues.

Alternative approaches exist for example (Gibson and Carmichael 1966), where perception is believed to be best understood in terms of systems rather than of a variety of independently operating sense organs. In this thesis however, I consider the problem formulation as the combination of cues. The cue approach to depth perception focuses on identifying information in the retinal image that is correlated with depth in the scene (Howard 2012).

Monocular cues include relative size, texture gradient, occlusion, linear perspective, contrast differences, and motion parallax (Burton 1945). Binocular cues, including stereopsis and eye convergence, yield depth from binocular vision through exploitation of parallax.

Figure 2.1: Scene taken in Lassonde Building 2001C

Monocular Cues

Monocular cues are the cues that can be obtained from the visual information available from only one eye. Monocular cues include accommodation, which is the change in the shape of the lens that occurs when we focus on objects at various distances; pictorial cues, which is depth information that can be depicted in a two-dimensional picture; and movement-based cues, which are based on depth information created by motion (Goldstein and Brockmole 2016).

Pictorial cues are sources of depth information that can be depicted in a picture, which includes occlusion, relative size, etc. Occlusion occurs when one object (completely or partially) blocks another from view. The partially hidden object is seen as farther away. In Figure 2.2, we can clearly tell which cakes are in front simply because they block the others. Occlusion does not provide

information about the quantitative or metric depth between objects but only indicates ordinal distance. However, since we are very familiar with the size of the cakes sold in super markets, their size can also potentially be used as cues. The cue of relative size is based on the fact that when two objects are of equal physical size, the one that is farther away will take up less of your field of view than the one that is closer. In Figure 2.2, the chocolate cake at bottom right (circled in blue) is perceived to be further away, consistent with its smaller image than the others, e.g., the pumpkin cake (circled in orange) at bottom left.

Figure 2.2: Relative size and familiar size cues.

As an observer moves, optical flow on one hand arises due to the relative motion between the observer and the objects (Gibson 1950); on the other hand, motion parallax created by the change in the relative position of two or more stationary objects gives hints about their relative distance (Rogers and Graham 1979a). Motion parallax can provide absolute depth information if direction

and velocity of movement is known (Ferris 1972).

Depth perception is a complex process, where multiple cues are typically used together. For instance in Figure 2.3, the images of parallel lines between paving stones converge in the distance, which gives a cue for depth, known as linear perspective. The trees occluding the building indicating that buildings are further away, i.e., an **occlusion cue**. For the cars in the parking lot, the images of some are bigger, while others are small. This provides an indication of their relative distance based on **relative size cues**. If the observer moves, the relative position of different objects will change. Thus, the **motion parallax cue** may be used to perceive depth.

Figure 2.3: Mixed depth cues.

Binocular Cues

Binocular cues depend on the images from both of the eyes. Our eyes are placed about 6 cm apart on average for adults and hence they each get a different view of the scene and the objects in the environment. Stereopsis usually refers to the perception of solidity and 3-dimensional structure. Because human eyes are located at the front, the overlap of the view from both eyes is significant. However, there are still differences in the images at each view point. These differences are called binocular disparity. The information provided by disparity can assist depth perception (Howard and Rogers 1995; Goldstein and Brockmole 2016).

Disparities are first processed in the visual cortex of the brain to yield depth perception (Ohzawa and Freeman 1986). While binocular disparities are naturally present when viewing a real 3-dimensional scene with two eyes, they can also be simulated by artificially presenting two different images separately to each eye using a method called stereoscopy. The perception of depth and volume in such cases is also referred to as "stereoscopic depth" (Goldstein and Brockmole 2016). Figure 2.4 gives an example of a stereoscopic pair.

2.2 Depth Perception in Virtual Reality

2.2.1 VR Devices

Modern commercial VR systems usually employ a head-mounted display (HMD) design (e.g., Oculus, HTC Vive). A head-mounted display (HMD) is a display

(a) Left eye image for parallel fusion. (b) Right eye image. (c) Left eye image for cross fusion.

Figure 2.4: Stereoscopic photograph (Goldstein and Brockmole 2016).

device, worn on the head or as part of a helmet. The device typically provides separate left and right images and thereby can present stereoscopic views to the user. The rendered views for both eyes usually need to be updated in real time according to the user's head position and orientation, depending on the use case.

The VR device used in this thesis, the Oculus Rift S is built with six degrees of freedom (6 DoF) tracking. Together with 80 Hz refresh rate, the device can generate realistic immersive VR experiences.

2.2.2 Depth Perception

Depth perception plays an important role in the user experience of VR. A large amount of work has been conducted to understand how humans perceive 3D shape. Rogers et. al. showed that parallax information can be a subtle and powerful cue to the shape and relative depth of three-dimensional surfaces (Rogers and Graham 1979a). Ono et. al. showed that depth perception arising from motion parallax is contingent upon distance information (Ono, Rivest,

and Ono 1986). Most cues to depth vary with egocentric distance of the object. The ability to perceive consistent depth at different distances is known as depth constancy (Wallach and Zuckerman 1963). Todd and Norman (2003) found that observers' judgments of 3D metric structure can be systematically distorted and often exhibit large failures of constancy over changes in viewing distance and/or orientation.

Humans rapidly adjust the calibration of their motions (walking, turning and throwing) and adapt to changing circumstances such as unexpected acceleration/deceleration in a real environment (Rieser et al. 1995). These circumstances are similar to first-person perspective VR tasks such as roller coaster video games and redirected walking/navigating which can be seen as changing the relationship between real motion and virtual motion (motion gain). Depth perception often co-exists with awareness of the change in motion gain (Cornilleau-Pérès and Gielen 1996). Depth perception under motion gain also affects user experience and performance (Campagnoli, Hung, and Domini 2022). Selzer, Larrea, and Castro (2022) studied virtual walking with large translational gain and discovered that humans could tolerate translational gains up to 1.5 times the real motion without the perception of presence being impacted. Brument et al. (2021) studied rotation gains under the combination of translational and rotational motions and found that rotation gains are less perceivable under low rotational motion speed and translational motion. Steinicke et al. (2009) found that rotation and translational gain can be scaled up by 49% and 26%, respectively or reduced 20% and 14% without being noticed.

Motivated by these interesting findings, in this thesis, a series of experiments

was designed to study the systematic disruption of perceived shape resulting from the mismatch (gain) between an observer's head motion and the virtual motion. Under such gain distortions, observers could rely on combined or compromised visual content in the VR and kinesthesia to perceive self motion. These two competing sources may indicate different self motion and consequently affect the perception of object shape. Gain levels were varied to study how motion distortion affects depth perception under different viewing and head motion conditions.

Chapter 3

Methodology

In this chapter, Section 3.1 and 3.2 describe the apparatus and virtual environment, respectively. Methods for the experiments are described in Sections 3.3, 3.4 and 3.5. Furthermore, a mathematical model is proposed to predict and analyze the impact of the experimentally manipulated factors in Section 3.6.

3.1 Apparatus

An Oculus Rift S head-mounted display (HMD) was used to present the virtual environments (VEs), as shown in Figure 3.1(a). The headset contains a pair of LCD displays with a resolution of 1280 $\times$ 1440 pixels per eye. The device refresh rate was 80 Hz and diagonal field of view was 115°. I set the frame rate as 60 Hz due to the limit of the power source of the laptop. This model has built-in cameras for 6 DoF (Degrees of Freedom) location/orientation tracking, which was used to update virtual head pose in real time.

The left and right hand controllers were used for participant input, a close shot is shown in Figure 3.1(b). The button X on the left hand controller was used to initialize the virtual position/orientation of the virtual environment to align with the participant's location at the moment it was pressed. The left-right axis of the joystick on the right hand controller was used for angle adjustment, shown as the white arrow in Figure 3.1(b). Observers tilted it to the left for adjusting smaller angles and right for bigger angles. The forward-backward axis was used for distance adjustment, shown as the blue arrow in Figure 3.1(b). Observers tilted it forward to increase distance and backward to reduce distance. Tilting in other directions was not taken as an input. The button A of the right hand controller was used for activating the next trial for stationary conditions and data registration for all conditions. The button B was used for early termination, in case that the observers experienced severe discomfort. The joystick of the left hand controller, other buttons, triggers and hand tracking feature were not used in this study.

Simulations were generated with PsychXR (Cutone and Wilcox 2021) and PsychoPy (Peirce 2007) packages, which allowed us to manipulate the motion gain in VR. The experiment setup is shown in Figure 3.1(c). The insert at top left shows the screenshot of the scene in PsychoPy when observers sways to the right.

(a) (b) (c)

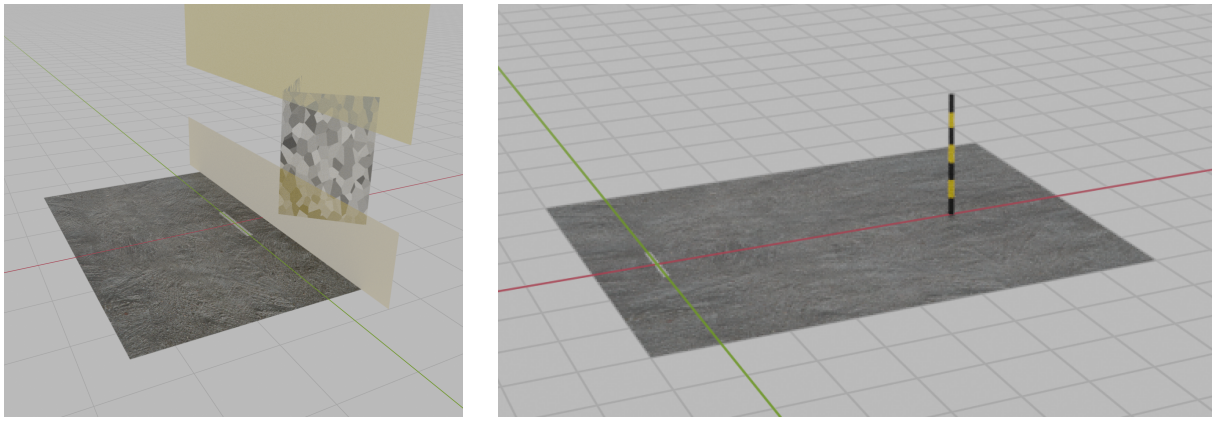
Figure 3.1: Experiment setup: (a) Oculus Rift S with hand controllers; (b) joystick axes and buttons of the hand controllers; (c) over-the-shoulder view.

3.2 Virtual environment

Fold Scene. A snow covered mountain skybox was centered around the viewpoint regardless of translations, providing the illusion that the scene was at infinite distance. The test stimulus was a 'fold' object, which was a convex dihedral angle formed by two wall-oriented planes connected at a common vertical edge as shown in Figure 3.2(a). Figure 3.3(a) shows the HMD view, taken by Oculus Mirror. The height and width of the planes forming the fold were 3 meters and $\sqrt{2} \cdot s$ meters, where s is the scaling factor of width. The scaling factor, s, was adopted to avoid the observers using the distance between left and right edges for angle judgement. Both planes were covered by a Voronoi texture image generated by Blender (Community 2018) (gray scale, F1 Euclidean with no lighting source). The Voronoi density at each distance was set based on visual comparisons between the rendered distances of the fold, to equate the size of the retinal projection. The upper and lower edges of the fold were occluded

from the observers' vantage point, so that the participants could not utilize the intersections of the top or bottom edges of the test object to perform the task. The occluding aperture was formed by two frontal parallel planes with a gap between. They were set at a distance of 0.8 m, covered by a black texture. The aperture in Figure 3.2(a) is shown as semi-transparent for illustrative purposes. The observers stood on top of a 10×3 m^2 ground plane. To guarantee that observers could not use height in the field for the task, the shorter dimension of the ground plane ended at the position of aperture, so that it did not appear as the background of the fold. The ground plane was covered by a igneous-like texture with no regular pattern. A white bar, sized 20×3.5 cm^2, was placed on the ground plane, indicating where to stand and thus the origin of head position. In order to maintain consistent distances between the participant and the fold stimuli, a distance range restriction of 7 cm about the ground bar along the anterior-posterior axis (front and back, 3.5 cm for each direction) was applied. The VR environment was dimmed once the observers went out of range to indicate that a distance correction was needed. In physical space, along a medial-lateral axis, a minimum of 20 cm was required to allow robust motion parallax from self motion. A square guardian of 3×3 m^2 (or 16.5×16.5 virtual marks in Oculus App) was drawn to keep the scale in the VR consistent with the real world. Note that it's not required that the physical space has as much space as the VR world.

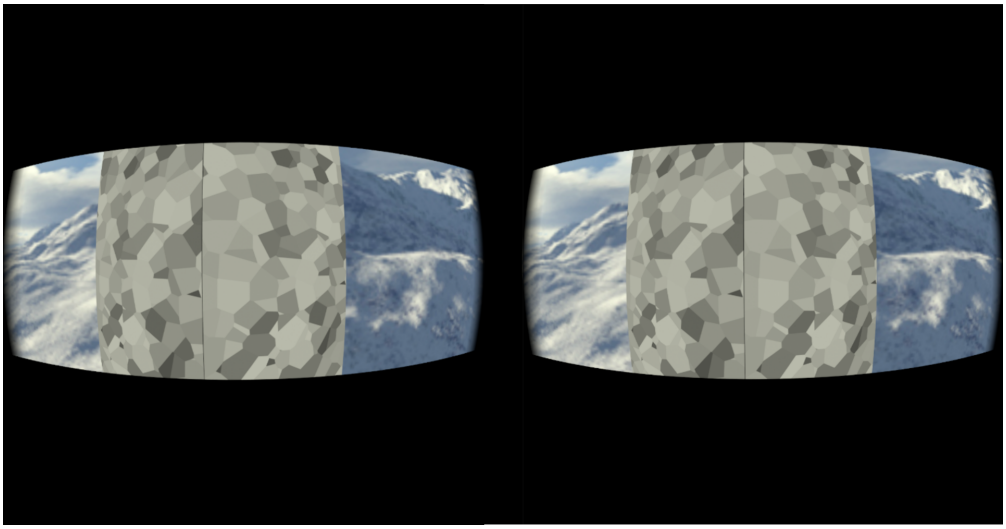
Pole Scene. A vertical pole with black and yellow stripes was positioned directly in front of the observer. The pole had a diameter of 5 cm, height 2.2 m and its position was adjustable on the anterior-posterior axis, as shown in



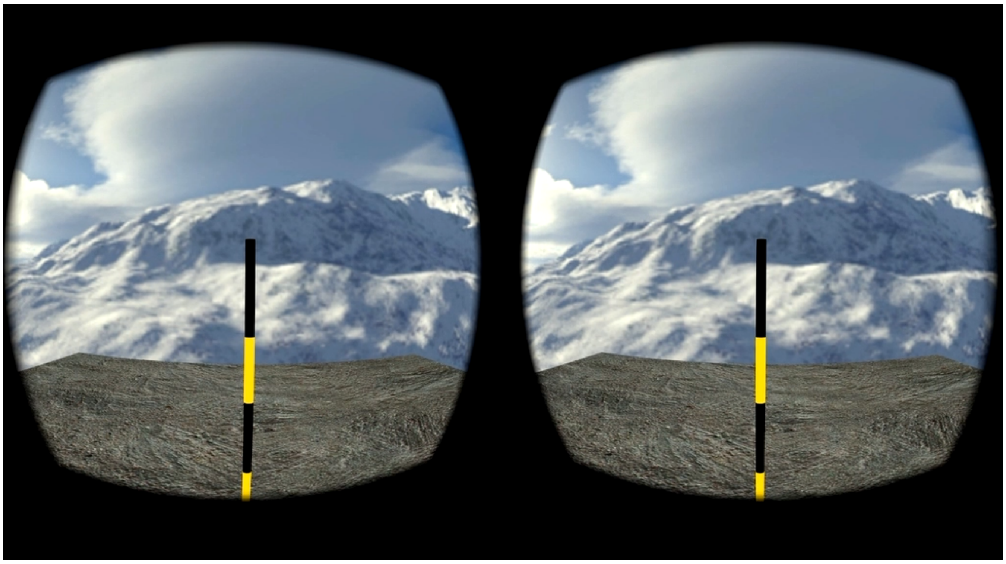
(a) (b)

Figure 3.2: 3D depictions of (a) fold scene and (b) pole scene, both generated using Blender.

Figure 3.2(b). In order to give the observers as many contextual cues as possible in this scene, apertures were removed. The ground plane changed to 5×10 m^2 to provide additional ground-based distance cues. Figure 3.3(b) shows the hmd view of this VE.



(a)



(b)

Figure 3.3: HMD views of VEs. Top: fold scene, bottom: pole scene.

3.3 Experiment 1

The fold scene was used as the VE to evaluate the effect of gain on the perception of object depth.

3.3.1 Participants

All nine observers (four female and five male, age range 22 – 58 years, average = 34.3) had normal or corrected to normal vision. Among them, four were experienced psychophysical observers (including the author) and five were naive observers. In the group of experienced observers, one was naive to the purpose of this experiment as were all the naive observers.

3.3.2 Task and Procedure

The task in this experiment was to adjust the angle of the fold until the surfaces appeared perpendicular. In each trial, observers were asked to adjust the angle between the planes using the left-right axis of the joystick on the right hand controller while maintaining their motion status. For motion parallax conditions, observers swayed at 0.5 Hz to a 60 beats-per-minute (BPM) metronome audio. Motion parallax indicators provided feedback to ensure that the observers moved at least 10 cm to each side of the center along the medial-lateral axis (totalling 20 cm physical distance from left to right). The indicators showed the observers how far they needed to move when doing motion parallax. The indicators were red light objects with a Gaussian gradient, rendered at the far side of each eye's view to avoid blocking the main view when activated (translation $\geq \pm$ 10 cm).

After one full cycle of motion parallax, the target fold stimulus was presented. To vary the resulting virtual fold scene, this kinesthetic motion parallax was tracked and multiplied by the gain to produce the virtual viewpoint used to render the scene. Viewing distance and width of the planes forming the fold were varied so the observers could not rely on them to judge depth across trials. Stationary trials were indicated by the absence of the metronome audio, each trial was activated by pressing the button A on the right hand controller and observers were instructed to stand still while doing the adjustment. When they were satisfied with the adjustment, they pressed the button A on the right hand controller to register their setting.

3.3.3 Design

The experiment was divided into two parts by Motion Status: stationary conditions as a baseline and motion parallax conditions. Both parts employed a fully within-subjects factorial design. The independent variables were as follows:

- Motion Status (motion parallax, stationary)
- Gain (motion parallax with $g = \frac{1}{2}, \frac{2}{3}, \frac{4}{5}, 1, \frac{5}{4}, \frac{3}{2}, 2$ and stationary with $g = 1$ which repeated 7 times.)
- Viewing (monocular, binocular)
- Fold Distance (1.3 m, 1.4 m, 1.5 m)
- Width ($w = \sqrt{2}, 1.125\sqrt{2}, 1.25\sqrt{2}$ m - with scaling factor $s = 1, 1.125, 1.25$, respectively)

where monocular is defined as showing the virtual scene to the right eye while the view in the left eye was blank (blackout).

Each participant completed 4 blocks (2 Viewing conditions $\times$ 2 Motion Statuses) of 63 motion trials (7 Gains $\times$ 3 Fold Distances $\times$ 3 Widths) for a total of 252 trials. The order of blocks was randomized across observers.
The dependent variables were:

- Self Motion (head position with respect to time. This is measured by the built-in head tracking of Oculus device.)
- Angle (input from the observers)
- Depth (distance between the apex and bottom edge of the fold, derived from Angle and Width. See Equation 4.2 for details.)

Due to the COVID-19 pandemic, relevant health measures had to be implemented to ensure the safety of participants. Participants were given the option to conduct the experiment at York University or at their own home (by themselves). For experiments at York University, I strictly followed the public health measures, e.g., social distancing, equipment sanitation, etc. At the same time, I ensured that participants followed the correct motion requirements. For experiments at the participants' home, the same HMD apparatus and laptop were delivered to the observers. Self Motion produced by all observers was confirmed against expectations. Thus, the quality and consistency of experiments was ensured.

3.4 Experiment 2

The fold and pole scenes were used to assess the effect of motion gain on perceived depth and distance, respectively.

3.4.1 Participants

Nine participants completed the second experiment (three female and six male, age range 27–58, average = 38.0). Six were from the previous group who completed Experiment 1 (three experienced and three naive) along with three additional experienced participants who were naive to the purpose of the experiment.

3.4.2 Task and Procedure

On each trial, the Angle adjustment task was tested under the same conditions as in Experiment 1. After the input of the Angle adjustment, the fold scene gradually faded into a blackout image over 1 second, then the pole scene was presented. The task in the pole scene was to adjust the distance of the pole until it appeared at the same distance as the peak of the fold that had just been seen. The pole was adjusted through the forward-backward axis of the joystick on the right hand controller. Once the adjustment was done, observers registered their response by pressing the button A as in the Angle adjustment task. The distance matching task was always binocular and stationary regardless of the test condition in the angle adjustment task; the metronome audio stopped for motion parallax trials and the ground plane was changed to 10×10 m^2. These

changes were designed so that the observers had enough perceptual distance cues for the distance matching task. When the distance matching task was completed, the trial was complete and the next trial was activated according to the Motion Status and Viewing condition of the angle adjustment task. The pole's initial position at the start of each distance matching task was set to be either in between 0.4–0.6 m or 2.5–6 m from the origin (i.e., observer).

3.4.3 Design

This experiment was also divided into two parts based on Motion Status and both parts employed a fully within-subjects factorial design. Independent variables and their levels were the same as in Experiment 1 but the seven repetitions of the stationary trials ($g = 1$) was removed in Experiment 2. Each participant completed 2 blocks of 9 stationary trials and 2 blocks of 63 motion parallax trials:

1 Gain(stationary) $\times$ 3 Fold Distances $\times$ 2 Viewings $\times$ 3 Widths + 7 Gains (motion parallax)$\times$ 3 Fold Distances$\times$ 2 Viewings $\times$ 3 Widths = 144 per participant.

The dependent variables were:

- Self Motion (same as in Experiment 1)
- Angle (same as in Experiment 1)
- Depth (same as in Experiment 1)
- Pole Distance (input from observers)

3.5 Experiment 3

In this experiment, a further and larger range of Fold Distances was tested to confirm if the effects found in Experiment 2 persisted. For Gain (g levels), I included a subset from Experiment 1 and 2.

3.5.1 Participants

Six out of the nine participants from Experiment 2 completed this experiment (two female and four male, age range 27 – 38, average = 32.3). Among them, three were experienced observers and three naive. Four of them also completed Experiment 1.

3.5.2 Procedure

The procedure of Experiment 3 was the same as Experiment 2.

3.5.3 Design

This experiment employed a similar design as the Experiment 2. The independent variables were the same as in Experiment 2 but some of the levels were different:

- Motion Status (same as in Experiment 1 and 2)
- Gain (motion parallax with $g = \frac{1}{2}$, $\frac{2}{3}$, 1, $\frac{3}{2}$, 2 and stationary with $g = 1$.)
- Viewing (same as in Experiment 1 and 2)

- Fold Distance (1.5 m, 3.0 m, 6.0 m)
- Width ($w = \sqrt{2}\cdot$ Fold Distance, $1.125\sqrt{2}\cdot$ Fold Distance, $1.25\sqrt{2}\cdot$ Fold Distance - with scaling factor $s = 1\cdot$Fold Distance, $1.125\cdot$Fold Distance, $1.25\cdot$Fold Distance, respectively)

Each participant completed 2 blocks of 9 stationary trials and 2 blocks of 45 motion trials:
1 Gain(stationary) $\times$ 3 Fold Distances $\times$ 2 Viewings $\times$ 3 Widths + 5 Gains (motion parallax)$\times$ 3 Fold Distances$\times$ 2 Viewings $\times$ 3 Widths = 108 per participant.
The dependent variables were the same as in Experiment 2.

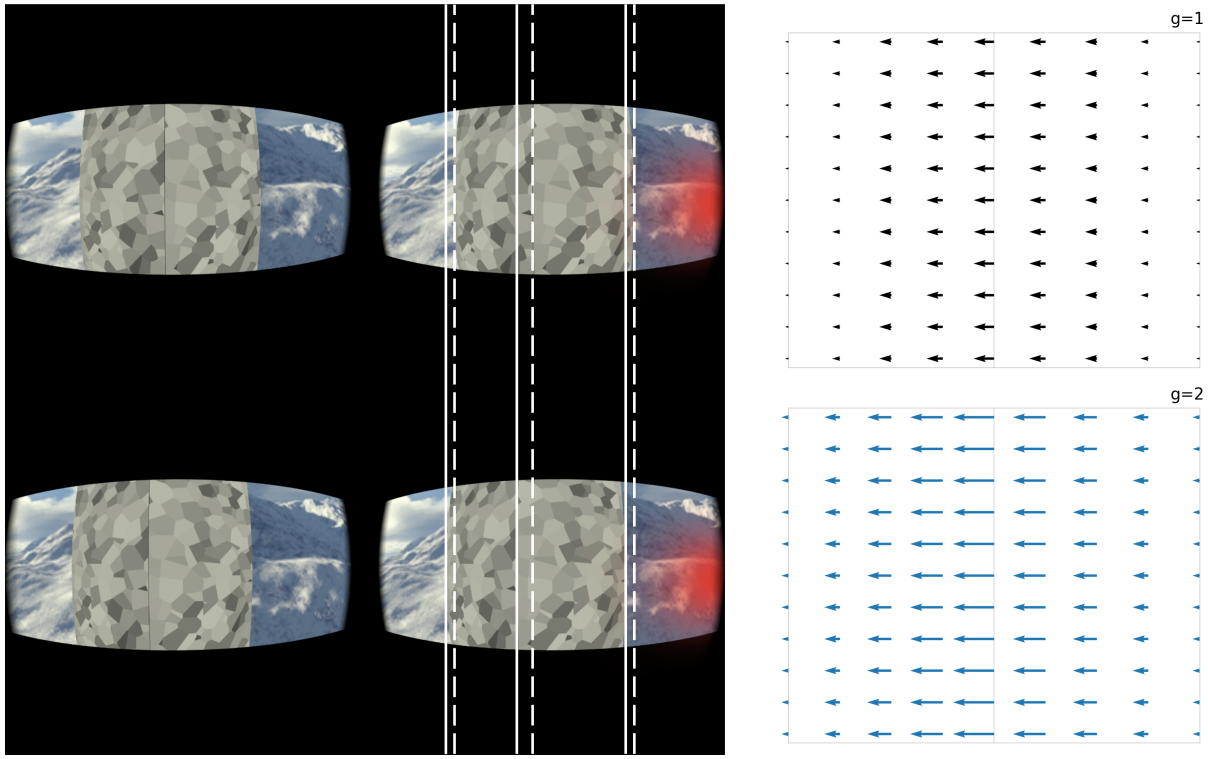
3.6 Theoretical Analysis

3.6.1 Gain Distortion

The motion in the VR scene was manipulated with a factor g, defined as gain:

$$g = \frac{m_g}{m_s} \tag{3.1}$$

which determines how the virtual camera used to render the VR image moves relative to the observer's actual physical Self Motion. The rendered virtual motion is denoted m_g, while m_s is the physical/kinesthetic Self Motion. The gain factor describes the difference between the VR rendered scene and real physical world. When $g = 1$, $m_g = m_s$, it indicates that there was no gain distortion.



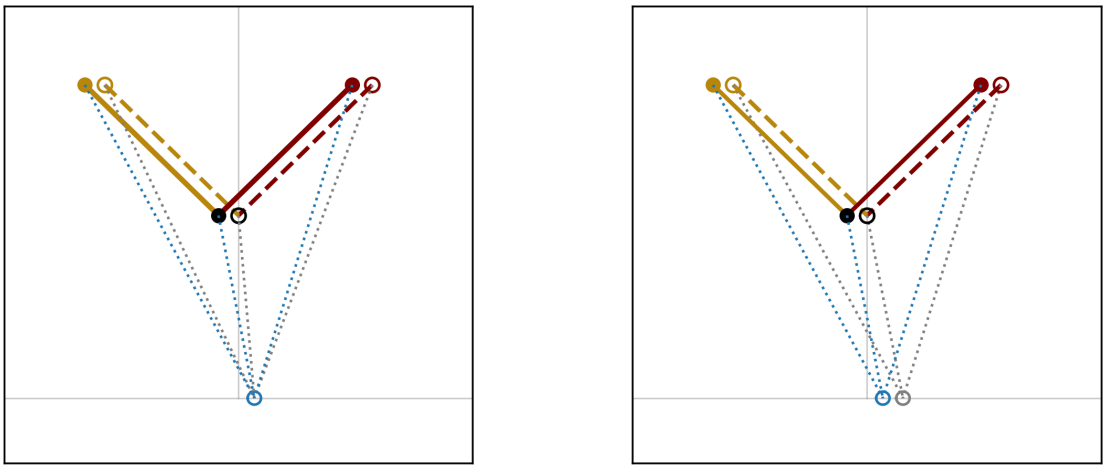
(a) (b)

Figure 3.4: Impact of gain when observers swayed to the right.

Figure 3.4 shows the impact of gain under two example gain situations, where the 'fold' is presented in front of the observer making a rightward motion. Top row shows the situation with $g = 1$ (without gain distortion), bottom shows the situation with $g = 2$ distortion. Figure 3.4(a) shows the HMD views. The white lines highlight the vertical edges in right eye's view. Dashed lines highlight the edges under $g = 1$ and solid lines highlight the edges under $g = 2$. Both situations demonstrate the scenes under the same amount of motion parallax (10 cm) to the right side boundary and thus the red light indicator was activated. Figure 3.4(b) shows the corresponding optical flow diagrams, each vector illustrates the pixel displacement of a point on the fold stimuli between two time instants. With $g = 2$ shown as the bottom row, the rendered motion is twice the amount of Self Motion, causing exaggerated flow vectors and the rendered stimulus moves more than the observer expects.

Figure 3.5 shows top-down views of the test stimulus and the physical or virtual eye position during rightward Self Motion. The dashed red/yellow lines represent the left/right planes of the fold with no motion gain manipulation ($g = 1$) while the solid red/yellow lines show them with $g = 2$ gain distortion. The open and filled markers denote these two conditions respectively, showing the far left, middle peak and far right points of the fold. Circles at the bottom denote the position of observer's eye. Grey dotted lines are visual lines for when $g = 1$ while blue dotted lines correspond to $g = 2$. If m_o is defined as the object motion. $m_{o,1}$ and $m_{o,2}$ represent the object motions in the virtual scene under the two gains.

The left diagram, Figure 3.5(a), shows that Self Motion m_s produced two



(a) (b)

Figure 3.5: Top-down view of fold stimulus and human eye: (a) different VR images are produced by the same Self Motion; (b) the same VR images are produced by different Self Motions.

different object motions under different gains ($m_{o,1} \neq m_{o,2}$). For $g = 1$ situation (grey), as the observer sways, the VR rendered fold remains at the original position ($m_{o,1} = 0$), as in the real world. For $g = 2$ situation (blue), the object points moves away from the origin by the same amount of Self Motion ($m_{o,2} = m_s$), producing twice the translation between the object and observer compared to the $g = 1$ situation.

Another set of situations is demonstrated in Figure 3.5(b) where $g = 1$ and $g = 2$ share the same VR image, i.e., parallel visual lines. The $g = 1$ situation (grey) shows twice the physical Self Motion as in Figure 3.5(a) ($m_{s,1} = 2m_s$), while the Self Motion in $g = 2$ situation (blue) is identical to the case in the left diagram ($m_{s,2} = m_s$). It's obvious to see that for the same viewpoint, observers' visual angles are the same. This is because both VR images are rendered based on the same amount of virtual motion m_g:
By the definition of HMD image rendering, virtual motion can be expressed as the sum of Self Motion and object motion,

$$m_g = m_s + m_o \tag{3.2}$$

For the $g = 1$ situation in Figure 3.5(b), object motion is 0 since it stays stationary, the virtual motion in the image can be represented as:

$$m_{g,1} = m_{s,1} + m_{o,1} = 2m_s + 0 = 2m_s$$

While for $g = 2$ situation:

$$m_{g,2} = m_{s,2} + m_{o,2} = m_s + m_s = 2m_s = m_{g,1}$$

Based on the above discussion, it can be concluded that parallel visual lines

between two viewing conditions are equivalent so that their m_g and VR images are the same. The above interpretations represent theoretical cases where gain manipulation, object and Self Motion are all perceived accurately. Next we will see in Section 3.6.2 that the scale of perception is a potential perceptual factor.

3.6.2 Geometry

In this section we analyze the perception of gain distortions in the physical space. Assume f_s is the Self Motion factor, defined as:

$$f_s = \frac{m_s^p}{m_s} \tag{3.3}$$

which describes how observers judge their Self Motion, where m_s is the actual physical self movement, as previously defined in Equation (3.1), m_s^p refers to the perceived self motion. Similarly, the object motion factor is defined as:

$$f_o = \frac{m_o^p}{m_s} \tag{3.4}$$

where m_o^p refers to the perceived object motion. f_o represents the factor relating perceived object motion to the physical motion made by observers, shown as the displacement between open and closed markers of object points in Figure 3.5. The sum of f_s and f_o accounts for the total perceived gain g^p:

$$g^p = f_s + f_o \tag{3.5}$$

The physical meaning of Equation (3.5) is that if the observer perceives the object stays stationary with no virtual object motion, $f_o = 0$, all motion is accounted as perceived self motion:

$$g^p = f_s$$

On the other hand, if observers attribute all the gain distortion to object motion in the virtual image and no distortion in perceived self motion, $f_s = 1$:

$$g^p = f_o + 1$$

By multiplying both sides of Equation (3.5) with m_s, it is reformulated into

$$g^p \cdot m_s = f_s \cdot m_s + f_o \cdot m_s \tag{3.6}$$

where $g^p \cdot m_s = m_g^p$, which is the amount of perceived virtual motion. Together with Equations (3.1), (3.3) and (3.4), Equation (3.6) can be rewritten as:

$$m_g^p = \frac{m_s^p}{m_s} \cdot m_s + \frac{m_o^p}{m_s} \cdot m_s = m_s^p + m_o^p \tag{3.7}$$

If observers perceive the total motions correctly, $m_g^p = m_g$, $m_s^p = m_s$ and $m_o^p = m_o$. Thus, Equation (3.7) becomes:

$$m_g = m_s + m_o$$

which aligns with Equation (3.2).

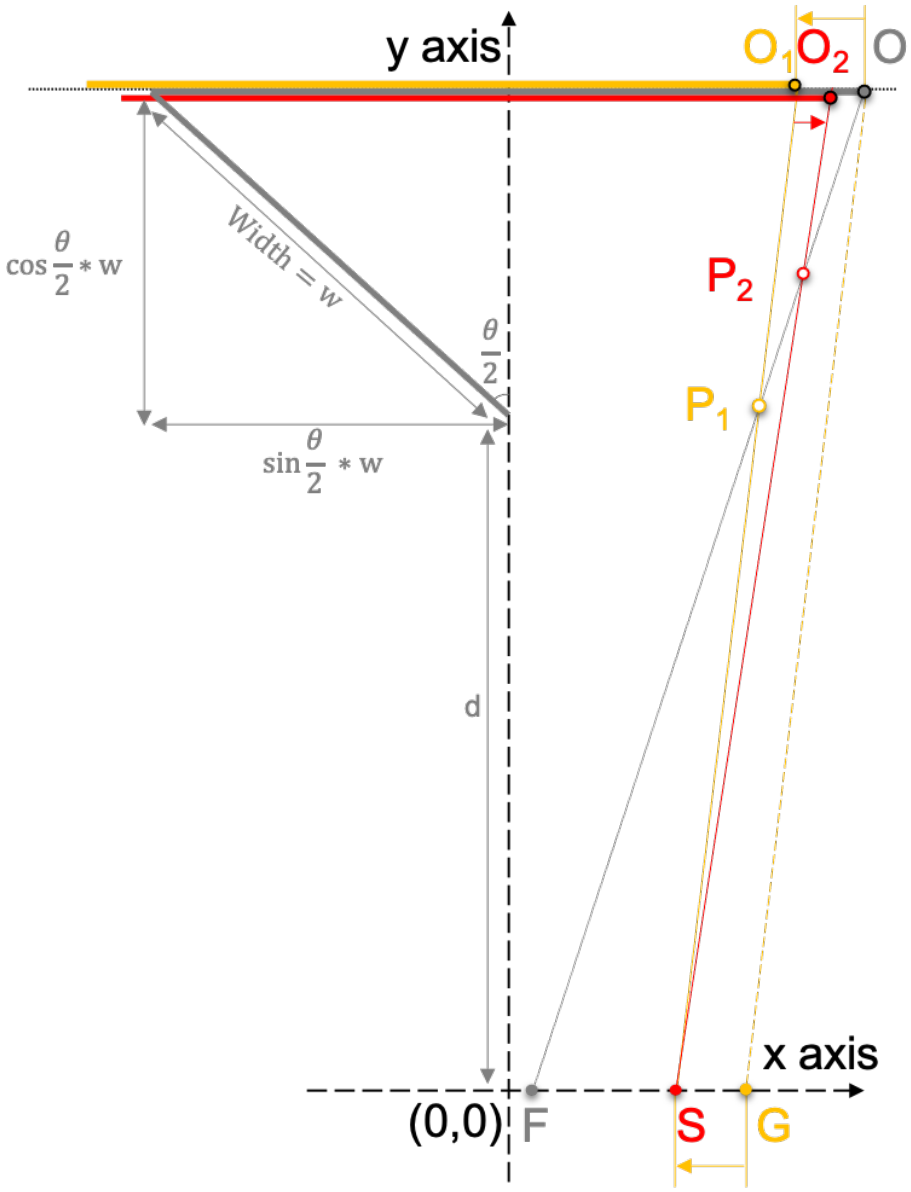


Figure 3.6: Possible perceptions under motion gain distortion

Figure 3.6 demonstrates possible perceptions of the right half of a fold under a motion gain manipulation. The x and y axes are the medial-lateral and anterior-posterior axes respectively. Solid markers (F, S and G) are eye positions, open markers (P_1 and P_2) are predicted object positions based on

Equation (3.7) and black outlined markers (O_1, O_2 and O) are the projected positions of P_1 and P_2, at a certain distance (dotted line). For a fold with angle θ located at a distance of d, the objective peak and far right object point O can be calculated as $(0, d)$ and $(\sin\frac{\theta}{2} \cdot w, d + \cos\frac{\theta}{2} \cdot w)$, where w is the width of each plane forming the fold stimulus. Defining the inter-pupillary distance as IPD, the distance between the midpoint of the eyes in the head and right eye is $\frac{IPD}{2}$. Now the initial vantage point F $(0, \frac{IPD}{2})$ is taken as the position of right eye when head position is at the origin $(0, 0)$. When standing at the origin, the observers' translation is 0 ($m_s = 0$). Based on Equation (3.1), it can be seen that $m_g = g \cdot m_s$. Considering $m_s = 0$, it is inferred that $m_g = 0$. This means the VR always outputs zero distortion for eye position F, so the object far right point is at the original position O. Line $\overline{FO}$ is the visual line toward the far right point.

Suppose after motion parallax translation, the head moves to $(p, 0)$ under motion distortion factor g (position not shown in the figure). Suppose the virtual scene under this condition presents the same VR image as presented at position G (right eye) under no motion distortion, then the coordinates of G is $(p \cdot g + \frac{IPD}{2}, 0)$. The scene appears exactly as it would at G, distortion is only apparent during motion. The visual direction of object point O from G is shown as the yellow dashed line. However, the head actually moved only $1/g$ times this distance. If point S is the perceived self position of the right eye for this image and the observer does not perceive any distorted object motion ($f_o = 1$), the location of O must be at the intersection of $\overline{FO}$ and a line in direction parallel to $\overline{GO}$ from S. Therefore, the visual line should start with S and keep the same

visual direction to maintain the same m_g, shown as the yellow solid line $\overline{SO_1}$ which is parallel to $\overline{GO}$. According to Equation (3.3), the coordinates of point S can be obtained as $(p \cdot f_s, 0)$. The displacement between G and S is the same as the displacement between O and O_1 since the two yellow lines are parallel, so this displacement can be calculated as:

$$\overline{OO_1} = \overline{GS} = p \cdot (g - f_s)$$

Combining $\overline{OO_1}$ and coordinates of O, the coordinates of O_1 can be calculated as:

$$(x_{O_1}, y_{O_1}) = (w \cdot \sin\frac{\theta}{2} - p \cdot (g - f_s), d + w \cdot \cos\frac{\theta}{2})$$

which is the object point for right eye at S. The visual line $\overline{SO_1}$ for perceived eye position S can be inferred. Putting together the 2 visual lines $\overline{SO_1}$ and $\overline{FO}$, their intersection P_1 can be calculated, which is the predicted perceived far right point as the observers move from $(0, 0)$ to $(p, 0)$ under gain g, with perceived self motion factor f_s and perceived object motion factor as 0.

If perceived object motion is not 0 ($f_o \neq 0$), observers notice the point on the object shifts as they perceive self movement from F to S. Suppose they perceive that the object point appears at O_2 instead of O_1, based on Equation (3.4), it can be seen that

$$\overline{O_1O_2} = p \cdot (f_o)$$

Combining $\overline{O_1O_2}$ with the coordinates of O_1, the predicted coordinates of O_2 can be calculated:

$$(x_{O_2}, y_{O_2}) = (w \cdot sin\frac{\theta}{2} - p \cdot (g - f_s - f_o), d + w \cdot cos\frac{\theta}{2})$$

which means if observers perceive all motions correctly, $f_s + f_o = g^p = g$, we have $g - f_s - f_o = 0$ and perceived object point O_2 will be at ground truth O.

Similarly, P_2 being the intersection of $\overline{FO}$ and $\overline{SO_2}$ is the perceived far right point under gain g, with perceived self motion factor f_s and perceived object motion factor f_o. Using the same geometric analysis, perceptions of the peak and far left point of the fold are also derived. Connecting the perceived points, the possible perceived fold can be modelled as a triangle. For different constrained combinations of f_s and f_o, a series of respective perceived folds can be computed under motion parallax p. The task is adjusting the fold to be a right angle, so it is reasonable if the following assumption holds, i.e.,

- Peak angle of the perceived triangle is close to $\frac{\pi}{2}$;
- Bottom edge of the triangle is almost parallel to frontal axis;

Figure 3.7 shows two possible perceived right angles (black) based on an adjusted angle θ (yellow).

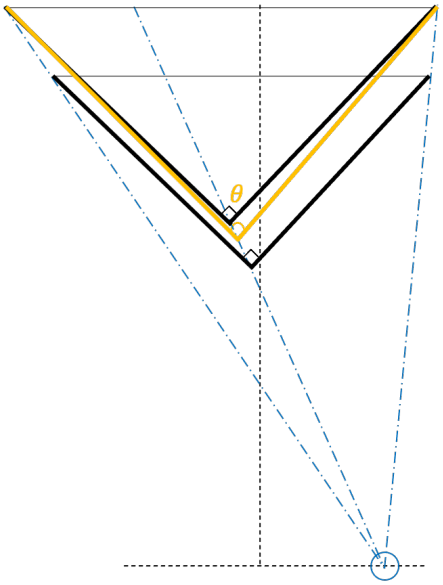


Figure 3.7: Possible perceptions (black) based on an adjusted angle θ (yellow) under gain distortion ($g > 1$). The VR presents the fold with extra object motion thus the yellow angle sways to the opposite direction (left) as observer performs motion parallax to the right.

3.6.3 Problem formulation

Keeping the vertices of the triangle on the visual lines, a series of right angles can be fit to model possible perceptions of the fold as shown in Figure 3.7. Amongst all possible perceptions, sorting out the most parsimonious one could be formulated into a multi-objective optimization problem:

$$\begin{aligned} &\underset{\vec{x}}{\text{minimize}} \quad f_1(\vec{x}), f_2(\vec{x}), \ldots, f_k(\vec{x}) \\ &\text{subject to} \quad \vec{x} \in \mathbf{X}. \end{aligned} \tag{3.8}$$

where $k \geq 2$ is the number of objectives, $\vec{x}$ is a feasible solution and set $\mathbf{X}$ is the feasible set that includes all feasible solutions. A feasible solution in this scenario

is an isosceles right angle that falls on the visual lines while its hypotenuse is parallel to frontal axis. Set $\mathbf{X}$ as a result is a group of right triangles at different distances under different scales (with different depths and medial-lateral object motion). The objectives of this optimization are for the following 3 hypotheses: observers are aware of 1) the extent of their Self Motion (Warren Jr 1995), 2) motion gain distortion in VR (Hettinger and Riccio 1992) and 3) rigidity of the test stimulus (Shiffrar and Pavel 1991; Koenderink and Van Doorn 1975).

Hypothesis 1 *(Self Motion) Observers have accurate sense of proprioception and efference through the vestibular system, active and passive limb motion, etc.*

$$f_1 = |f_s - 1| \tag{3.9}$$

which minimizes the difference between perceived self motion and the physical Self Motion, perceived self motion determines the perceived location of the eye thus affects origin of the visual lines shown as the solid lines in Figure 3.6. As a result, perceived right triangles that come from the intersection of the visual lines are also affected by perceived self motion.

Hypothesis 2 *(Object motion) Observers notice gain manipulations in VR and compensate the visual distortion as object translations.*

$$f_2 = |f_s + f_o - g| \tag{3.10}$$

By minimizing the difference between perceived total motion and gain, observers are assumed to have appropriate expectations on how the object moves.

Hypothesis 3 *(Constancy) The perceived right triangle should maintain a scale that is similar to the scale of the fold.*

$$f_3 = |depth_{fold} - depth_{perceived}|$$

Observers naturally expect the test object to be rigid while swaying laterally. Consequently, scales of the actual fold and perceived right triangle should not differ greatly. Minimizing the difference between the depths of these 2 triangles keeps the test object rigid, therefore potentially predicts the most viable perception.

Set $\mathbf{X}$ from Equation 3.8 can be considered as the following three objectives: 4) perceived points form perpendicular lines; 5) far left and right points form a bottom line that's parallel to x axis and 6) length of left and right edges are the same.

$$f_4 = |angle_{peak} - \frac{\pi}{2}|$$
$$f_5 = |slope_{bottom}|$$
$$f_6 = |length_{left} - length_{right}|$$

Since f_1, f_2, ..., f_6 have different scales, Equation (3.8) can be transformed into a normalized optimization function:

$$\underset{\vec{x}}{\text{minimize}} \quad \sum_{i=1,\cdots,6} \frac{f_i(\vec{x}) - \mu_i}{\theta_i} \tag{3.11}$$

where μ_i and θ_i are the mean and standard deviation for f_i, respectively. μ_i and θ_i were obtained by iterating all possible solutions of each f_i numerically.

Chapter 4

Results and Discussion

4.1 Data Integrity

When analysing the results, it was apparent that on some trials no adjustment inputs were made, which suggested that an accidental button press ended the trial early. As a result, 4, 9 and 1 trials were deemed as outliers and thus excluded from the results and analysis in Experiment 1, 2 and 3, respectively. The excluded data accounted for less than 1.0% of total trials in each experiment.

4.2 Motion Monitoring

As mentioned in Chapter 3, there were two Motion Statuses in each experiment, i.e., stationary and motion parallax. In motion parallax blocks, participants were instructed to sway laterally through a minimum distance of 20 cm at 0.5 Hz timed to a metronome beat. Thus, Self Motion was recorded to verify the

validity of trials and conformance to these instructions.

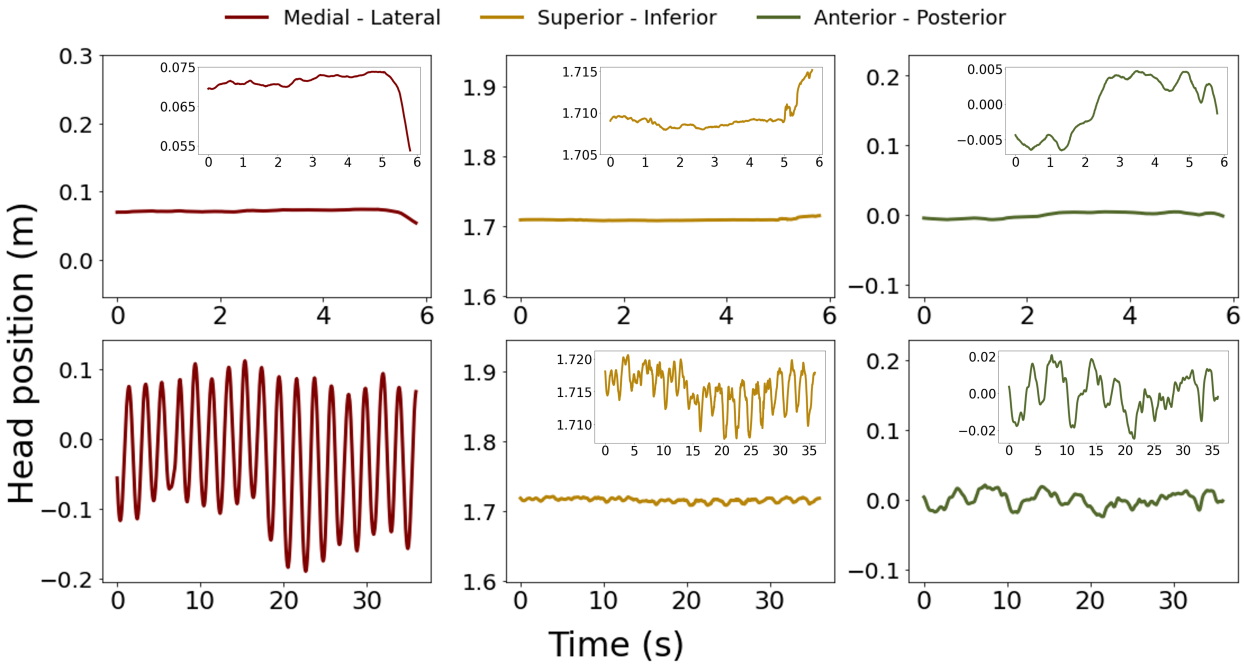


Figure 4.1: Self Motion (head position as a function of time) during 2 trials of a representative observer. Top: a stationary trial in Experiment 1. Insets show magnified views. Bottom: a motion parallax trial under $g = 2$ in Experiment 2.

Figure 4.1 shows the Self Motion of a representative observer during two trials. The horizontal axis represents the time and the vertical axis is the head position. The top plots are from a stationary trial and bottom plots are a motion parallax trial under Gain of 2. The left panel (red) is the motion parallax axis (medial-lateral); the middle yellow plots represent the height axis (superior-inferior) and the right side green plots are depth (anterior-posterior) axis. During the stationary trial, Self Motion in both parallax and depth directions were within 1 cm; while in the motion parallax condition, the observers swayed across 20 cm in the expected motion parallax direction (medial-lateral axis) and 4

cm in the depth direction (anterior-posterior axis), confirming that the depth restriction worked well. The stationary trial took about 20% (6 seconds) of the time of the motion parallax trial (36 seconds). As shown in Figure 4.1, there are roughly 18 peaks in a 36-s period, which shows conformance to the expected 0.5 Hz.

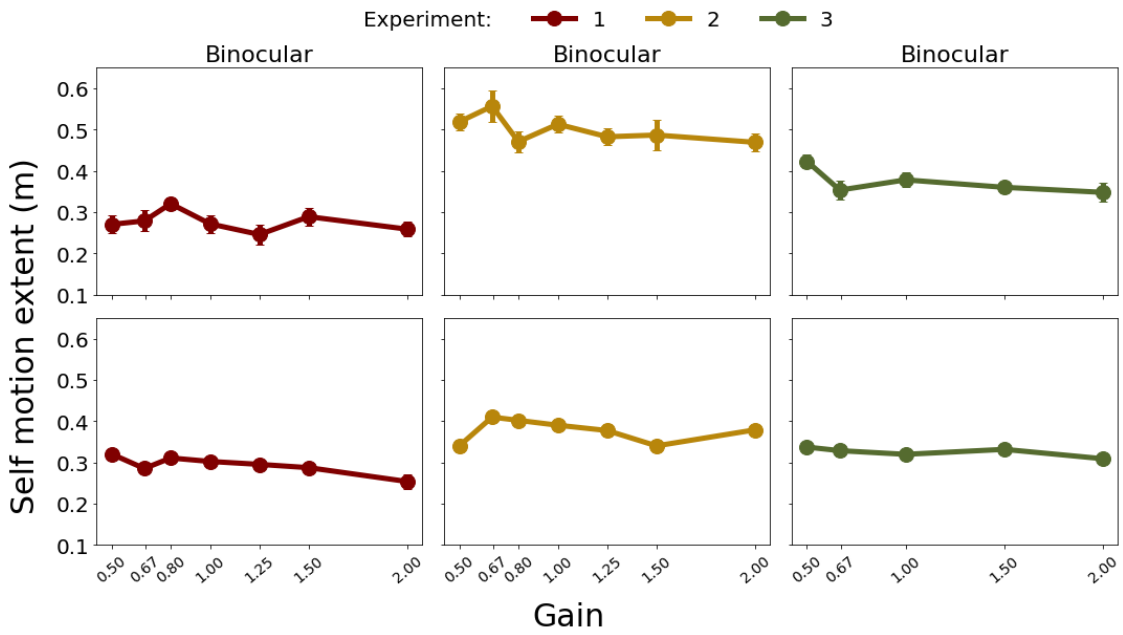


Figure 4.2: Self Motion extent of all motion parallax trials from a representative observer. Self Motion extents were first separated according to the Viewing condition then averaged by each Gain. Error bars represent ± 1 SEM.

The extent of Self Motion in each trial was calculated as the average lateral distance from peak to peak as in Figure 4.1. Figure 4.2 shows the extent of Self Motion across all Gains from the same representative observer. The extent of Self Motion in Experiment 1 is shown in red on the left panel, Experiment 2 is shown in yellow (middle panel) and Experiment 3 is shown in green (right panel). Top plots are monocular conditions and bottom plots are binocular.

The horizontal axis is Gain and vertical is Self Motion (motion parallax swaying distance) in meters.

Ideally the extent of Self Motion should not vary with Width, Gain, Fold Distance or Viewing. To assess this, Linear Mixed Models (LMM) and Likelihood Ratio Tests (LRT) were used to analyze the impact of the independent variables and compare LMM models with an α of 0.05. The analyses were performed in R (version 4.0.4, 2021-02-15) using the lme4 (Bates et al. 2015) and lmtest (Zeileis and Hothorn 2002) package. The extent of Self Motion (averaged in each trial) was the dependent variable; Viewing, Gain, Fold Distance and Width of the dihedral planes were fixed effects and observer was a random effect. I started with a base model which included main effects of the independent variables including Gain, Fold Distance and Width and their interactions with Viewing to identify whether theses main effects were modulated by binocular Viewing. Then I used LRT to probe for significant two- and three-way interactions. No interaction between any pair of Gain, Fold Distance and Width produced significant improvements in the model. Based on the LRT result, the model specification was chosen as follows:

$$\begin{aligned} Self\ Motion \quad \sim \quad & Viewing \times Gain \quad + \quad Viewing \times Fold\ Distance \\ & + \quad Viewing \times Width \quad + \quad (1|Observer) \end{aligned} \tag{4.1}$$

LMM analysis suggest that Gain had a main effect on Self Motion in all three experiments ($p < 0.01$). This suggests that observers relied on virtual motion shown in the scene to adjust the extent of Self Motion. But the effect of Gain was relatively small (-3.69 ± 1, -1.82 ± 0.4 and -1.34 ± 0.4 cm/Gain unit,

95% Confidence Interval (CI), in Experiment 1, 2 and 3, respectively) so I still consider Self Motion invariant of Gain distortion. The movement feedback also ensured that the participants made the minimum required motion in all conditions.

Observers swayed less under binocular Viewing in Experiment 2. This could be because that in Experiment 2 and 3, observers were presumably alerted to the fact the fold was rendered at different Fold Distances (distance matching task was introduced) and they relied on stereo cues to determine their extent of Self Motion while accounting for distance. However, in Experiment 3, the test range was greatly extended (up to 6 m). Parallactic change is very limited at a large distance, e.g., 6 m, which may explain why the (binocular) Viewing effect was not observed. However, like the Gain effect, the (binocular) Viewing effect was also small (2.81 cm difference between binocular and monocular Viewing conditions in the model) and I consider it had little effect on Self Motion.

Fold Distance and Width did not have any effect on Self Motion. No interaction effects between the independent variables were found on Self Motion either.

4.3 Experiment 1

In the experiments, participants were instructed to adjust the fold Angle to 90°. All participants reported that the task that they were asked to perform was difficult. As a result, none of them were confident with their judgments. However, their data showed that the judgments they made were fairly consistent,

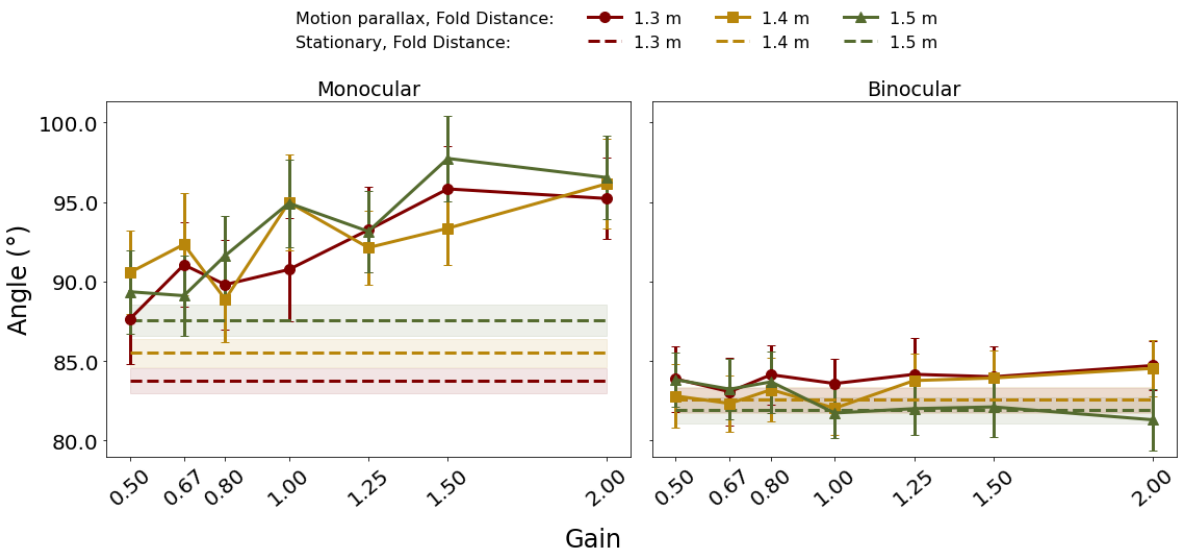


Figure 4.3: Angle settings as a function of Gain averaged across observers in Experiment 1. Stationary settings are plotted as a line across the x-axis for easy visualization. Error bars and regions around dashed lines represent ± 1 SEM.

suggesting that observers have reliable perception across visual manipulations.

4.3.1 Angle and Depth

Figure 4.3 shows the Angle settings plotted as a function of Gain levels. The Angle data was first separated according to Fold Distance and Viewing, then averaged across observers at each Gain. Observers set the Angles closer to veridical (90°) when viewed monocularly. Binocular Viewing generally yielded Angles less than 90°. Intra- and inter-subject differences were also small. Individual Angle plots are shown in Figure B.1.

With each Angle input, I calculated the Depth of the fold. Since width was manipulated across trials, Depth was normalized by the scaling factor when

comparing across conditions:

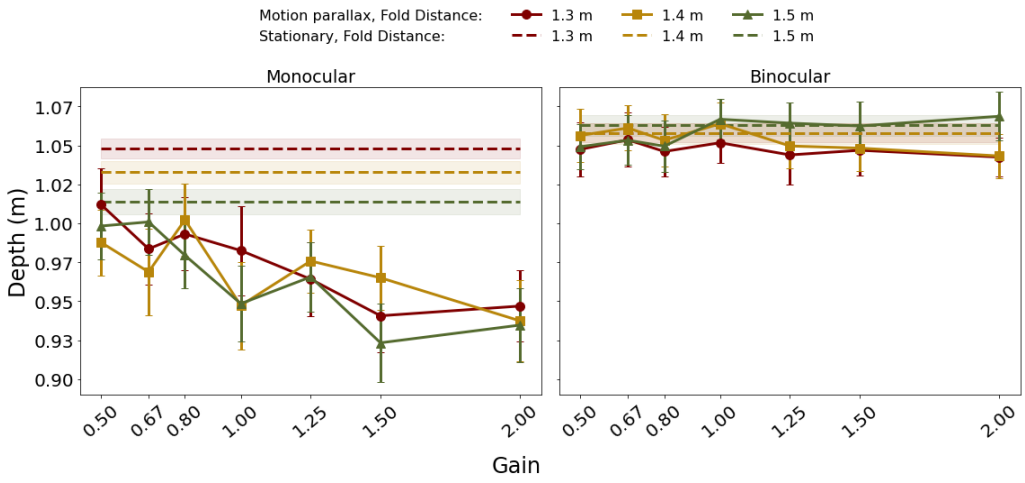
$$d = \frac{w \cdot cos\frac{\theta}{2}}{s} \tag{4.2}$$

where d is the Depth in each trial; w is the width of the planes; s is the scaling factor of the width. If all Angles were adjusted to perfect 90° then $\theta = 90°$ and $d = 1$ m. If not, $d > 1$ m would mean the current condition yields more Depth and $d < 1$ m would mean less Depth than a objectively 90° fold.

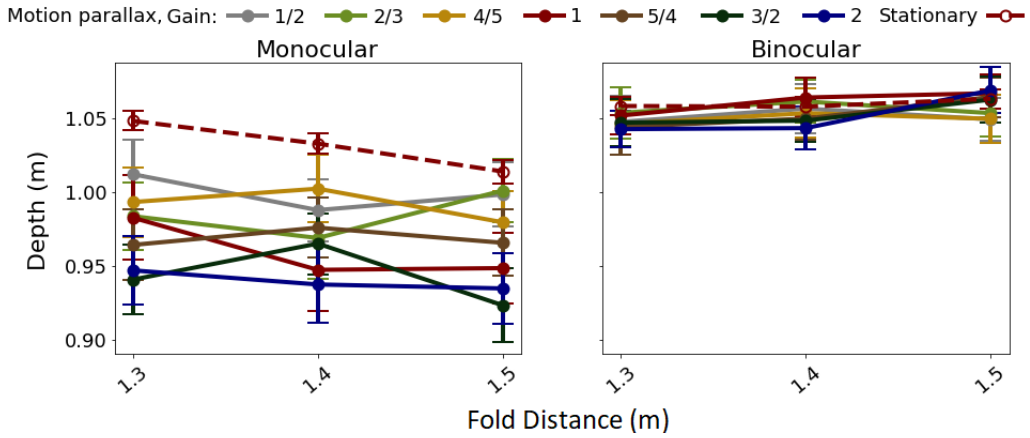
Similar to Model 4.1, the LMM model on Depth was chosen as:

$$\begin{aligned} Depth \quad \sim \quad & Viewing \times Gain \quad + \quad Viewing \times Fold\ Distance \\ & + \quad Viewing \times Width \quad + \quad (1|Observer) \end{aligned} \tag{4.3}$$

Under monocular Viewing, observers set the fold close to veridical (Depth, d close to 1) under compressive Gain ($g < 1$). With increasing Gain, Depth tended to decrease, as shown in Figure 4.4(a). That is, observers' perceived 90° settings corresponded to objectively flatter folds under expansive Gains ($g > 1$). This was confirmed by LMM findings of a significant interaction between Viewing and Gain on Depth ($p < 0.001$). In post-hoc analysis, General Linear Hypotheses contrast tests (GLHT) were performed using the multcomp package (Hothorn, Bretz, and Westfall 2008). The Gain×Viewing interaction was analyzed through simple main effects of Gain at each level of Viewing. The results show a significant negative slope of Gain under monocular Viewing (-4.16 ± 1.3 cm/Gain unit, 95% CI, $p < 0.001$ as shown in Table 4.1. Significant effect sizes are highlighted in bold fonts.) and close to and not significantly different from zero effect of Gain under binocular Viewing. Binocular cues under binocular Viewing apparently removed the Gain effect.



(a) Depth settings as a function of Gain averaged across observers in Experiment 1. Stationary settings are plotted as a line across the x-axis for easy visualization.



(b) Depth settings as a function of Fold Distances averaged across all observers in Experiment 1.

Figure 4.4: Depth settings averaged across all observers in Experiment 1. Error bars and regions around dashed lines represent ± 1 SEM.

Table 4.1: GLHT results of Gain effect on Depth under different Viewing conditions in Experiment 1.

Viewing	Gain effect (m)	2.5% bound (m)	97.5% bound (m)
Monocular	**-0.046**	**-0.063**	**-0.029**
Binocular	4.3×10^{-3}	-0.022	0.013

Table 4.2: GLHT results of Viewing effect on Depth at different Fold Distances under Gain of 1 in Experiment 1.

Motion Status	Viewing	Fold Distance (m)	Effect (m)	2.5% bound (m)	97.5% bound (m)
Motion parallax	**Binocular vs. monocular**	**1.3**	**0.070**	**0.050**	**0.089**
		1.4	**0.081**	**0.061**	**0.10**
		1.5	**0.089**	**0.070**	**0.11**
		all distances	**0.080**	**0.072**	**0.087**
Stationary	**Binocular vs. monocular**	1.3	0.002	-0.009	0.030
		1.4	**0.003**	**0.005**	**0.044**
		1.5	**0.050**	**0.030**	**0.070**
		all distances	**0.028**	**0.020**	**0.036**

Separate analyses were performed for the stationary baseline and motion parallax conditions and the effect sizes are shown in Table 4.2. LMM results suggested Viewing had a main effect on Depth ($p < 0.001$ under both motion parallax and stationary conditions), binocular Viewing yielded bigger Depth settings, shown in the right panel of Figure 4.4(a).

Figure 4.4(b) shows the averaged Depth data plotted as a function of Fold Distance. Individual plots are shown in Figure B.2(b). LMM results found increasing Fold Distance produced a decreasing trend in Depth when observers stayed stationary under monocular Viewing, see dashed lines in the left panel of Figure 4.4(b). Contrast results are shown in Table 4.3. Stationary and

Table 4.3: GLHT results of Fold Distance effect on Depth under different Viewing conditions and Gain of 1 in Experiment 1. Fold Distance is categorical but its linear and quadratic effect were also tested.

Motion condition	Viewing mode	Linear effect	Quadratic effect	Pairwise contrast (m)	Pairwise effect (°)	2.5% bound	97.5% bound
Motion parallax	Monocular	-0.009	0.001	1.4 vs. 1.3	-0.005	-0.021	0.011
				1.5 vs. 1.3	-0.009	-0.025	0.007
				1.5 vs. 1.4	-0.004	-0.020	0.012
	Binocular	0.011	0.002	1.4 vs. 1.3	0.006	-0.009	0.022
				1.5 vs. 1.3	0.011	-0.005	0.027
				1.5 vs. 1.4	0.004	-0.011	0.020
Stationary	**Monocular**	**0.024**	0.003	**1.4 vs. 1.3**	**-0.014**	**-0.030**	**-0.003**
				1.5 vs. 1.3	**-0.034**	**-0.051**	**-0.018**
				1.5 vs. 1.4	**-0.020**	**-0.037**	**-0.004**
	Binocular	0.004	0.002	1.4 vs. 1.3	<0.001	-0.016	0.017
				1.5 vs. 1.3	0.005	-0.011	0.022
				1.5 vs. 1.4	0.005	-0.011	0.022

monocular observers may not have appreciated the difference in Fold Distance, since the texture of the stimulus was scaled with distance. They adjusted the far folds flatter while perceiving them as 90°, which could be due to an underestimation on the egocentric distance of the far apexes. This effect was not present under either binocular Viewing or motion parallax thus the 'illusion' that all folds were rendered at the same close Fold Distance was not present.

4.3.2 Prediction

In the previous sections, I showed the ground truth of Gain g, Fold Distance and Depth as either independent or dependent variables. In this analysis, the predicted perceptions of these variables is shown, including perceived gain g^p, perceived self motion f_s^p, perceived depth d^p and perceived distance of the best matching 90° solution for the same conditions from our predictive model[1]. The predictive model is based on motion parallax so there is no prediction for stationary conditions.

Figure 4.5 gives the prediction of perceived self motion averaged across all observers. Figure B.3 shows the prediction results for individual observers. The model predicted that observers perceived self motion almost accurately $f_s^p \approx 1$ and the plots of the predictions are flat lines.

Figure 4.6 and B.4 show the prediction result on gain perception (g^p) or equivalently, total motion perception ($f_s^p + f_o^p$). If Gain is perceived accurately,

[1]Note that the previous section presented the actual set Depth rendered for the apparently 90° hinges. This was calculated from the fold adjusted by observers in Section 4.3.1, which is not necessarily 90°. In distinction, for the perceived depth prediction in this section, I am showing the depth of the perceptual 90° fold predicted by Model 3.11 based on the participant's setting.

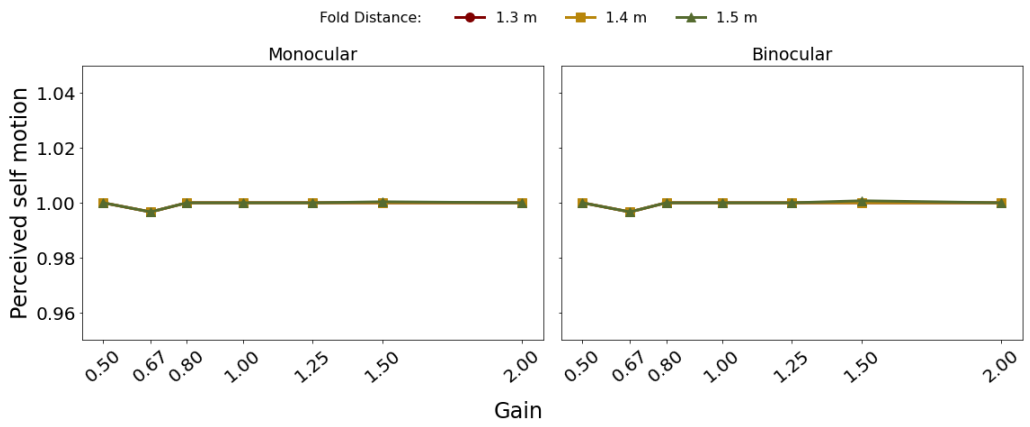


Figure 4.5: Predictions of perceived self motion as a function of Gain averaged across all observers in Experiment 1. Error bars (too small to be seen in this plot) represent ± 1 SEM.

$g^p = g$, a diagonal line will present.

The general trend of predictions was consistent across observers. The model predicted that Gain was perceived nearly perfectly under compressive/no Gain distortion ($g \leq 1$) but underestimated under expansive Gains ($g > 1$) starting from Gain of 1.25. Or in other words, the optimization model predicted that observers had different compensatory mechanisms under expansive ($g > 1$) versus compressive ($g < 1$) Gain distortions. Subsequently, for each Angle that the observers adjusted, I used the perceived self motion f_s^p and object motion f_o^p to predict the perceived depth and position (distance) of the adjusted fold. The previous results of this section showed the model prediction of close to veridical perceived self motion f_s^p and perceived gain g^p, especially when $g \leq 1$. This suggests most of the compressive Gains were seen as object motion not Depth difference.

Figure 4.7 shows the prediction of perceived depth as a function of Gain,

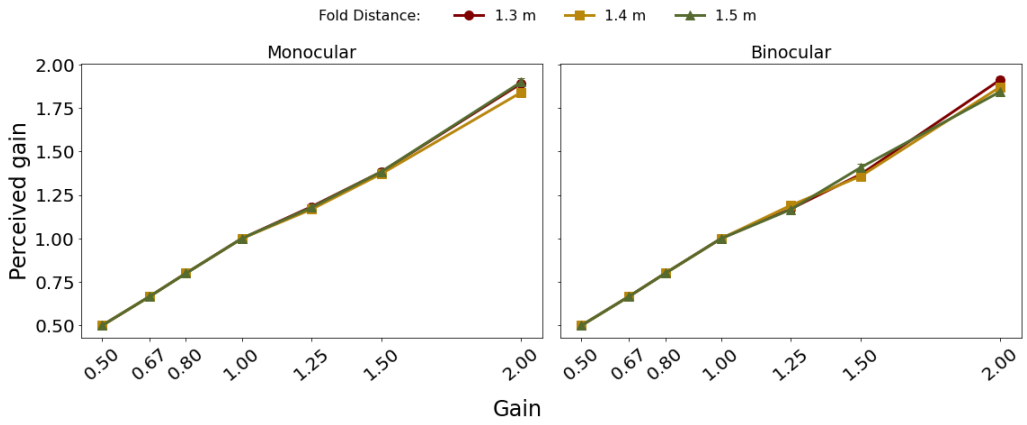


Figure 4.6: Predictions of perceived gain as a function of Gain averaged across all observers in Experiment 1. Error bars (too small to be seen in this plot) represent ± 1 SEM.

averaged across all observers (see Figure B.5 for individual plots).

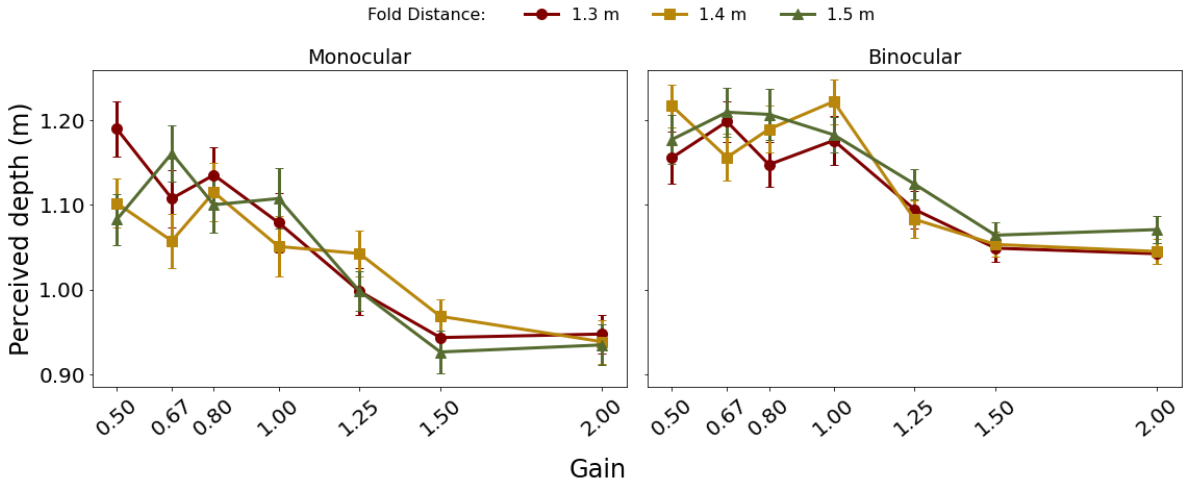


Figure 4.7: Predictions of perceived depth of fold as a function of Gain averaged across all observers in Experiment 1. Stationary settings are plotted as a line across the x-axis for easy visualization. Error bars represent ± 1 SEM.

The trend of perceived depth is similar to the trend of actual Depth decreasing with increasing Gain under monocular Viewing. For binocular Viewing

conditions, our model still predicted a similar trend as in monocular conditions. This could be because the model only considered monocular cues (motion parallax) but not binocular cues under binocular Viewing, such as binocular disparity or vergence angle. The model predicted observers would perceive more depth than for an objectively a 90° fold ($d^p > 1$) under compressive Gains ($g < 1$). The overestimation of perceived depth gradually decreased to a slight underestimation with increasing expansive Gains ($g > 1$). Taken together with the prediction of perceived gain (g^p) in Figure 4.6, the underestimation of expansive Gain ($g^p < g | g > 1$) resulted in close to perfect depth perception ($d^p \approx 1$), while perfectly perceived compressive Gain ($g^p \approx g | g < 1$) produced overestimated depth perception ($d^p > 1$). In other words, although observers successfully perceived the direction of compressive or expansive Gain, they may not have perfect perception of the absolute extent of Gain thus the perceived depth is stretched.

Related to the decreasing trend on perceived depth with increasing Gain, the model predicted that observers should perceive variations in the distance of the apex. Figure 4.8 shows the prediction from the proposed model, averaged across all observers. Individual predictions are shown in Figure B.6. Clearly, the model predicted the perceived fold distance to be underestimated under expansive Gains, while under compressive Gains it is predicted to be perceived accurately. The expansive Gain was not fully attributed to perceived self and object motion in the model ($f_s^p + f_o^p < g$ or equivalently $g^p < g$ given $g > 1$, as reflected in the underestimation in Figure 4.6 when $g > 1$), and distortion of perceived fold distance was predicted.

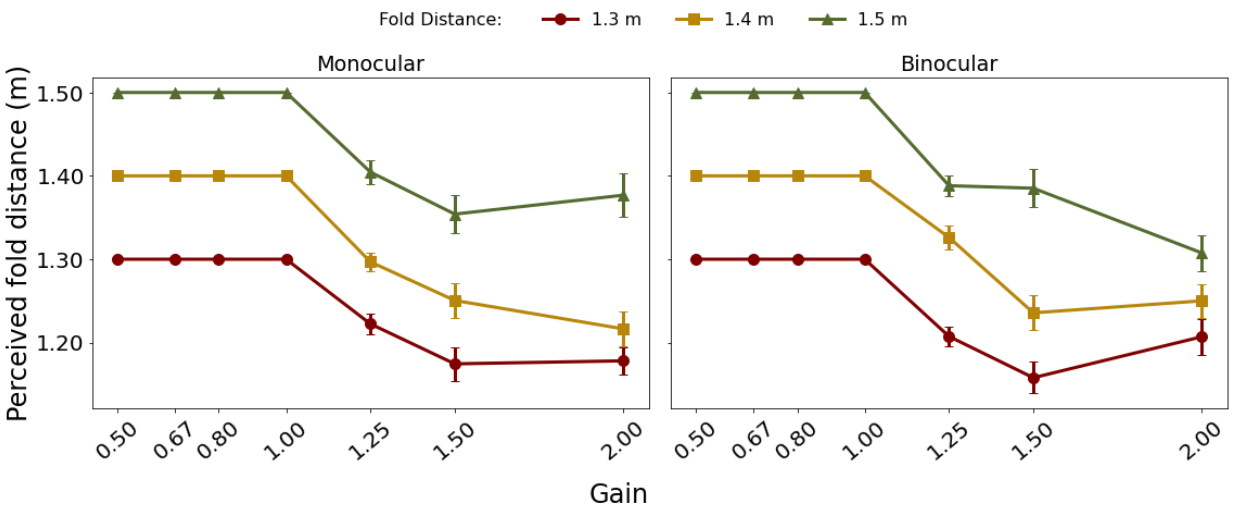


Figure 4.8: Predictions of perceived fold distance as a function of Gain averaged across observers in Experiment 1. Error bars (some are too small to be seen in this plot) represent ± 1 SEM.

To conclude Experiment 1, observers adjusted the Depth of fold to be smaller with increasing Gain under monocular Viewing. The proposed model suggests observers perceptually favor closer folds with smaller depths under expansive Gains ($g > 1$). If this prediction is the case, observers should perceive not only underestimated Fold Distances, but also smaller scales of fold according to the geometry shown in Figure 3.7 (black triangles) in Section 3.6.2.

Fold Distance had a decreasing effect on monocular Depth only under stationary Motion Status. Introducing more distance cues of either motion parallax or binocular Viewing weakened this effect.

Since the model suggested that most observers should perceive decreasing distance with increasing Gain, a second set of experiments were conducted to test this prediction.

4.4 Experiment 2

In order to verify the predicted decrease in the apparent distance of the fold with increasing Gain, Experiment 2 added an additional evaluation of Pole Distance. In each trial, after the Angle adjustment task, observers matched a pole to the remembered apex of the previously seen fold. Except for that, similar analyses as in experiment 1 were performed.

4.4.1 Angle and Depth

Figure 4.9 and 4.10 show the Angle and Depth settings. The LMM results showed that the interaction between monocular Viewing and Gain on Depth remained significant ($p < 0.001$); its effect sizes are shown in Table 4.4 as simple main effects of Gain for each Viewing condition. The same decreasing Depth/increasing Angle settings with increasing Gain found in Experiment 1 were observed for most monocular conditions in this experiment. Although no Gain by Fold Distance interaction was found, visually the effects seem especially pronounced at 1.3 m and 1.4 m.

Again Viewing had a significant effect on Depth in both motion parallax and stationary blocks (both $p's < 0.001$, effect sizes of Viewing under different Fold Distances are reported in Table 4.5) showing observers adjusted bigger Depth under binocular Viewing under Gain of 1.

Figure 4.10(b) and B.8(b) show the average and individual Depth data as a function of Fold Distance, the plots are visually similar to Experiment 1. However, in Experiment 2 the decreasing trend on Depth with increasing Fold

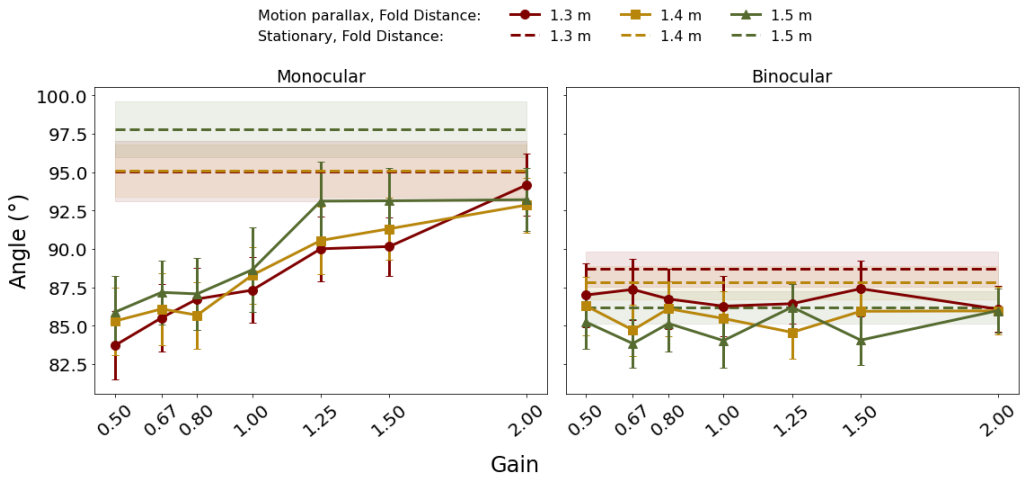


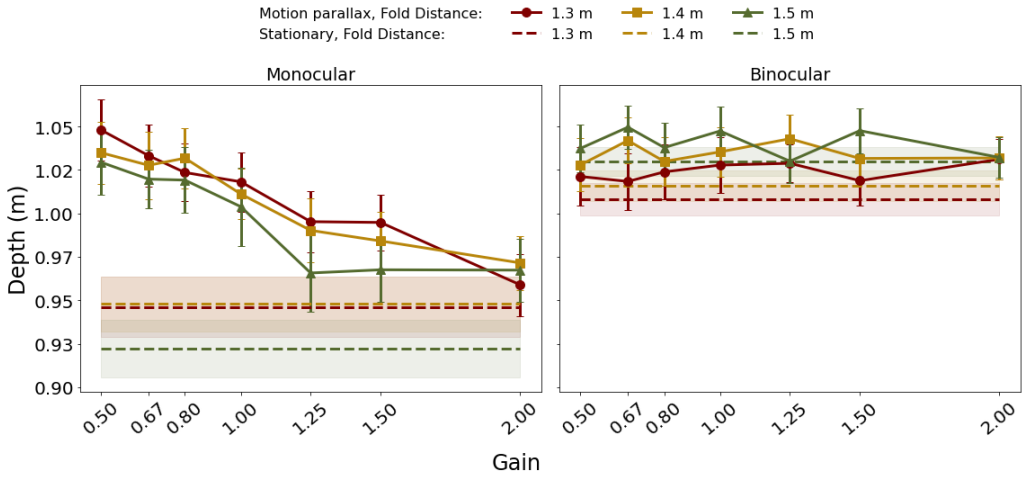
Figure 4.9: Angle settings as a function of Gain averaged across all observers in Experiment 2. Error bars and regions around dashed lines represent ± 1 SEM.

Table 4.4: GLHT results of Gain effect on Depth under different Viewing conditions in Experiment 2.

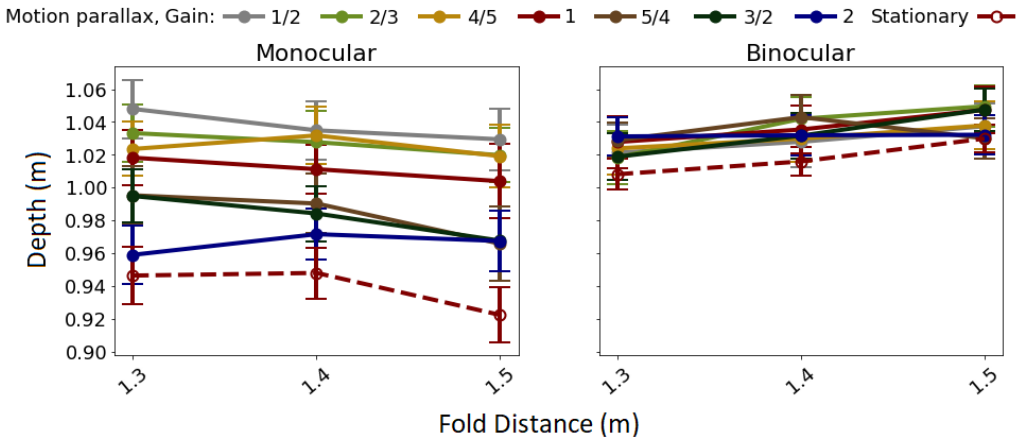
Viewing	Gain effect (m)	2.5% bound (m)	97.5% bound (m)
Monocular	**-0.061**	**-0.075**	**-0.048**
Binocular	-0.011	-0.024	0.003

Table 4.5: GLHT results of Viewing effect on Depth at different Fold Distances under Gain of 1 in Experiment 2.

Motion condition	Viewing	Fold Distance (m)	Effect (m)	2.5% bound (m)	97.5% bound (m)
Motion parallax	**Binocular vs. monocular**	1.3	0.010	-0.0005	0.020
		1.4	**0.022**	**0.012**	**0.032**
		1.5	**0.039**	**0.028**	**0.049**
		all distances	**0.023**	**0.017**	**0.029**
Stationary	**Binocular vs. monocular**	**1.3**	**0.062**	**0.023**	**0.10**
		1.4	**0.072**	**0.033**	**0.11**
		1.5	**0.10**	**0.065**	**0.14**
		all distances	**0.079**	**0.064**	**0.094**



(a) Depth settings as a function of Gain averaged across all observers in Experiment 2. Stationary settings are plotted as a line across the x-axis for easy visualization.



(b) Depth settings as a function of Fold Distances averaged across all observers in Experiment 2.

Figure 4.10: Depth settings averaged across all observers in Experiment 2. Error bars and regions around dashed lines represent ± 1 SEM.

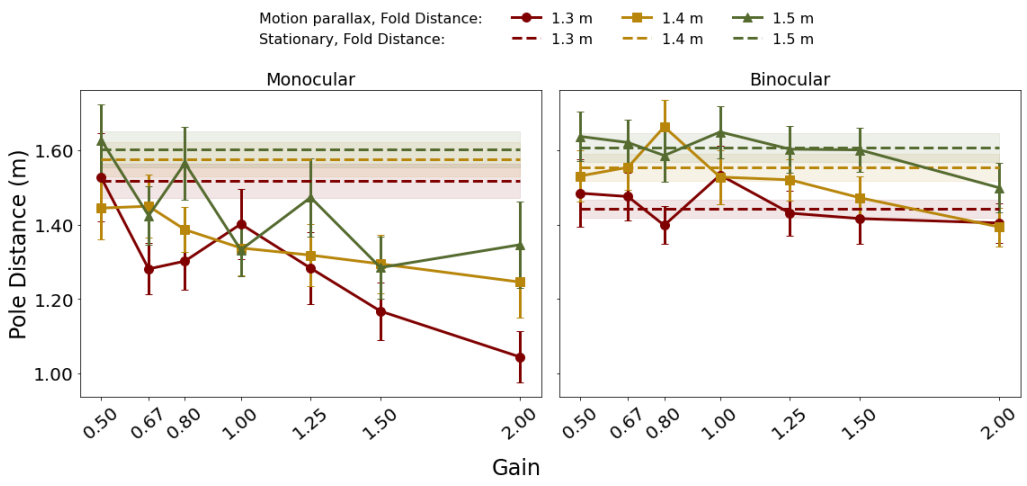
Table 4.6: GLHT results of Fold Distance effect on Depth under different Viewing conditions and Gain of 1 in Experiment 2. Fold Distance is categorical but its linear and quadratic effect were also tested.

Motion condition	Viewing mode	Linear effect	Quadratic effect	Pairwise contrast (m)	Pairwise effect (°)	2.5% bound	97.5% bound
Motion parallax	**Monocular**	**-0.013**	-9.27×10^{-3}	1.4 vs. 1.3	-0.002	-0.014	0.010
				1.5 vs. 1.3	**-0.013**	**-0.025**	**-0.001**
				1.5 vs. 1.4	-0.011	-0.023	0.001
	Binocular	**0.016**	4.50×10^{-3}	1.4 vs. 1.3	0.010	-0.002	0.022
				1.5 vs. 1.3	**0.016**	**0.004**	**0.028**
				1.5 vs. 1.4	0.006	-0.006	0.018
Stationary	Monocular	0.014	5.22×10^{-3}	1.4 vs. 1.3	-0.004	-0.035	0.028
				1.5 vs. 1.3	-0.020	-0.051	0.011
				1.5 vs. 1.4	-0.016	-0.048	0.015
	Binocular	0.016	3.32×10^{-3}	1.4 vs. 1.3	0.007	-0.024	0.038
				1.5 vs. 1.3	0.022	-0.009	0.054
				1.5 vs. 1.4	0.015	-0.016	0.047

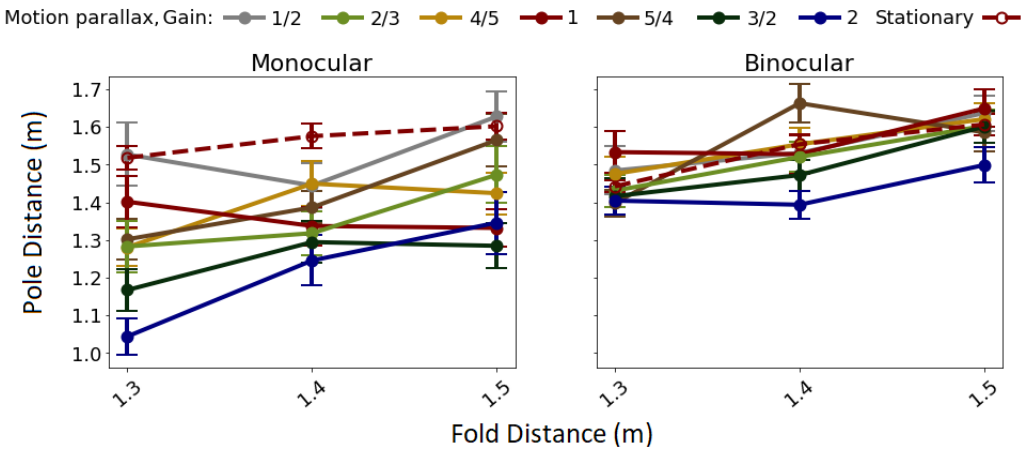
Distance was not present in the monocular and stationary block. Conversely, this effect was observed in the monocular and motion parallax block although the effect size was smaller. Moreover, the Fold Distance produced a small increasing effect on Depth in the binocular and motion parallax block (interaction between Viewing and Fold Distance, $p < 0.001$). This could be caused by the introduction of the second task since observers were presumably alerted about the Fold Distance manipulations especially when binocular and motion parallax cues were both available. On one hand observers could have increasingly overestimated the distance of far apexes to address the Fold Distance variations, on the other, because of the general underestimation of the distance in VR (Cutting and Vishton 1995; Wright 1995), observers could have compressed the virtual distance of the left/right edges of far folds thus perceived sharper folds (with bigger depths) as 90°.

4.4.2 Pole Distance

Figure 4.11 shows the Pole Distance plots averaged across all participants, individual plots are shown in Figure B.9. LMM results in Table A.2(a) show that Viewing did not have a main effect on Pole Distance in motion parallax blocks but there was a significant interaction between Viewing and Gain. Effect sizes of Viewing and Gain are shown in Table 4.7 and 4.8. Pole Distance decreased more with increasing Gain under monocular than binocular Viewing. Monocularly Viewing observers overestimated the Pole Distance under compressive Gains ($g < 1$) but underestimated under expansive Gains ($g > 1$). They perceived the distance to be closer as Gain increased (solid lines in Figure 4.11) which



(a) Pole Distance settings as a function of Gain averaged across all observers in Experiment 2. Stationary settings are plotted as a line across the x-axis for easy visualization.



(b) Pole Distance settings as a function of Fold Distances averaged across all observers in Experiment 2.

Figure 4.11: Pole Distance settings averaged across all observers in Experiment 2. Error bars and regions around dashed lines represent ± 1 SEM.

Table 4.7: GLHT results of Viewing effect on Pole Distance at different Fold Distances under Gain of 1 in Experiment 2. Distance is categorical but its linear and quadratic effect were also tested.

Motion Status	Viewing	Distance (m)	Effect (m)	2.5% bound (m)	97.5% bound (m)
Motion parallax	**Binocular vs. monocular**	**1.3**	**0.153**	**0.092**	**0.214**
		1.4	**0.158**	**0.097**	**0.219**
		1.5	**0.151**	**0.090**	**0.212**
		all distances	**0.154**	**0.118**	**0.190**
Stationary	Binocular vs. monocular	1.3	-0.063	-0.200	0.075
		1.4	0.005	-0.140	0.150
		1.5	0.005	-0.134	0.145
		all distances	-0.017	-0.126	0.091

Table 4.8: GLHT results of Gain effect on Pole Distance under different Viewing conditions in Experiment 2.

Viewing	Gain effect (m)	2.5% bound (m)	97.5% bound (m)
Monocular	**-0.189**	**-0.247**	**-0.130**
Binocular	**-0.082**	**-0.140**	**-0.023**

generally aligns with the prediction from Experiment 1.

Figure 4.11(b) shows the averaged Pole Distance plotted as a function of Fold Distance, individual results are shown in Figure B.9. The general trend of Pole Distance under monocular Viewing aligns with the prediction from Experiment 1 but the binocular condition trend does not. Observers compressed the absolute distance with increasing expansive Gain ($g > 1$) under monocular Viewing. Further, under compressive Gains ($g > 1$) or binocular Viewing, Pole Distance was overestimated although observers might have set the virtual pole further than they actually perceived in these conditions because of the general

Table 4.9: GLHT results of Fold Distance effect on Pole Distance under different Viewing conditions and Gain of 1 in Experiment 2. Fold Distance is categorical but its linear and quadratic effect were also tested.

Motion condition	Viewing mode	Linear effect	Quadratic effect	Pairwise contrast (m)	Pairwise effect (m)	2.5% bound (m)	97.5% bound (m)
Motion parallax	**Monocular**	**0.151**	0.014	1.4 vs. 1.3	0.068	-0.004	0.141
				1.5 vs. 1.3	**0.151**	**0.078**	**0.223**
				1.5 vs. 1.4	**0.083**	**0.010**	**0.155**
	Binocular	**0.149**	0.002	**1.4 vs. 1.3**	**0.074**	**0.001**	**0.146**
				1.5 vs. 1.3	**0.15**	**0.076**	**0.221**
				1.5 vs. 1.4	**0.075**	**0.003**	**0.148**
Stationary	Monocular	0.090	-0.017	1.4 vs. 1.3	0.054	-0.054	0.161
				1.5 vs. 1.3	0.090	-0.017	0.196
				1.5 vs. 1.4	0.036	-0.072	0.144
	Binocular	**0.158**	−0.085	**1.4 vs. 1.3**	**0.122**	**0.015**	**0.229**
				1.5 vs. 1.3	**0.158**	**0.051**	**0.265**
				1.5 vs. 1.4	0.036	-0.071	0.144

underestimation in VR systems. Pole Distance increased with Fold Distance in both motion parallax (solid lines, $p < 0.001$) and stationary blocks (dashed lines, $p < 0.001$). Complete LMM results are shown in Table A.2(a) and A.2(b). This implies that observers had correct perception on relative distance. Depths under binocular Viewing are bigger than monocular at the farthest Fold Distance 1.5 m under both Motion Statuses although no interaction effect were found between Viewing and Fold Distance. Contrast test results of Fold Distance effect are shown in Table 4.9.

4.4.3 Prediction

Similar trends of predicted perceptions of Self Motion, Gain, Depth and Fold Distance from Experiment 1 were also predicted in this experiment (average and individual results are shown in Appendix B).Gain dependence for predicted perception of depth in this experiment was less aggressive starting from Gain of 1–1.25 than in Experiment 1. The model again predicts that observers perceived stretch in depth under compressive Gains but compression on depth and distance under expansive Gains.

To conclude Experiment 2, both Depth and Pole Distance decreased with increasing Gain under monocular Viewing. Under binocular Viewing, the effect of Gain on Depth was not observed but a Gain effect on Pole Distance was present. As the Pole Distance matching task presumably alerted the observers about the difference in Fold Distances, an increase in Depth settings with increasing Fold Distance was found in conditions with rich distance/depth cue conditions (the motion parallax and binocular Viewing block).

In the natural world, objects that are closer to laterally swaying observers will present more observed motion. This aligns with our finding in this experiment that amplified motion parallax from expansive Gain distortion ($g > 1$) made the observers perceive a nearer Fold Distance; compressive Gain distortions ($g < 1$) that reduced the extent of motion parallax made them perceive a relatively further Fold Distance. At the same time, Depth of the stimulus seen as a right Angle was less under large Gains. Taken together, observers could see the stimulus as bigger and farther under compressive Gain effect but smaller and nearer (especially the apex) under expansive Gain effect.

The range of Fold Distances used in this experiment was limited as the farthest condition was 1.5 m, so I asked: whether adjusted depth and distance will present the same trends if the Fold Distance is much farther? This motivated Experiment 3, in which distance was extended to 6.0 m and I used a subset of Gain levels to repeat the same tasks as in Experiment 2.

4.5 Experiment 3

Observers presented increasing compression of depth and absolute distance as Fold Distance increased from 1.3 to 1.5 m when viewed monocularly in Experiment 2. The testing range of Fold Distance in Experiment 1 and 2 was only 0.2 m and the furthest testing stimulus was only 1.5 m which was close to the observers. As a follow-up on Experiment 2, a farther and larger range of Fold Distances, 1.5, 3 and 6 m, was tested in Experiment 3 to confirm if Depth continued to decrease with Fold Distance over a larger range. For Gain, I used a subset (5 out of 7 levels) from Experiment 1 and 2: $\frac{1}{2}$, $\frac{2}{3}$, 1, $\frac{3}{2}$ and 2.

4.5.1 Angle and Depth

Figure 4.12 and 4.13 show the Angle and Depth results as a function of Gain. In Experiments 1 and 2, increasing Gain produced decreasing Depth settings only under monocular Viewing. In Experiment 3 however, the interaction between Viewing condition and Gain was significant but marginal ($p = 0.049$), which appears to be due to a smaller but still significant monocular effect of Gain under the extended distance range, see Table 4.10 for effect sizes. Depth responses at

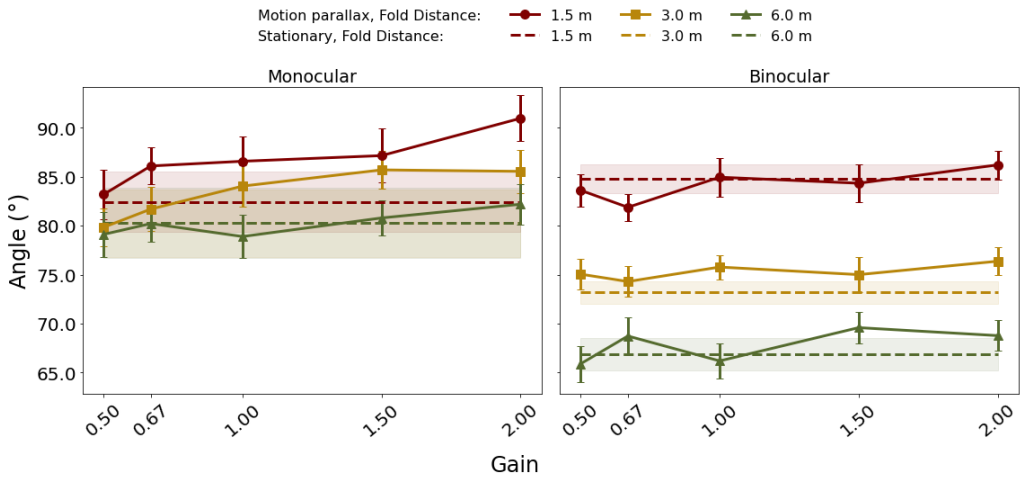


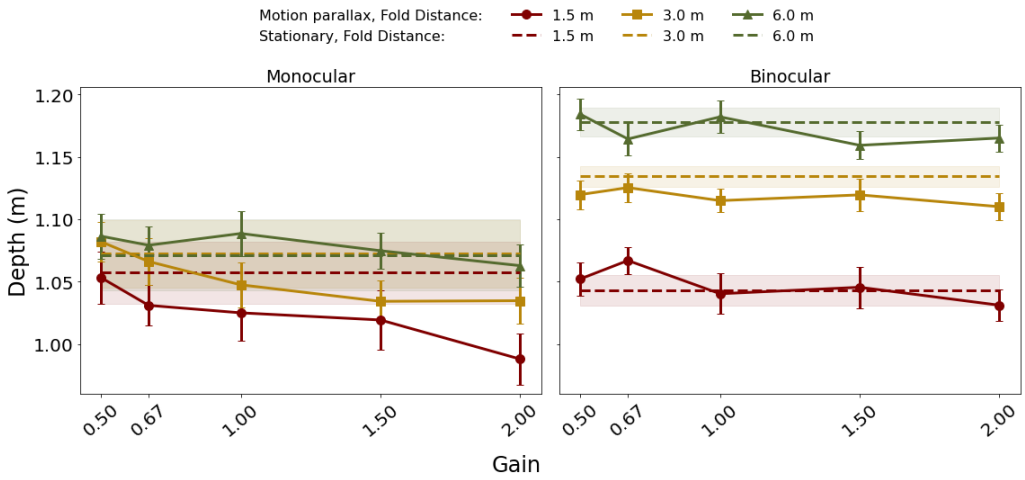
Figure 4.12: Angle settings as a function of Gain averaged across all observers in Experiment 3. Error bars and regions around dashed lines represent ± 1 SEM.

Table 4.10: GLHT results of Gain effect on Depth under different Viewing conditions in Experiment 3.

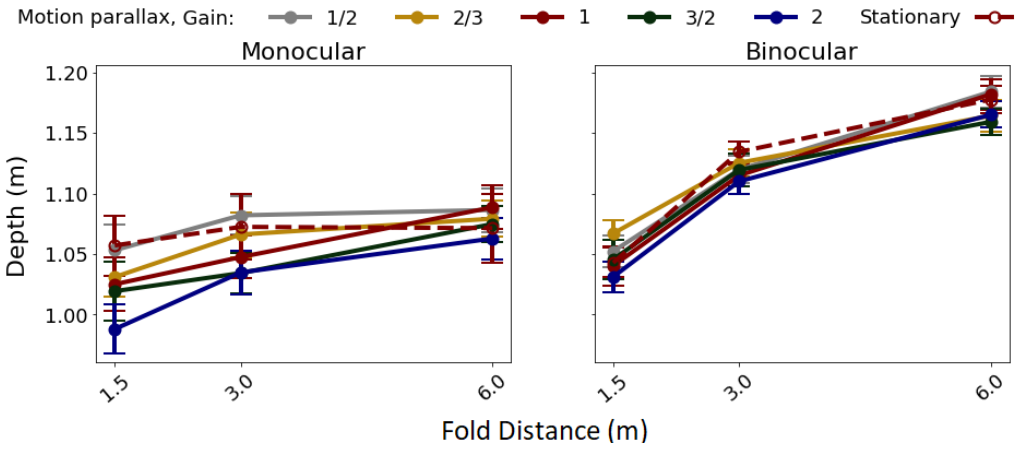
Viewing	Gain effect (m)	2.5% bound (m)	97.5% bound (m)
Monocular	**-0.039**	**-0.057**	**-0.021**
Binocular	-0.023	-0.042	0.0053

1.5 and 3.0 m were consistent with the previous findings in Experiments 1 and 2. The decreasing Depth trend with increasing Gain under monocular Viewing at relatively close distances indicates that the fold was set to be flatter and thus observers perceived the fold as more slanted than their input settings under expansive Gains. No obvious effect of Gain on Depth was present at 6.0 m or under binocular Viewing. Extended distance range tested clearly weakened the Gain effect.

Observers made sharper folds with larger depth settings under binocular Viewing compared to monocular Viewing. A main effect of Viewing was found



(a) Depth settings as a function of Gain averaged across all observers in Experiment 3. Stationary settings are plotted as a line across the x-axis for easy visualization.



(b) Depth settings as a function of Fold Distances averaged across all observers in Experiment 3.

Figure 4.13: Depth settings averaged across all observers in Experiment 3. Error bars and regions around dashed lines represent ± 1 SEM.

Table 4.11: GLHT results of Viewing effect on Depth at different Fold Distances under Gain of 1 in Experiment 3.

Motion Status	Viewing	Fold Distance (m)	Effect (m)	2.5% bound (m)	97.5% bound (m)
Motion parallax	**Binocular vs. monocular**	1.5	0.022	<0.001	0.043
		3.0	**0.063**	**0.041**	**0.085**
		6.0	**0.091**	**0.069**	**0.11**
		all distances	**0.059**	**0.050**	**0.067**
Stationary	**Binocular vs. monocular**	1.5	-0.014	-0.079	0.051
		3.0	0.062	-0.0031	0.13
		6.0	**0.11**	**0.041**	**0.17**
		all distances	**0.051**	**0.026**	**0.077**

($p < 0.001$) and the effect sizes are shown in Table 4.11. Complete LMM results are shown in Table A.3.

Figure 4.13(b) shows the average Depth as a function of Fold Distance and B.19 shows the individual data. Compared to Experiment 1 and 2, this experiment further explored the Fold Distance manipulation by extending the range to up to 6 m, thus the effect of Fold Distance was amplified. Increasing Fold Distance produced an increasing effect on Depth in motion parallax blocks under both Viewing conditions (distance main effect, $p < 0.001$). The Fold Distance effect on depth was more pronounced under binocular Viewing (interaction between Viewing and Fold Distance, $p < 0.001$), consistent with a stronger effect of Fold Distance on binocular Depth than monocular and the difference between binocular and monocular Depths increasing with increasing Fold Distance. The effect of Fold Distance under binocular Viewing was approximately 1.5–2 times as big as it was under monocular Viewing, when comparing under the same conditions (see Table 4.12). Under stationary conditions, the effect of Fold

Table 4.12: GLHT results of Fold Distance effect on Depth under different Viewing conditions and Gain of 1 in Experiment 3. Fold Distance is categorical but its linear and quadratic effect were also tested.

Motion condition	Viewing mode	Linear effect	Quadratic effect	Pairwise contrast (m)	Pairwise effect (°)	2.5% bound (m)	97.5% bound (m)
Motion parallax	**Monocular**	**0.055**	-0.004	**3.0 vs. 1.5**	**0.030**	**0.012**	**0.047**
				6.0 vs. 1.5	**0.055**	**0.038**	**0.073**
				6.0 vs. 3.0	**0.025**	**0.008**	**0.043**
	Binocular	**0.124**	-0.018	**3.0 vs. 1.5**	**0.071**	**0.053**	**0.088**
				6.0 vs. 1.5	**0.124**	**0.106**	**0.141**
				6.0 vs. 3.0	**0.053**	**0.035**	**0.071**
Stationary	Monocular	0.014	-0.016	3.0 vs. 1.5	0.015	-0.038	0.069
				6.0 vs. 1.5	0.014	-0.039	0.068
				6.0 vs. 3.0	-0.001	-0.054	0.052
	Binocular	**0.135**	-0.048	**3.0 vs. 1.5**	**0.092**	**0.038**	**0.145**
				6.0 vs. 1.5	**0.135**	**0.082**	**0.188**
				6.0 vs. 3.0	0.043	-0.010	0.097

Distance was significant only under binocular Viewing (interaction between Viewing and Fold Distance, $p < 0.001$).

Table 4.13: GLHT results of Viewing effect on Pole Distance at different Fold Distances under Gain of 1 in Experiment 3.

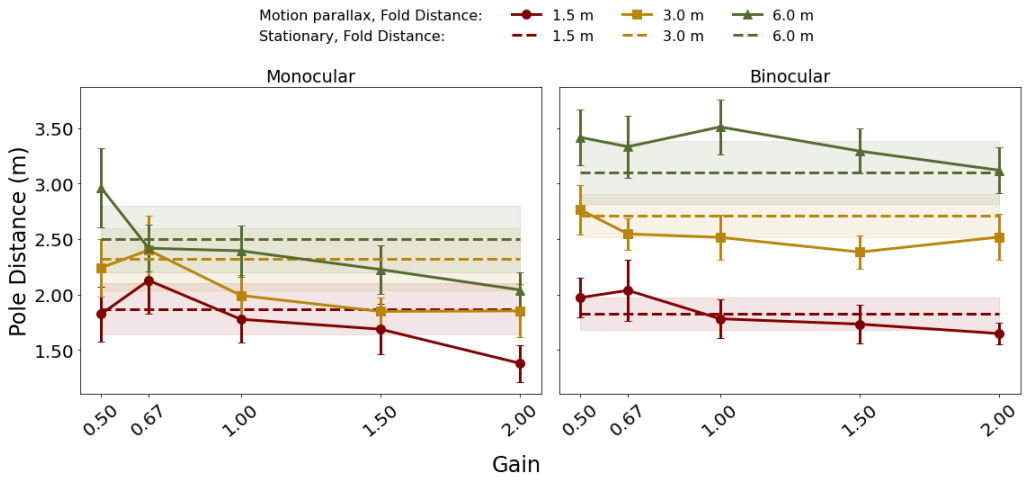
Motion Status	Viewing	Fold Distance (m)	Effect (m)	2.5% bound (m)	97.5% bound (m)
Motion parallax	**Binocular vs. monocular**	1.5	0.047	-0.176	0.270
		3.0	**0.450**	**0.227**	**0.673**
		6.0	**0.900**	**0.676**	**1.124**
		all distances	**0.466**	**0.335**	**0.597**
Stationary	Binocular vs. monocular	1.5	-0.217	-1.058	0.624
		3.0	0.221	-0.620	1.062
		6.0	0.428	-0.413	1.269
		all distances	0.144	-0.507	0.796

4.5.2 Pole Distance

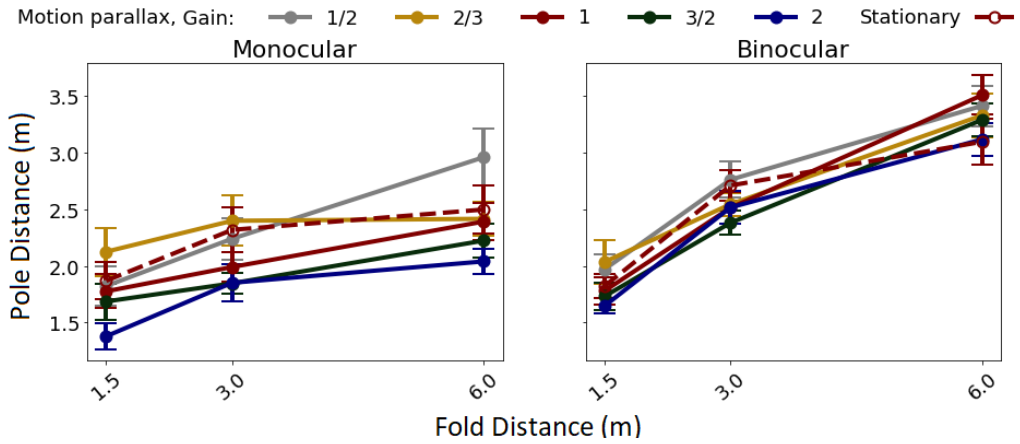
Figure 4.14 shows the Pole Distance data averaged across all participants. Individual plots are shown in Figure B.20. Observers compressed the range of absolute Fold Distance under both Viewing conditions, especially under expansive Gains ($g > 1$) or at far distances.

An interaction between Viewing and Fold Distance on Pole Distance was observed, see Table 4.13 for effect sizes. Binocular Viewing resulted in bigger Pole Distances at Fold Distances of 3.0 and 6.0 m. Increasing Gain produced increasingly underestimated Pole Distance responses under both Viewing conditions (main effect of Gain on Pole Distance, $p < 0.001$). Effect sizes are shown in Table 4.14. Pole Distance was underestimated on average by 62.5% and 60.2% under monocular and binocular Viewing, respectively.

Figures 4.14(b) shows the Pole Distance plotted as a function of Fold Distance. Pole Distance increased with Fold Distance in both motion parallax



(a) Pole Distance settings as a function of Gain averaged across all observers in Experiment 3. Stationary responses are plotted as a line across the x-axis for easy visualization.



(b) Pole Distance settings as a function of Fold Distances averaged across all observers in Experiment 3.

Figure 4.14: Pole Distance settings averaged across all observers in Experiment 3. Error bars and regions around dashed lines represent ± 1 SEM.

Table 4.14: GLHT results of Gain effect on Pole Distance under different Viewing conditions in Experiment 3.

Viewing	Gain effect (m)	2.5% bound (m)	97.5% bound (m)
Monocular	**-0.401**	**-0.587**	**-0.215**
Binocular	**-0.191**	**-0.379**	**-0.005**

and stationary blocks (both $p < 0.001$, effect sizes shown in Table 4.15). This effect indicates that observers had correct perception of relative distance. The interaction between Viewing and Fold Distance on Pole Distance was significant only in motion parallax blocks ($p < 0.001$). Increasing Fold Distance produced more underestimation on Pole Distance at 3.0 and 6.0 m than at 1.5 m under both Viewing conditions. This could be because at far Fold Distances there was not enough support for distance perception even under binocular Viewing, since binocular disparity and convergence angle were too small. Observers indicated more, yet still underestimated, Pole Distance under binocular Viewing compared to monocular Viewing.

4.5.3 Prediction

Prediction results were similar to Experiment 1 and 2. Observers set monocular folds flatter and closer throughout the range of increasing Gain at 1.5 and 3.0 m while the model predicted similar decreasing perceptions under both Viewings at all distances but only for $g > 1$ conditions.

To conclude Experiment 3, the same effect of decreasing Depth with increasing Gain under monocular Viewing persisted at 1.5 and 3.0 but not 6.0 m.

Table 4.15: GLHT results of Fold Distance effect on Pole Distance under different Viewing conditions and Gain of 1 in Experiment 3. Fold Distance is categorical but its linear and quadratic effect were also tested.

Motion condition	Viewing mode	Linear effect	Quadratic effect	Pairwise contrast (m)	Pairwise effect (m)	2.5% bound (m)	97.5% bound (m)
Motion parallax	**Monocular**	**0.648**	**0.034**	**3.0 vs. 1.5**	**0.307**	**0.043**	**0.572**
				6.0 vs. 1.5	**0.649**	**0.384**	**0.913**
				6.0 vs. 3.0	**0.341**	**0.077**	**0.605**
	Binocular	**1.501**	0.081	**3.0 vs. 1.5**	**0.710**	**0.446**	**0.975**
				6.0 vs. 1.5	**1.501**	**1.236**	**1.766**
				6.0 vs. 3.0	**0.791**	**0.526**	**1.056**
Stationary	**Monocular**	**0.629**	−0.270	3.0 vs. 1.5	0.449	-0.188	1.087
				6.0 vs. 1.5	0.629	-0.019	1.277
				6.0 vs. 3.0	0.180	-0.468	0.828
	Binocular	**1.274**	−0.500	**3.0 vs. 1.5**	**0.887**	**0.239**	**1.535**
				6.0 vs. 1.5	**1.274**	**0.626**	**1.922**
				6.0 vs. 3.0	0.387	-0.261	1.035

What's more, the same Fold Distance effect on Depth were observed not only under the same binocular motion parallax block as in Experiment 2 but also seen when either binocular or motion parallax was present. However, this effect was not present at 6.0 m. Larger Fold Distances produced a large gap between the ground truth Fold Distance and Pole Distance. This gap may be because the observers did not perceive the absolute Fold Distance accurately but only relative distance and depth, which may explain why they were able to ignore extremely unnatural Gain distortion. In spite of the perceptual compression on absolute distance, relative distance and depth relations were less sensitive and adjusted more closely to veridical.

Chapter 5

General Discussion and Conclusions

5.1 Compression of virtual space

Across all three experiments, folds were set as flatter with smaller Depth than a 90° angle under monocular Viewing. An interesting finding is that under impoverished depth/distance cue environment, Depth settings for the adjusted fold decreased with Fold Distance over the 1.3 to 1.5 m range tested under monocular or stationary Viewing. This could be a reflection of Gogel's specific distance tendency where, in the absence of reliable distance information, observers localize objects at a 'default' distance of typically around 2 m (Gogel and Tietz 1973). In our case, the default distance was further than the Fold Distances in Experiment 1. Thus, for that experiment, I predict that observers should see the dihedral fold as relatively stretched in depth if the specific

distance tendency is operating and this should be more pronounced for nearer Fold Distances which are further from the specific distance.

Introducing additional depth/distance cues via motion or binocular Viewing (with stereopsis), or extending the Fold Distance range to beyond the specific distance (i.e., 3.0 and 6.0 m) eliminated this specific distance effect or produced an opposite effect where Depth settings increased with Fold Distance. These variations in the effect of Fold Distance can be described in the confusion matrices in Table 5.1, where a ⊖ or ⊕ denotes the presence of a statistically significant effect in the negative or positive direction while an ✗ indicates no significant effect was found.

Table 5.1: Confusion Matrices for the effect of Fold Distance on Depth settings in Experiment 1, 2 and 3.

(a) Effect of Fold Distance on Depth - Experiment 1.

Range 1.3–1.5 m		Stereo	
		Yes	No
Motion Parallax	Yes	✗	✗
	No	✗	⊖

(b) Effect of Fold Distance on Depth - Experiment 2.

Range 1.3–1.5 m		Stereo	
		Yes	No
Motion Parallax	Yes	⊕	✗
	No	✗	⊖

(c) Effect of Fold Distance on Depth - Experiment 3.

Range 1.5–6.0 m		Stereo	
		Yes	No
Motion Parallax	Yes	⊕	⊕
	No	⊕	✗

Our results showed that observers were partially able to ignore or compensate for the effect of extreme Gain distortion on Depth and Fold Distance. This

is consistent with previous reports that the central nervous system is tolerant to visual-otolith conflicts generated by brief, low-acceleration head movements (Kim et al. 2021). Observers compressed virtual space in depth, which is aligned with the prediction that observers underestimated the expansive Gain while performing motion parallax (Figure 4.6). These results are also consistent with previous findings of general underestimation of distance in virtual reality at least for far distances (Gilinsky 1951; Johansson 1973). This tendency would also be relevant during the Pole Distance matching task but to a lesser extent due to the richer depth cue environment compared to the angle task. Thus, observers likely perceived even more compression of Depth than their already compressive Pole Distance matches indicated.

5.2 Role of Stereopsis

Previous research has shown that binocular cues are important when discriminating 3D surface orientation and depth (Allison, Gillam, and Vecellio 2009; Tsutsui et al. 2001). Allison et al. showed stereoscopic information was useful during ground surface discrimination tasks at the largest distances we used here and beyond (Allison, Gillam, and Palmisano 2009). In Oluk et al.'s experiments, the performance of slant perception task was completely determined by binocular cues given monocular cues such as perspective projection (Oluk et al. 2022). I extended these observations indicating the importance of stereoscopic depth cues in calibrating depth perception and maintaining perceptual stability in a VE. Binocular Viewing produced smaller dependence in portrayed Depth and

Pole Distance settings on Gain than observed when viewing monocularly.

5.3 Gain and stability in VR

Over three experiments I evaluated how motion Gain affects the perception of depth/shape. Information from motion parallax has been found to be a powerful cue to the shape and relative depth of three-dimensional surfaces (Rogers and Graham 1979b; Rogers and Graham 1982). Therefore, I reasoned that Gain distortions that scaled motion parallax would impact perception of 3D layout and object shape when combined with other cues. Motion gain could be introduced by various factors in current AR/VR systems including tracking error or poor calibration, display scaling that does not match the field of view (Kuhl, Thompson, and Creem-Regehr 2009; Psotka, Lewis, and King 1998), artistic intent to simulate large or small spaces or third-person views (Salamin, Thalmann, and Vexo 2006), or to serve interaction in approaches such as re-directed walking (Reimer et al. 2020; Steinicke et al. 2009; Razzaque, Kohn, and Whitton 2001) or seven-league boots (Interrante, Ries, and Anderson 2007).

Previous work has shown that users have considerable tolerance to gain distortions. For example, it has been demonstrated that gains greater than 1.8 are tolerated before users experienced overt instability of the world known as oscillopsia (Jaekl, Jenkin, and Harris 2005), although there was bias and less tolerance for gain reduction. Cybersickness has been associated with gain distortions but modest distortions are generally tolerable and Selzer, Larrea,

and Castro (2022) found significant cybersickness only for gains greater than 2.0. During actual locomotion (for example redirected walking), users are reportedly insensitive to a wide range of gain manipulation from 0.86 to 1.26 (Steinicke et al. 2010) although rich visuals and a view of one's feet may reduce this tolerance (Kruse, Langbehn, and Steinicke 2018).

I studied the impact of Gain over a range of distances (up to 6 m). Our results have interesting implications for the fundamental mechanisms of human vision and their impact on VR usability. Observers tended to distort apparent distance and object size in order to compensate for the Gain manipulation, favoring a stable world with less motion distortion. Figure 5.1 illustrates this distortion of perceived dihedral angle when the Self Motion is accurately perceived. Red angles depict folds set by the observer in the computer graphic (CG) virtual world and the black angles represent the corresponding right angle perceptions predicted by the model. The two folds are rendered at the same egocentric distance (Fold Distance) but they yield different perceived fold distances. Perceived fold distance of each black right angle was determined using a scaled world where the perceived right-angled fold (black angle) keeps its lateral position unchanged–directly in front of the observer's original position (grey eye).

The left depiction illustrates compressive Gain distortion ($g < 1$) where the virtual motion shown on the HMD is less than the physical motion made by the observer. According to Equation 3.5, this is achieved by moving the virtual object (red angle) in the same direction as the observer sways to the right. The right depiction illustrates expansive Gain distortion ($g > 1$), where the virtual

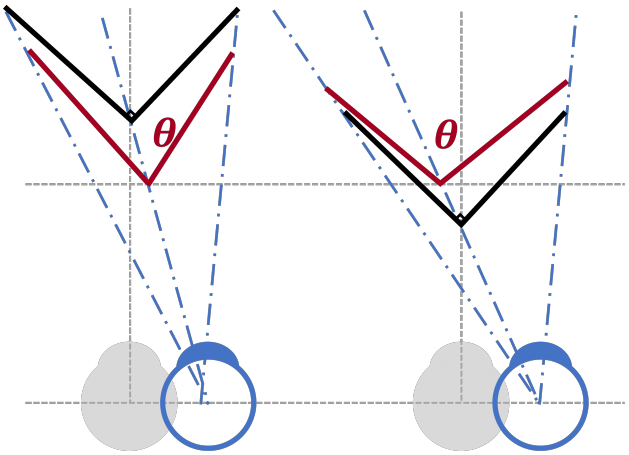


Figure 5.1: Two predictions in space under Gain distortions. Left: compressive ($g < 1$), right: expansive ($g > 1$).

object moves in the opposite direction, providing the scene as if the observer swayed more than the actual physical motion.

The figure shows that observers are expected to perceive more slant from a given fold under expansive Gain compared to compressive gain. Therefore users select a larger angle (flatter) fold in the CG model while reported seeing it as a 90° fold. Symmetrical reasoning predicts a smaller angle setting under compressive gain. Results for Pole Distance showed the same trend as the distance shifts in black angles in Figure 5.1. I established this geometrical model based on three assumptions: accurate perception of Self Motion, Gain distortion and object constancy. In order for the model to arrive at a solution, it predicted that different mechanisms were responsible for Depth perception under compressive versus expansive Gain distortions. Under compressive Gains (left), the model predicted observers perceived both Self Motion and Gain distortion correctly and compensated by seeing a right-angled fold in a larger scale (size) at a further distance. Conversely, under expansive Gains (right), it suggested

that observers compressed the absolute distance and size to account for the Gain distortion. Our model suggested under expansive gain, observers perceived less Gain than rendered (Figure 4.6). If this were true then the observers would not need to scale perceived size or distance fully. This prediction remains a focus for future studies.

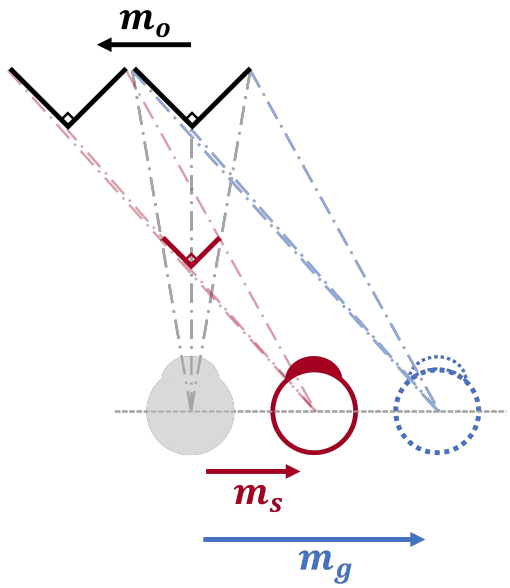


Figure 5.2: Depictions of expansive ($g > 1$) Gain effect on perceived fold distance.

LeClair et al. showed that monocular motion parallax from lateral sway produced the impression of relative motion, not depth, between objects (LeClair and Durgin 2010). Simulated objects appeared to become non-rigid as viewing distance increased. Our prediction and portrayed depth results seem to reflect similar processing. Figure 5.2 demonstrates the situation in which the observer makes a rightward head motion m_s (red) which is amplified by g to produce m_g and renders the fold from this vantage point as shown by the blue lines of sight.

If the observer experiences self-motion as m_g then the world should appear at the correct scale and stationary. The lack of an influence of Gain under binocular viewing suggests that this interpretation might have been preferred by the visual system when stereopsis was available to provide additional cues to scale and distance. If instead they experience their self-motion accurately then they should (1) perceive the object moving laterally as m_o, (2) perceive a stationary, rigid but scaled version of the scene as in the red fold, or (3) some compromise between scale and object motion. Our geometrical modelling, combined with the finding that Pole Distance varied with Gain, suggests observers were adopting interpretation 2 or 3.

These interpretations have interesting practical applications for VR. Flexibility in the mapping between real and virtual space is a key feature of numerous VR applications. While this study demonstrated that users' perception of Depth and Fold Distance were distorted when moving through scaled virtual space, the extent of the distortions were relatively small compared to the Gain distortions applied. The freedom to manipulate motion Gain in VR is determined by the distance range in the scene and the richness of the depth cues, particularly binocular vision. In the presence of rich depth cues where the scale of the scene is unambiguous, users seem quite tolerant to a large range of Gain manipulations. Conversely with impoverished depth cues or at close distances users are susceptible to illusory motion or scale distortions. Thus, developers may want to apply different rendering protocols at different virtual distances to tailor the experience to the environment.

5.4 Conclusions

Interaction with virtual objects in VR is a complex process and could be affected by a number of factors, among which perception of the depth and shape of the object are of vital importance. My thesis contribute to the literature on depth and shape perception in the presence of motion gain. In the past, the impact of motion gain on distance and shape perception have been studied separately. I systematically studied the effect of translational motion gain on depth and distance perception at the same time. In particular, my experiments covered not only close but also beyond moderate distances, i.e., 1.3 to 6.0 m. I found that Depth of the apparent right angle settings decreased with increasing Gain under monocular Viewing. On the contrary, Fold Distance had an increasing effect on depth under binocular Viewing when observers had the expectation that Fold Distance might be varied in rich depth/distance cue environment (Experiment 2 and 3). Relative distance perception was remarkably accurate while absolute distance was compressed especially at far distances. No interactions between Gain and Fold Distance were found. When both effects were present, the observations were: depth at each Fold Distance decreased with increases in gain, and with increases in Fold Distance. An interesting finding in the monocular stationary blocks of Experiment 1 and 2 was that when observers did not explicitly estimate Fold Distance and may not have been aware of the difference in Fold Distance, they acted as though they perceived all the folds located at a similar position and thus an 'inverse' (decreasing) effect of Fold Distance on adjusted depth was observed in the monocular stationary condition.

Motion parallax or binocular cues would contradict this impression since they both provide support that improved distance perception.

Furthermore, a mathematical model was proposed to predict observers' perceptions under the experimental manipulations. With the model, the theoretical values for Depth and Fold Distance perceptions were calculated. The model provided some insights on motion/distance compensation that potentially explained the increasing underestimation on Pole Distance. However, since this model did not include stereopsis information, the differences between binocular and monocular predictions was minimum. Another finding about this model is that it assumes observers perceived Gain perfectly and that they contributed almost all the Gain distortion as object motion. The fact that they did not means the observers needed to scale their perception on self motion accordingly. This implies that observers were able to ignore extremely unnatural motions, particularly under binocular viewing and motion parallax since they provided conflicting information against motion distortion.

The findings are especially useful for immersive VR devices. One of the most common application of VR technology is streaming high quality content, e.g., 4K or even 8K video. According to my results, humans are more tolerant to rendering mismatches in distant targets. Therefore, content creators can selectively focus on the quality of stimuli in near space, while sacrificing some of the distant contents if necessary. This can be extended to any VR system, commercial or academic, that requires the creation of contents at far distances.

5.5 Future work

In this thesis, I conducted systematic experiments to study how visual distortions affect depth and distance perception. The basic setting in all my experiments was that observers did not leave the standing point. Even for the motion parallax condition, the observers only did swaying around a fixed point. In most virtual reality setups, visual motion range may greatly exceed physical motion range. Thus, mappings or transformations are inevitable for smooth VR experience (McCauley and Sharkey 1992; Rieser et al. 1995; Thompson et al. 2011). Future experiments will be designed to study how physical motion, e.g., walking, affects depth perception, hopefully to explore the visual constraints in HMD VR.

Bibliography

Allison, Robert S, Barbara J Gillam, and Stephen A Palmisano (2009). "Stereoscopic discrimination of the layout of ground surfaces". In: *Journal of Vision* 9.12, pp. 8–8.

Allison, Robert S, Barbara J Gillam, and Elia Vecellio (2009). "Binocular depth discrimination and estimation beyond interaction space". In: *Journal of Vision* 9.1, pp. 10–10.

Anstis, Stuart (1995). "Aftereffects from jogging". In: *Experimental Brain Research* 103.3, pp. 476–478.

Armbrster, Claudia et al. (2008). "Depth perception in virtual reality: distance estimations in peri-and extrapersonal space". In: *Cyberpsychology & Behavior* 11.1, pp. 9–15.

Bates, Douglas et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Software* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01.

Boyd, D Eric and Bernadett Koles (2019). "Virtual reality and its impact on B2B marketing: A value-in-use perspective". In: *Journal of Business Research* 100, pp. 590–598.

Braunstein, Myron L et al. (1986). "Recovering viewer-centered depth from disparity, occlusion, and velocity gradients". In: *Perception & Psychophysics* 40.4, pp. 216–224.

Brument, Hugo et al. (2021). "Studying the Influence of Translational and Rotational Motion on the Perception of Rotation Gains in Virtual Environments". In: *Symposium on Spatial User Interaction*, pp. 1–12.

Burton, Harry Edwin (May 1945). "The Optics of Euclid". In: *Journal of the Optical Society of America* 35.5, p. 357. DOI: 10.1364/josa.35.000357.

Campagnoli, Carlo, Bethany Hung, and Fulvio Domini (2022). "Explicit and implicit depth-cue integration: evidence of systematic biases with real objects". In: *Vision research* 190, p. 107961.

Cheng, Alan, Lei Yang, and Erik Andersen (2017). "Teaching language and culture with a virtual reality game". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 541–549.

Community, Blender Online (2018). "Blender - a 3D modelling and rendering package". In: URL: http://www.blender.org.

Cornilleau-Pérès, V and CCAM Gielen (1996). "Interactions between self-motion and depth perception in the processing of optic flow". In: *Trends in Neurosciences* 19.5, pp. 196–202.

Cutone, Matthew and Laurie Wilcox (2021). *PsychXR (Version 0.2.4)*. https://github.com/mdcutone/psychxr/.

Cutone, Matthew, Laurie Wilcox, and Robert Allison (Oct. 2020). "The impact of motion gain on egocentric distance judgments from motion parallax". In:

Journal of Vision 20.11, p. 1426. DOI: 10.1167/jov.20.11.1426. URL: https://doi.org/10.1167/jov.20.11.1426.

Cutting, James E and Peter M Vishton (1995). "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth". In: *Perception of space and motion*. Elsevier, pp. 69–117.

Eggleston, Robert G, William P Janson, and Kenneth A Aldrich (1996). "Virtual reality system effects on size-distance judgements in a virtual environment". In: *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*. IEEE, pp. 139–146.

Ferris, Steven H. (1972). "Motion parallax and absolute distance." In: *Journal of Experimental Psychology* 95.2, pp. 258–263. DOI: 10.1037/h0033605.

Gibson, James J (1950). *The perception of the visual world*. Houghton Mifflin.

Gibson, James J and Leonard Carmichael (1966). *The senses considered as perceptual systems*. Vol. 2. 1. Houghton Mifflin Boston.

Gilinsky, Alberta S (1951). "Perceived size and distance in visual space". In: *Psychological review* 58.6, p. 460.

Gogel, Walter C and Jerome D Tietz (1973). "Absolute motion parallax and the specific distance tendency". In: *Perception & Psychophysics* 13.2, pp. 284–292.

Goldstein, E. Bruce and James Brockmole (2016). *Sensation and Perception*. Nelson Education.

Hettinger, Lawrence J and Gary E Riccio (1992). "Visually induced motion sickness in virtual environments". In: *Presence: Teleoperators & Virtual Environments* 1.3, pp. 306–310.

Hothorn, Torsten, Frank Bretz, and Peter Westfall (2008). "Simultaneous Inference in General Parametric Models". In: *Biometrical Journal* 50.3, pp. 346–363.

Howard, Ian P (2012). *Sensation and Perception*. New York: Oxford University Press.

Howard, Ian P and Brian J Rogers (1995). *Binocular vision and stereopsis*. New York: Oxford University Press.

Interrante, V., B. Ries, and L. Anderson (Mar. 2007). "Seven League Boots: A New Metaphor for Augmented Locomotion through Moderately Large Scale Immersive Virtual Environments". In: *IEEE Symposium on 3D User Interfaces, 2007. 3DUI '07.* Charlotte, NC: IEEE, pp. 167–170. DOI: 10.1109/3DUI.2007.340791.

Jaekl, P. M., M. R. Jenkin, and Laurence R. Harris (Apr. 2005). "Perceiving a stable world during active rotational and translational head movements". In: *Experimental Brain Research* 163.3, pp. 388–399. DOI: 10.1007/s00221-004-2191-8. URL: https://doi.org/10.1007/s00221-004-2191-8.

Johansson, Gunnar (1973). "Monocular movement parallax and near-space perception". In: *Perception* 2.2, pp. 135–146.

Kavanagh, Sam et al. (2017). "A systematic review of Virtual Reality in education". In: *Themes in Science and Technology Education* 10.2, pp. 85–119.

Kim, Juno et al. (Apr. 2021). "Effects of Linear Visual-Vestibular Conflict on Presence, Perceived Scene Stability and Cybersickness in the Oculus Go and Oculus Quest". In: 2. DOI: 10.3389/frvir.2021.582156. URL: https://doi.org/10.3389/frvir.2021.582156.

Koenderink, Jan J and Andrea J Van Doorn (1975). "Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer". In: *Optica acta* 22.9, pp. 773–791.

Kruse, Lucie, Eike Langbehn, and Frank Steinicke (2018). "I Can See on My Feet While Walking: Sensitivity to Translation Gains with Visible Feet". In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 305–312. DOI: 10.1109/VR.2018.8446216.

Kuhl, Scott A., William B. Thompson, and Sarah H. Creem-Regehr (Sept. 2009). "HMD Calibration and Its Effects on Distance Judgments". In: *ACM Trans. Appl. Percept.* 6.3, 19:1–19:20. ISSN: 1544-3558. DOI: 10.1145/1577755.1577762. URL: http://doi.acm.org/10.1145/1577755.1577762 (visited on 07/15/2019).

LeClair, A. and F. H. Durgin (2010). "Depth interval perception: Comparing binocular stereopsis with motion parallax in "action space"". In: *Journal of Vision* 8.6, pp. 857–857. DOI: 10.1167/8.6.857.

Lugrin, Jean-Luc et al. (2018). "Any "body" there? avatar visibility effects in a virtual reality game". In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, pp. 17–24.

McCauley, Michael E and Thomas J Sharkey (1992). "Cybersickness: Perception of self-motion in virtual environments". In: *Presence: Teleoperators & Virtual Environments* 1.3, pp. 311–318.

Ohzawa, IZUMI and RALPH D Freeman (1986). "The binocular organization of complex cells in the cat's visual cortex". In: *Journal of Neurophysiology* 56.1, pp. 243–259.

Oluk, Can et al. (Apr. 2022). "Stereo slant discrimination of planar 3D surfaces: Frontoparallel versus planar matching". In: *Journal of Vision* 22.5, p. 6. DOI: 10.1167/jov.22.5.6. URL: https://doi.org/10.1167/jov.22.5.6.

Ono, Mika E., Josée Rivest, and Hiroshi Ono (1986). "Depth perception as a function of motion parallax and absolute-distance information." In: *Journal of Experimental Psychology: Human Perception and Performance* 12.3, pp. 331–337. DOI: 10.1037/0096-1523.12.3.331. URL: https://doi.org/10.1037%2F0096-1523.12.3.331.

Peirce, Jonathan W. (May 2007). "PsychoPy—Psychophysics software in Python". In: *Journal of Neuroscience Methods* 162.1-2, pp. 8–13. DOI: 10.1016/j.jneumeth.2006.11.017. URL: https://doi.org/10.1016/j.jneumeth.2006.11.017.

Philbeck, John W and Jack M Loomis (1997). "Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions." In: *Journal of Experimental Psychology: Human Perception and Performance* 23.1, p. 72.

Psotka, Joseph, Sonya A. Lewis, and Donald King (1998). "Effects of Field of View on Judgments of Self-Location: Distortions in Distance Estimations

Even When the Image Geometry Exactly Fits the Field of View". In: *Presence: Teleoperators and Virtual Environments* 7.4, pp. 352–369. ISSN: 1054-7460. DOI: 10.1162/105474698565776. URL: http://dx.doi.org/10.1162/105474698565776.

Razzaque, Sharif, Zachariah Kohn, and Mary C. Whitton (2001). "Redirected walking". In: *Proceedings of EUROGRAPHICS*. Vol. 9, pp. 105–106.

Reimer, Dennis et al. (Mar. 2020). "The Influence of Full-Body Representation on Translation and Curvature Gain". In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 154–159. DOI: 10.1109/VRW50115.2020.00032.

Rieser, John J et al. (1995). "Calibration of human locomotion and models of perceptual-motor organization." In: *Journal of Experimental Psychology: Human Perception and Performance* 21.3, p. 480.

Rogers, Brian and Maureen Graham (Apr. 1979a). "Motion Parallax as an Independent Cue for Depth Perception". In: *Perception* 8.2, pp. 125–134. DOI: 10.1068/p080125. URL: https://doi.org/10.1068/p080125.

— (1979b). "Motion parallax as an independent cue for depth perception". In: *Perception* 8.2, pp. 125–134.

— (Jan. 1982). "Similarities between motion parallax and stereopsis in human depth perception". In: *Vision Research* 22.2, pp. 261–270. DOI: 10.1016/0042-6989(82)90126-2. URL: https://doi.org/10.1016/0042-6989(82)90126-2.

Salamin, Patrick, Daniel Thalmann, and Frédéric Vexo (Nov. 2006). "The benefits of third-person perspective in virtual and augmented reality?" In: *Proceedings*

of the ACM Symposium on Virtual Reality Software and Technology. VRST '06. New York, NY, USA: Association for Computing Machinery, pp. 27–30. ISBN: 978-1-59593-321-8. DOI: 10.1145/1180495.1180502. URL: https://doi.org/10.1145/1180495.1180502 (visited on 10/01/2022).

Selzer, Matias Nicolas, Martin Leonardo Larrea, and Silvia Mabel Castro (2022). "Analysis of translation gains in virtual reality: the limits of space manipulation". In: *Virtual Reality*, pp. 1–11.

Shiffrar, Maggie and M. Pavel (1991). "Percepts of rigid motion within and across apertures." In: *Journal of Experimental Psychology: Human Perception and Performance* 17.3, pp. 749–761. DOI: 10.1037/0096-1523.17.3.749. URL: https://doi.org/10.1037/0096-1523.17.3.749.

Smith, Michael James, Gavriel Salvendy, and Richard J. Koubek (Aug. 1997). *Design of Computing Systems: Cognitive considerations*. ISBN: 9780444821836.

Stanney, Kay M and Robert S Kennedy (1997). "The psychometrics of cybersickness". In: *Communications of the ACM* 40.8, pp. 66–68.

Stanney, Kay M, Ronald R Mourant, and Robert S Kennedy (1998). "Human factors issues in virtual environments: A review of the literature". In: *Presence* 7.4, pp. 327–351.

Steinicke, Frank et al. (2009). "Estimation of detection thresholds for redirected walking techniques". In: *IEEE transactions on visualization and computer graphics* 16.1, pp. 17–27.

— (2010). "Estimation of Detection Thresholds for Redirected Walking Techniques". In: *IEEE Transactions on Visualization and Computer Graphics*

16.1, 17–27. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2009.62.

Surdick, R Troy et al. (1994). "Relevant cues for the visual perception of depth: is where you see it where it is?" In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 38. 19. SAGE Publications Sage CA: Los Angeles, CA, pp. 1305–1309.

Swan, J Edward et al. (2007). "Egocentric depth judgments in optical, see-through augmented reality". In: *IEEE transactions on visualization and computer graphics* 13.3, pp. 429–442.

Thompson, William et al. (2011). *Visual perception from a computer graphics perspective*. CRC press.

Todd, James T. and J. Farley Norman (Jan. 2003). "The visual perception of 3-D shape from multiple cues: Are observers capable of perceiving metric structure?" In: *Perception & Psychophysics* 65.1, pp. 31–47.

Tsutsui, Ken-Ichiro et al. (Dec. 2001). "Integration of Perspective and Disparity Cues in Surface-Orientation–Selective Neurons of Area CIP". In: *Journal of Neurophysiology* 86.6, pp. 2856–2867. DOI: 10.1152/jn.2001.86.6.2856. URL: https://doi.org/10.1152/jn.2001.86.6.2856.

Wallach, Hans and Carl Zuckerman (1963). "The constancy of stereoscopic depth". In: *The American journal of psychology* 76.3, pp. 404–412.

Waller, David and Adam R Richardson (2008). "Correcting distance estimates by interacting with immersive virtual environments: Effects of task and available sensory information." In: *Journal of Experimental Psychology: Applied* 14.1, p. 61.

Warren Jr, William H (1995). "Self-motion: Visual perception and visual control". In: *Perception of space and motion*. Elsevier, pp. 263–325.

Welch, RB and MM Cohen (1988). "Adaptation to Fixed and Variable Prismatic Displacement". In: *Bulletin of the Psychonomic Society*. Vol. 26. 6, pp. 510–510.

Wright, Robert H (1995). *Virtual Reality Psychophysics: Forward and Lateral Distance, Height, and Speed Perceptions with a Wide-Angle Helmet Display*. Tech. rep. ARMY RESEARCH INST FOR THE BEHAVIORAL and SOCIAL SCIENCES ALEXANDRIA VA.

Zeileis, Achim and Torsten Hothorn (2002). "Diagnostic Checking in Regression Relationships". In: *R News* 2.3, pp. 7–10. URL: https://CRAN.R-project.org/doc/Rnews/.

Appendix A

Complete LMM Results

Table A.1: LMM results for Experiment 1.

(a) Motion parallax conditions

Dependent Variable	Independent Variable	Sum Sq	Mean Sq	NumDF	DenDF	F Value	$Pr(>F)$
	Viewing	**763.18**	**763.18**	**1**	**1111**	**14.29**	**< 0.001**
	Distance	1.27	0.63	2	1111	0.012	0.99
	Gain	**1519.78**	**1519.78**	**1**	**1111**	**28.45**	**< 0.001**
Angle	**Width**	**740.82**	**370.41**	**2**	**1111**	**6.93**	**< 0.001**
	Viewing:Distance	275.50	137.75	2	1111	2.58	0.076
	Viewing:Gain	**1498.11**	**1498.11**	**1**	**1111**	**28.05**	**< 0.001**
	Viewing:Width	46.38	23.19	2	1111	0.43	0.65
	Viewing	**0.065**	**0.065**	**1**	**1111**	**15.49**	**< 0.001**
	Distance	1.8×10^{-4}	8.8×10^{-5}	2	1111	0.021	0.98
	Gain	0.11	0.11	**1**	**1111**	**26.93**	**< 0.001**
Depth	**Width**	**0.059**	**0.030**	**2**	**1111**	**7.01**	**< 0.001**
	Viewing:Distance	0.019	0.0095	2	1111	2.25	0.11
	Viewing:Gain	**0.12**	**0.12**	**1**	**1111**	**27.44**	**< 0.001**
	Viewing:Width	0.0043	0.0021	2	1111	0.51	0.60

(b) Stationary conditions

Dependent Variable	Independent Variable	Sum Sq	Mean Sq	NumDF	DenDF	F Value	$Pr(>F)$
	Viewing	**3088.56**	**3088.56**	**1**	**1115**	**46.77**	**< 0.001**
	Distance	**454.09**	**227.05**	**2**	**1115**	**3.44**	**0.032**
Angle	**Width**	**1301.10**	**650.55**	**2**	**1115**	**9.85**	**< 0.001**
	Viewing:Distance	**979.65**	**489.82**	**2**	**1115**	**7.42**	**< 0.001**
	Viewing:Width	373.07	186.53	2	1115	2.82	0.060
	Viewing	**0.23**	**0.23**	**1**	**1115**	**49.45**	**< 0.001**
	Distance	**0.039**	**0.020**	**2**	**1115**	**4.29**	**0.014**
depth	**Width**	**0.096**	**0.048**	**2**	**1115**	**10.42**	**< 0.001**
	Viewing:Distance	**0.075**	**0.037**	**2**	**1115**	**8.15**	**< 0.001**
	Viewing:Width	**0.032**	**0.016**	**2**	**1115**	**3.49**	**0.031**

Table A.2: LMM results for Experiment 2.

(a) Motion parallax conditions

Dependent Variable	Independent Variable	Sum Sq	Mean Sq	NumDF	DenDF	F Value	$Pr(>F)$
Angle	**Viewing**	**516.06**	**516.06**	**1**	**1114**	**15.73**	**< 0.001**
	Distance	41.86	20.93	2	1114	0.64	0.53
	Gain	**2434.34**	**2434.34**	**1**	**1114**	**74.21**	**< 0.001**
	Width	**1010.46**	**505.23**	**2**	**1114**	**15.40**	**< 0.001**
	Viewing:Distance	**482.48**	**241.24**	**2**	**1114**	**7.35**	**< 0.001**
	Viewing:Gain	**2275.17**	**2275.17**	**1**	**1114**	**69.36**	**< 0.001**
	Viewing:Width	**274.71**	**137.35**	**2**	**1114**	**4.19**	**0.015**
Depth	**Viewing**	**0.034**	**0.034**	**1**	**1114**	**13.95**	**< 0.001**
	Distance	0.0034	0.0017	2	1114	0.68	0.51
	Gain	**0.17**	**0.17**	**1**	**1114**	**68.62**	**< 0.001**
	Width	**0.071**	**0.035**	**2**	**1114**	**14.35**	**< 0.001**
	Viewing:Distance	**0.040**	**0.020**	**2**	**1114**	**8.04**	**< 0.001**
	Viewing:Gain	**0.17**	**0.17**	**1**	**1114**	**69.16**	**< 0.001**
	Viewing:Width	**0.021**	**0.010**	**2**	**1114**	**4.21**	**0.015**
Pole Distance	Viewing	0.1008	0.10	1	1114	1.12	0.29
	Distance	**4.23**	**2.12**	**2**	**1114**	**23.54**	**< 0.001**
	Gain	**4.85**	**4.85**	**1**	**1114**	**53.94**	**< 0.001**
	Width	0.073	0.037	2	1114	0.41	0.67
	Viewing:Distance	0.0026	0.0013	2	1114	0.015	0.99
	Viewing:Gain	**0.76**	**0.76**	**1**	**1114**	**8.47**	**0.004**
	Viewing:Width	0.045	0.023	2	1114	0.25	0.78

(b) Stationary conditions

Dependent Variable	Independent Variable	Sum Sq	Mean Sq	NumDF	DenDF	F Value	$Pr(>F)$
Angle	**Viewing**	**6677.1**	**6677.1**	**1**	**359.02**	**101.28**	**< 0.001**
	Distance	0.7	0.4	2	359.02	0.0054	0.99
	Width	**517.8**	**258.9**	**2**	**359.02**	**3.93**	**0.021**
	Viewing:Distance	389.4	194.7	2	359.02	2.95	0.053
	Viewing:Width	**561.7**	**280.9**	**2**	**359.02**	**4.2602**	**0.015**
depth	**Viewing**	**0.59**	**0.59**	**1**	**359.06**	**105.78**	**< 0.001**
	Distance	0.00018	0.00009	2	359.06	0.016	0.98
	Width	**0.047**	**0.024**	**2**	**359.06**	**4.24**	**0.015**
	Viewing:Distance	0.030	0.015	2	359.06	2.68	0.070
	Viewing:Width	**0.056**	**0.028**	**2**	**359.06**	**5.02**	**0.0071**
Pole Distance	Viewing	0.093	0.093	1	358.89	1.45	0.23
	Distance	**1.014**	**0.51**	**2**	**358.89**	**7.89**	**< 0.001**
	Width	0.28	0.14	2	358.89	2.20	0.11
	Viewing:Distance	0.097	0.049	2	358.89	0.76	0.47
	Viewing:Width	**0.51**	**0.26**	**2**	**358.89**	**3.98**	**0.020**

Table A.3: LMM results for Experiment 3.

(a) Motion parallax conditions

Dependent Variable	Independent Variable	Sum Sq	Mean Sq	NumDF	DenDF	F Value	$Pr(>F)$
Angle	**Viewing**	**846.5**	**846.5**	**1**	**522**	**22.34**	**< 0.001**
	Distance	**11789.7**	**5894.9**	**2**	**522**	**155.56**	**< 0.001**
	Gain	**937.8**	**937.8**	**1**	**522**	**24.75**	**< 0.001**
	Width	**994.0**	**497.0**	**2**	**522**	**13.12**	**< 0.001**
	Viewing:Distance	**2160.9**	**1080.4**	**2**	**522**	**28.51**	**< 0.001**
	Viewing:Gain	121.4	121.4	1	522	3.20	0.074
	Viewing:Width	142.1	71.0	2	522	1.87	0.15
Depth	**Viewing**	**0.048**	**0.048**	**1**	**522**	**19.08**	**< 0.001**
	Distance	**0.72**	**0.36**	**2**	**522**	**144.06**	**< 0.001**
	Gain	**0.061**	**0.061**	**1**	**522**	**24.21**	**< 0.001**
	Width	**0.065**	**0.033**	**2**	**522**	**13.06**	**< 0.001**
	Viewing:Distance	**0.11**	**0.054**	**2**	**522**	**21.42**	**< 0.001**
	Viewing:Gain	**0.0098**	**0.0098**	**1**	**522**	**3.89**	**0.049**
	Viewing:Width	0.011	0.0055	2	522	2.2092	0.11
Pole Distance	Viewing	1.69	1.69	1	522	2.98	0.085
	Distance	**103.80**	**51.90**	**2**	**522**	**91.25**	**< 0.001**
	Gain	**14.37**	**14.37**	**1**	**522**	**25.26**	**< 0.001**
	Width	0.28	0.14	2	522	0.25	0.78
	Viewing:Distance	**16.33**	**8.165**	**2**	**522**	**14.36**	**< 0.001**
	Viewing:Gain	1.80	1.80	1	522	3.17	0.076
	Viewing:Width	0.22	0.11	2	522	0.19	0.83

(b) Stationary conditions

Dependent Variable	Independent Variable	Sum Sq	Mean Sq	NumDF	DenDF	F Value	$Pr(>F)$
Angle	**Viewing**	**981.26**	**981.26**	**1**	**93**	**13.66**	**< 0.001**
	Distance	**1885.41**	**942.70**	**2**	**93**	**13.13**	**< 0.001**
	Width	**465.10**	**232.55**	**2**	**93**	**3.24**	**0.044**
	Viewing:Distance	**1134.81**	**567.40**	**2**	**93**	**7.90**	**< 0.001**
	Viewing:Width	107.91	53.95	2	93	0.75	0.47
depth	**Viewing**	**0.071**	**0.071**	**1**	**93**	**15.81**	**< 0.001**
	Distance	**0.11**	**0.053**	**2**	**93**	**11.84**	**< 0.001**
	Width	**0.030**	**0.015**	**2**	**93**	**3.28**	**0.042**
	Viewing:Distance	**0.067**	**0.034**	**2**	**93**	**7.45**	**< 0.001**
	Viewing:Width	0.0073	0.0036	2	93	0.81	0.45
Pole Distance	**Viewing**	**2.67**	**2.67**	**1**	**93**	**4.02**	**0.048**
	Distance	**17.18**	**8.59**	**2**	**93**	**12.90**	**< 0.001**
	Width	3.20	1.60	2	93	2.40	0.096
	Viewing:Distance	1.95	0.98	2	93	1.47	0.24
	Viewing:Width	0.56	0.28	2	93	0.42	0.66

Appendix B

Additional Figures

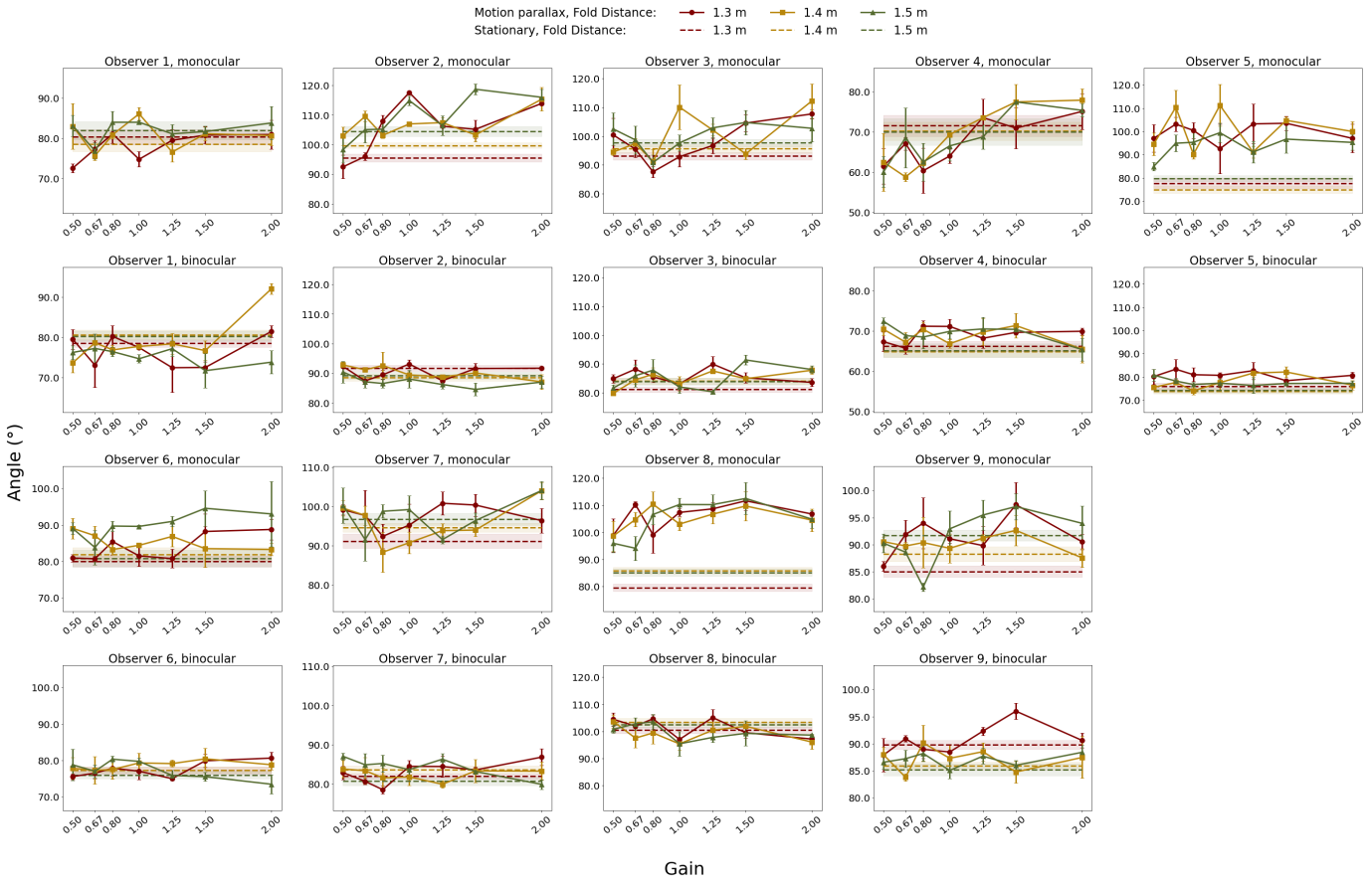
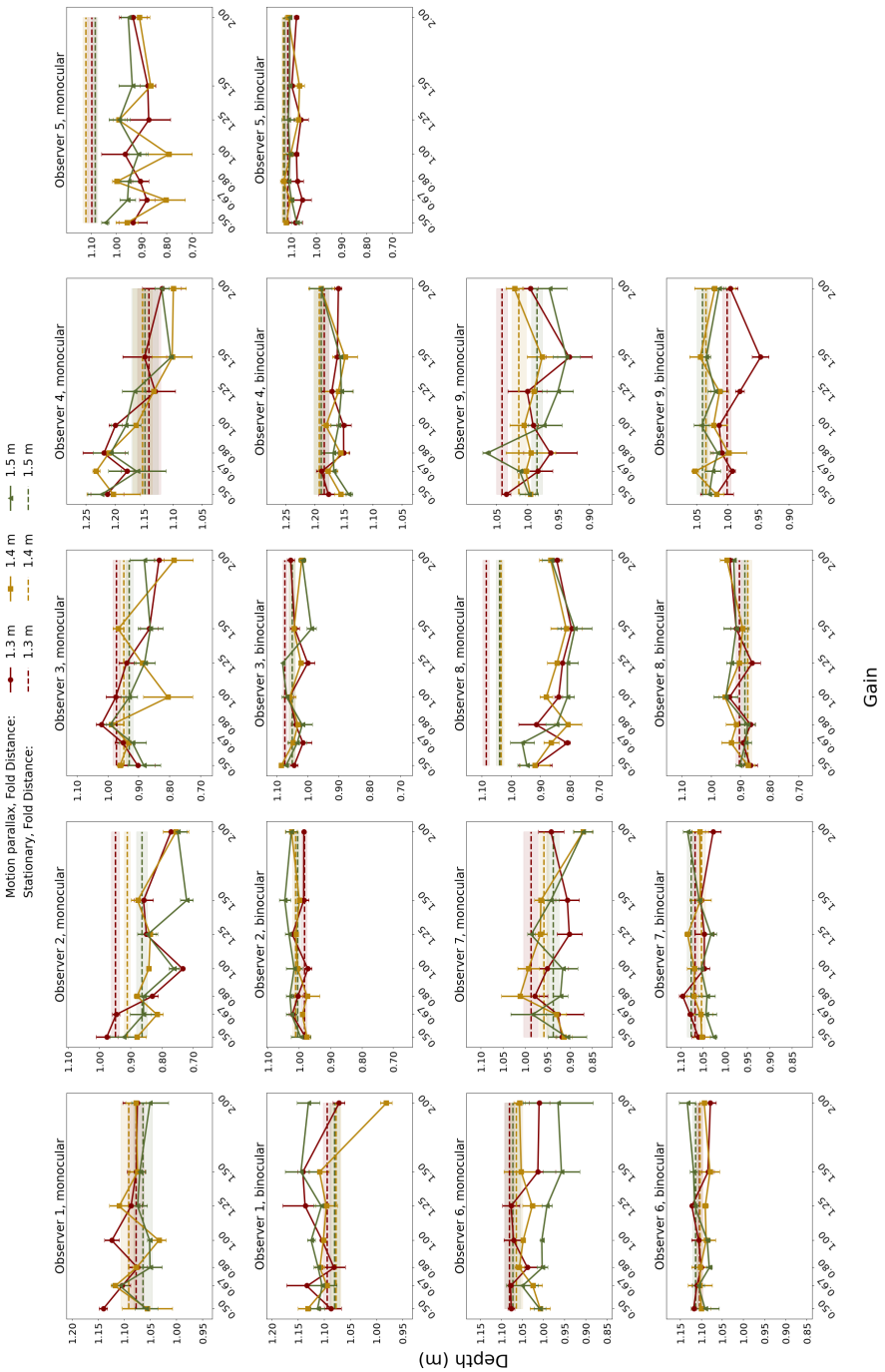
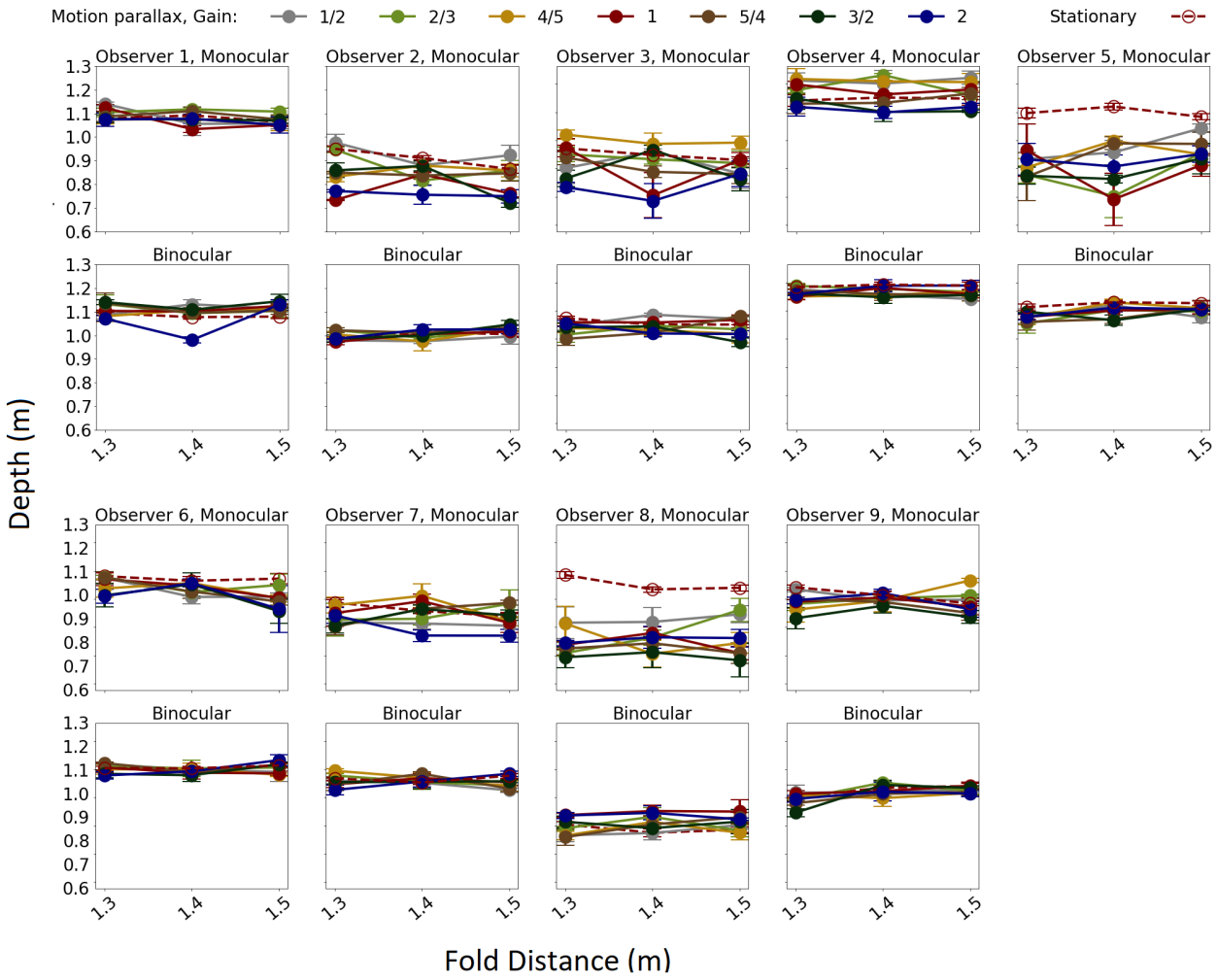


Figure B.1: Individual Angle settings as a function of Gain in Experiment 1. Stationary settings are plotted as a line across the x-axis for easy visualization. Error bars and regions around dashed lines represent ± 1 SEM.



(a) Individual Depth settings as a function of Gain in Experiment 1. Stationary settings are plotted as a line across the x-axis for easy visualization.



(b) Individual Depth settings as a function of Fold Distance in Experiment 1.

Figure B.2: Individual Depth settings in Experiment 1. Error bars and regions around dashed lines represent ± 1 SEM.

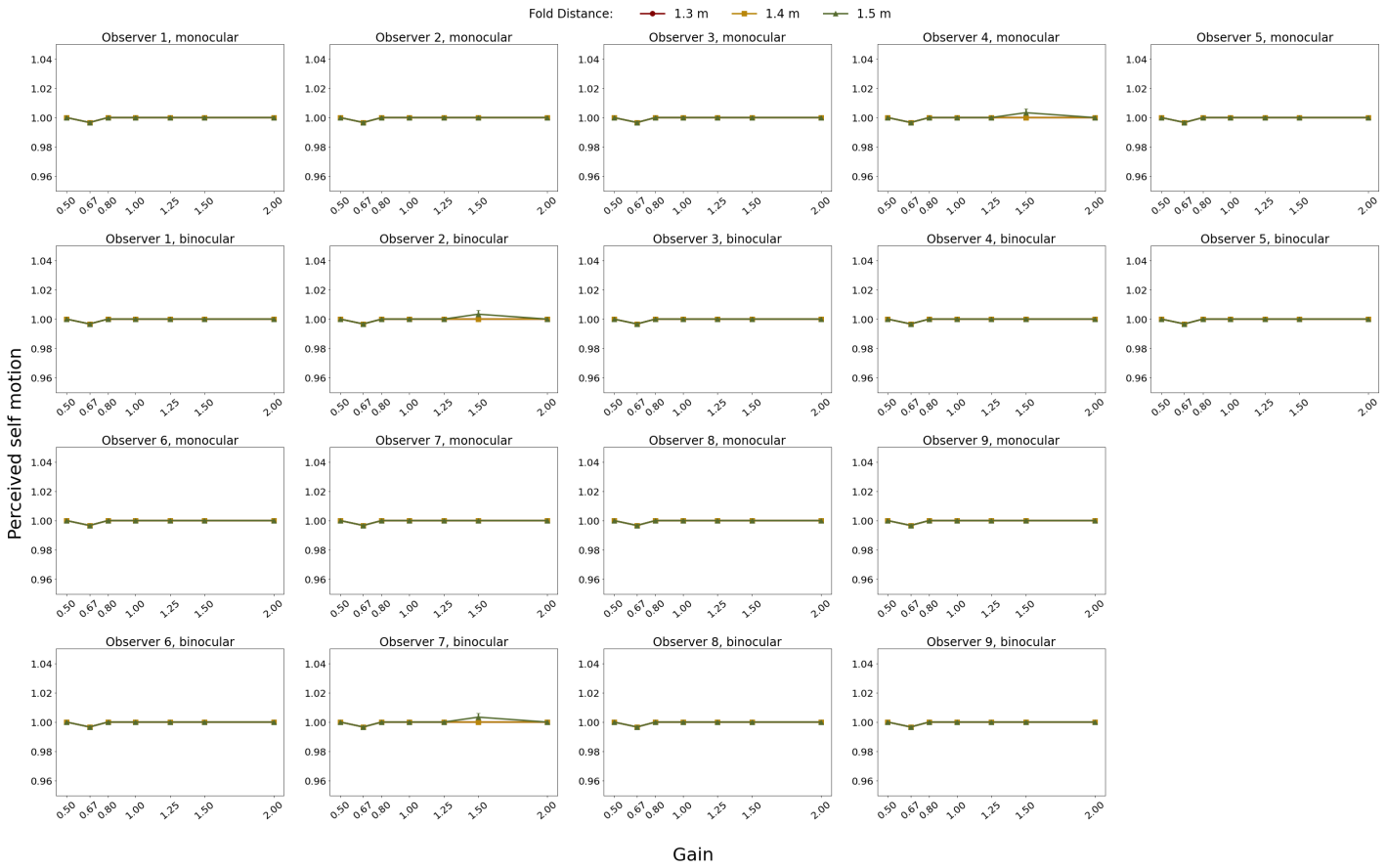


Figure B.3: Individual predictions of perceived self motion as a function of Gain in Experiment 1. Error bars (too small to be seen in this plot) represent ±1 SEM.

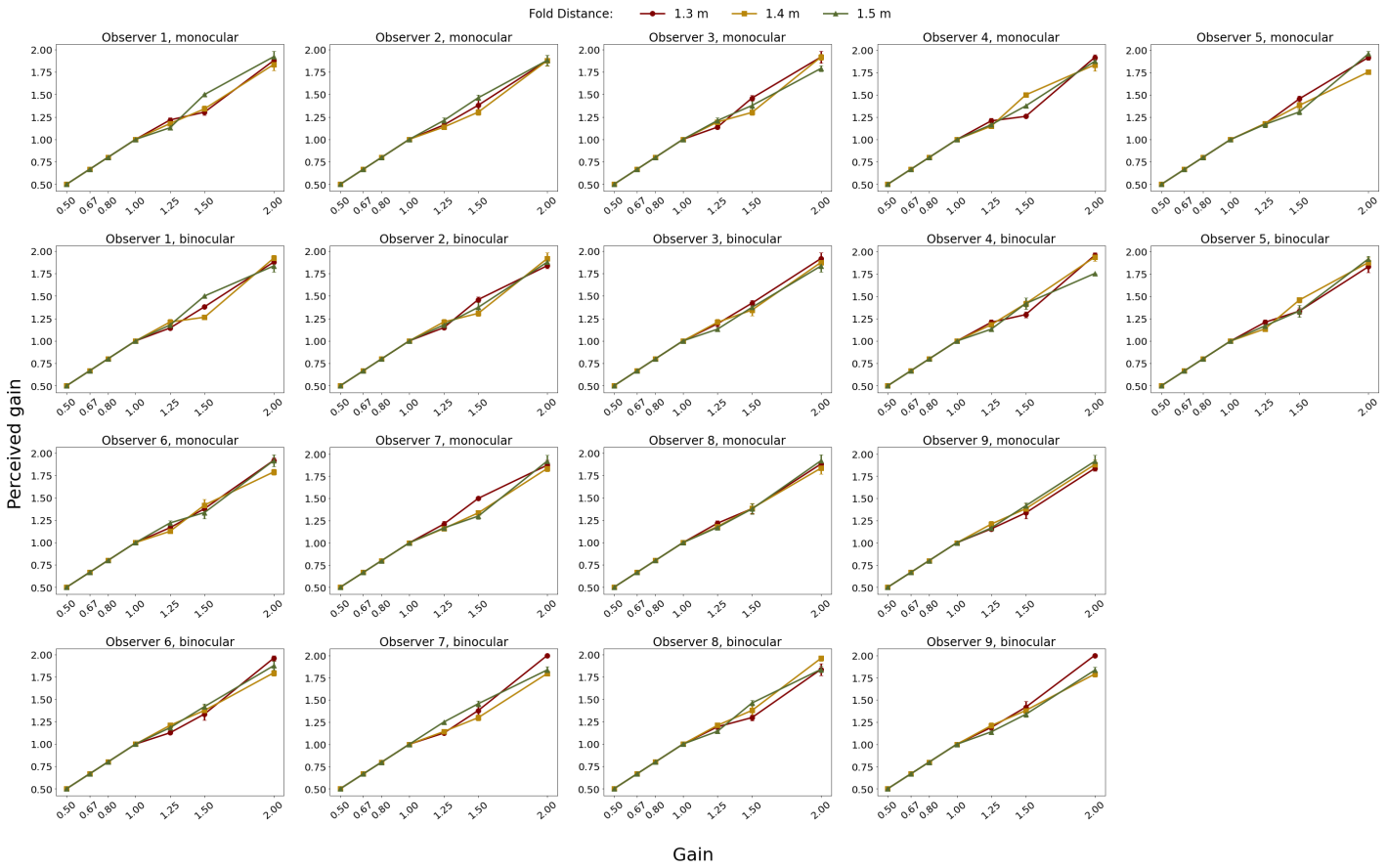


Figure B.4: Individual predictions of perceived gain as a function of Gain in Experiment 1. Error bars (some are too small to be seen in this plot) represent ±1 SEM.

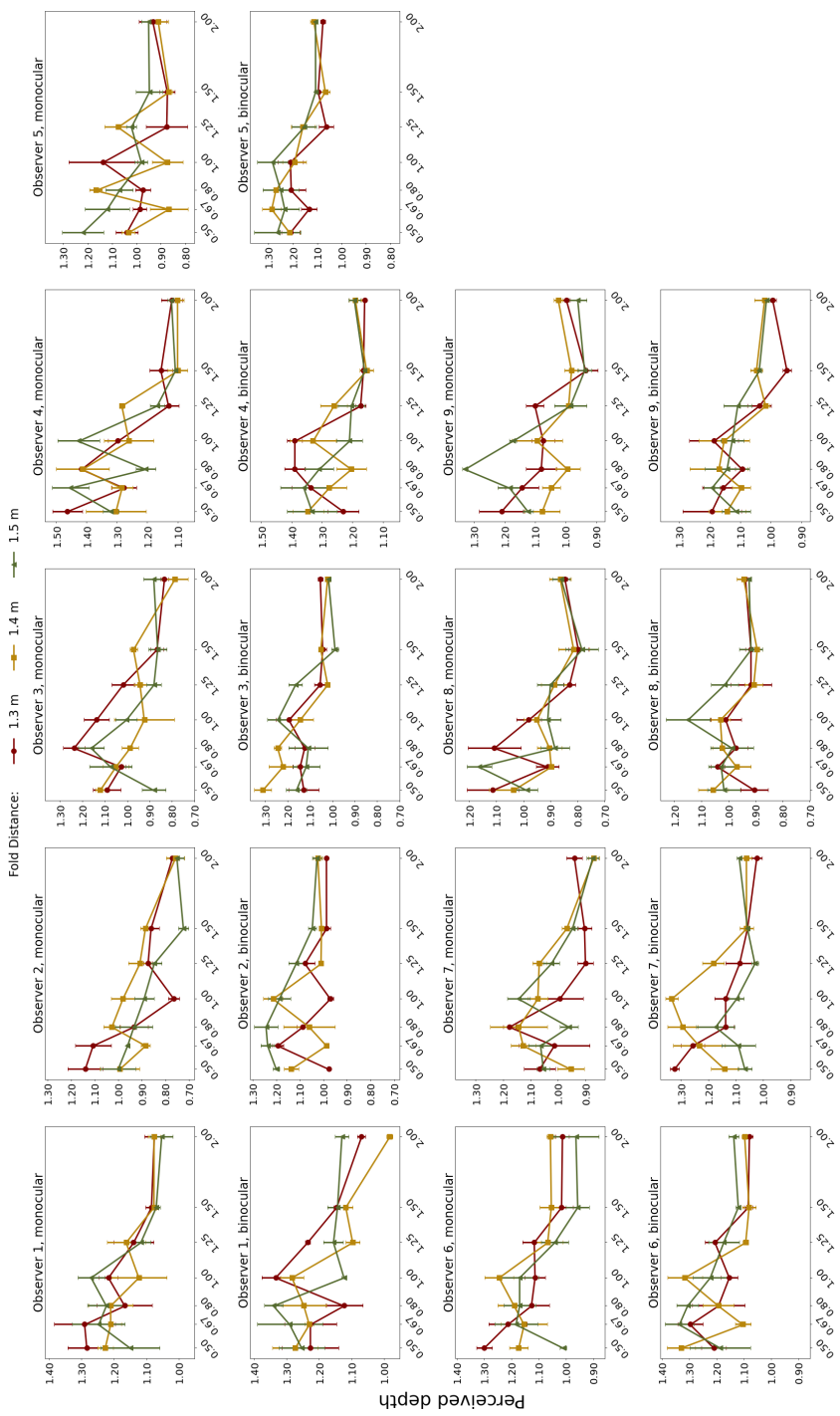


Figure B.5: Individual predictions of perceived depth as a function of Gain in Experiment 1. Error bars represent ±1 SEM.

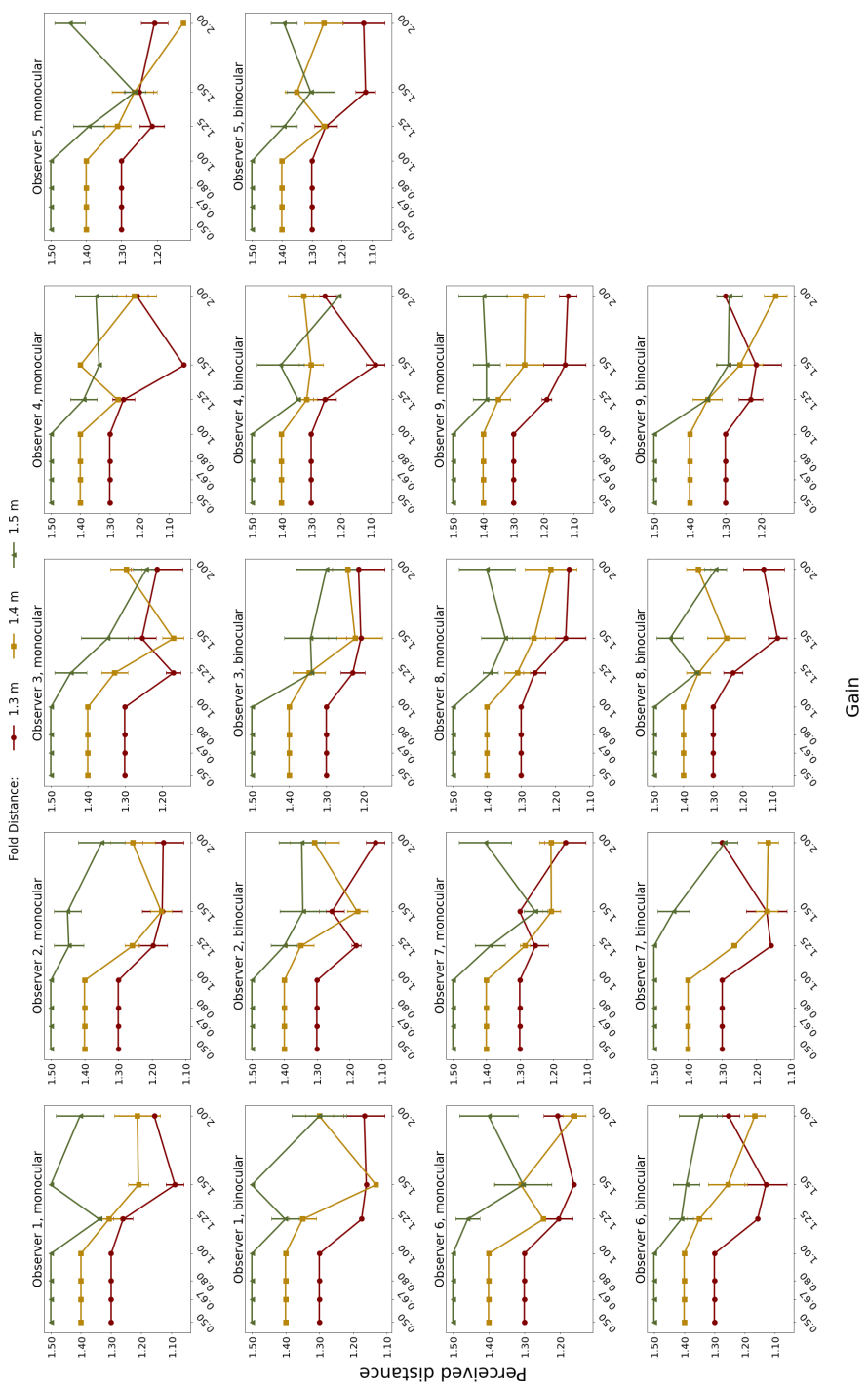


Figure B.6: Individual predictions of perceived fold distance as a function of Gain in Experiment 1. Error bars represent ±1 SEM.

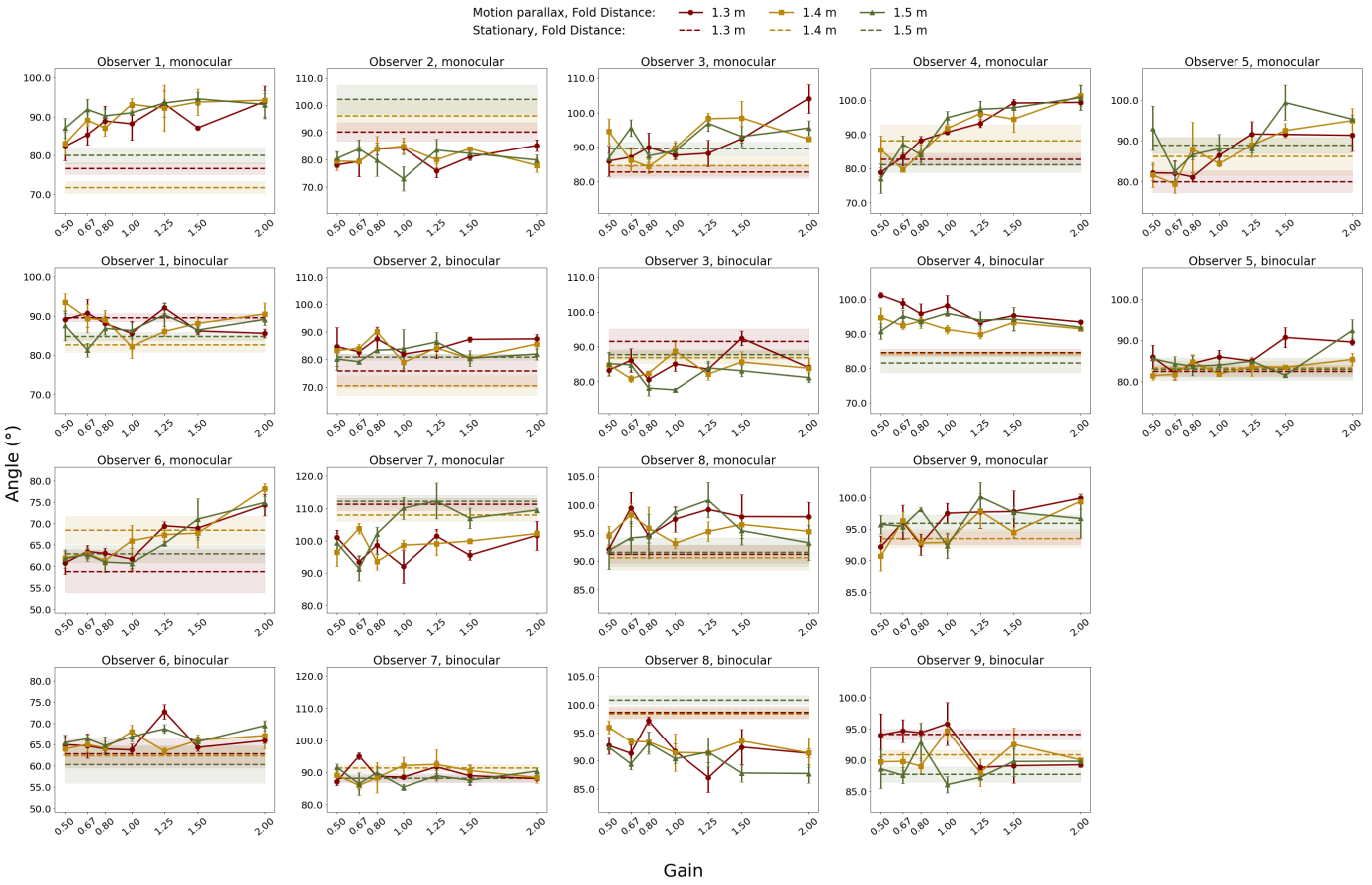
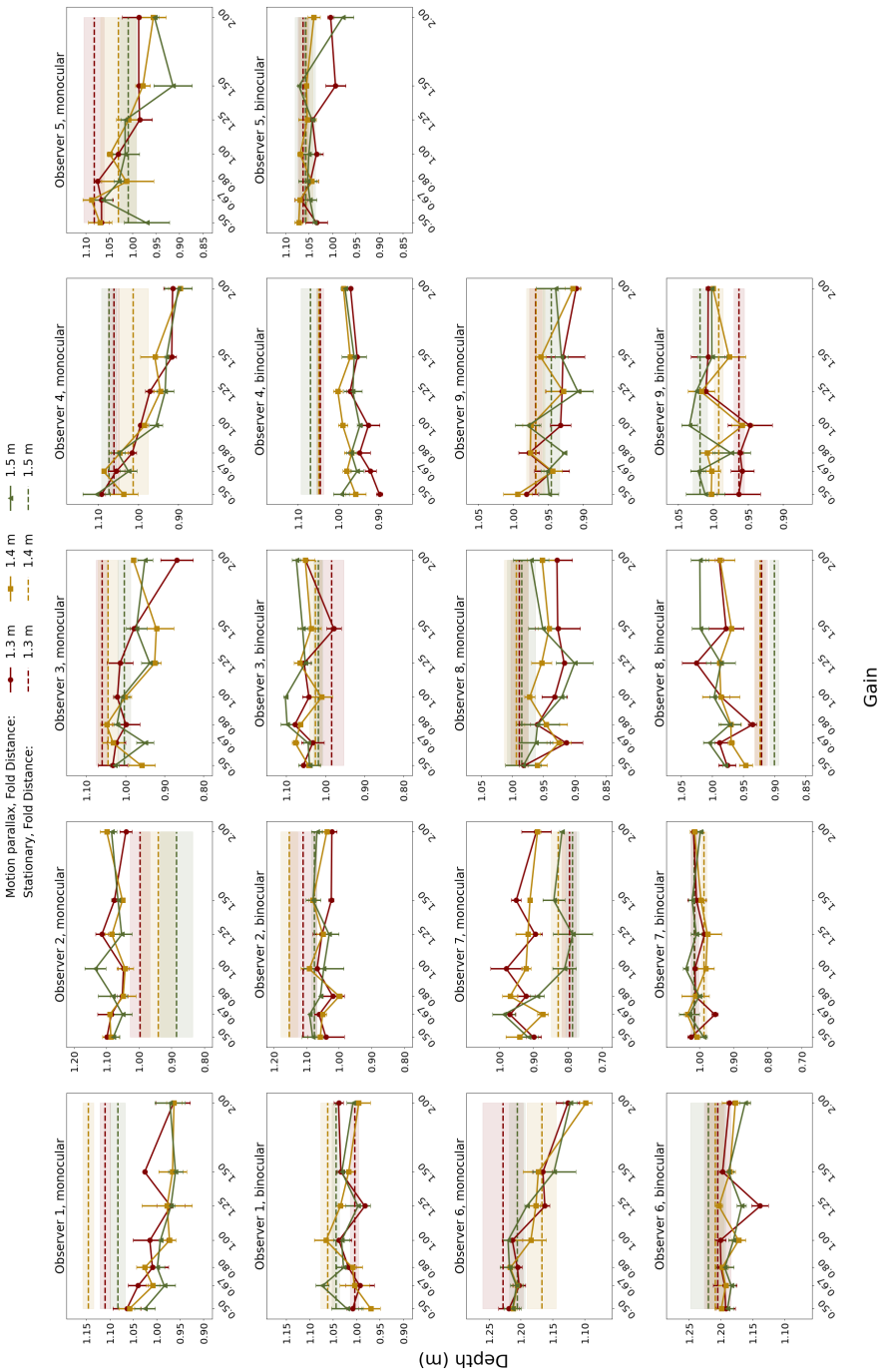
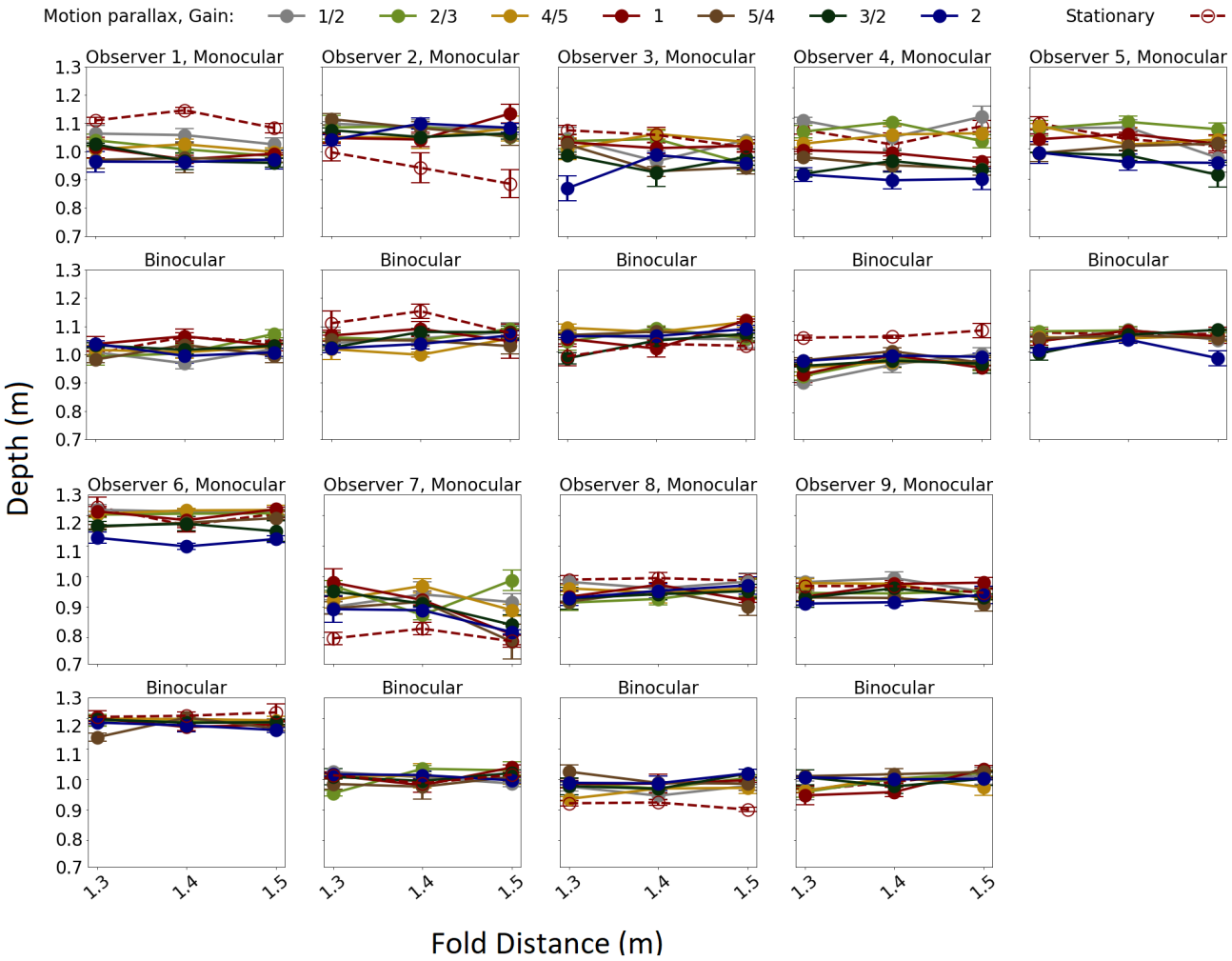


Figure B.7: Individual Angle settings as a function of Gain in Experiment 2. Stationary settings are plotted as a line across the x-axis for easy visualization. Error bars and regions around dashed lines represent ± 1 SEM.

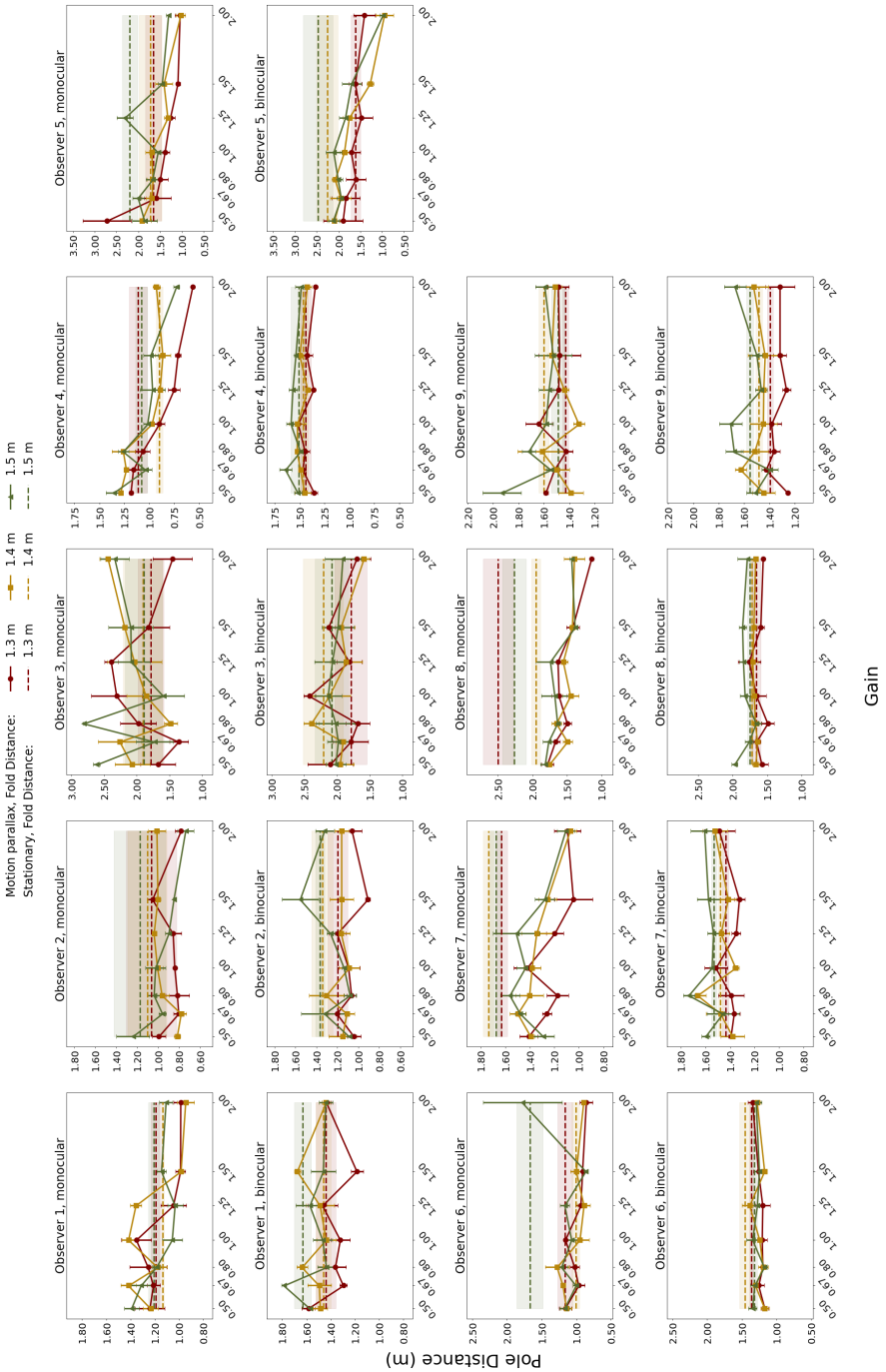


(a) Individual Depth settings as a function of Gain in Experiment 2. Stationary settings are plotted as a line across the x-axis for easy visualization.

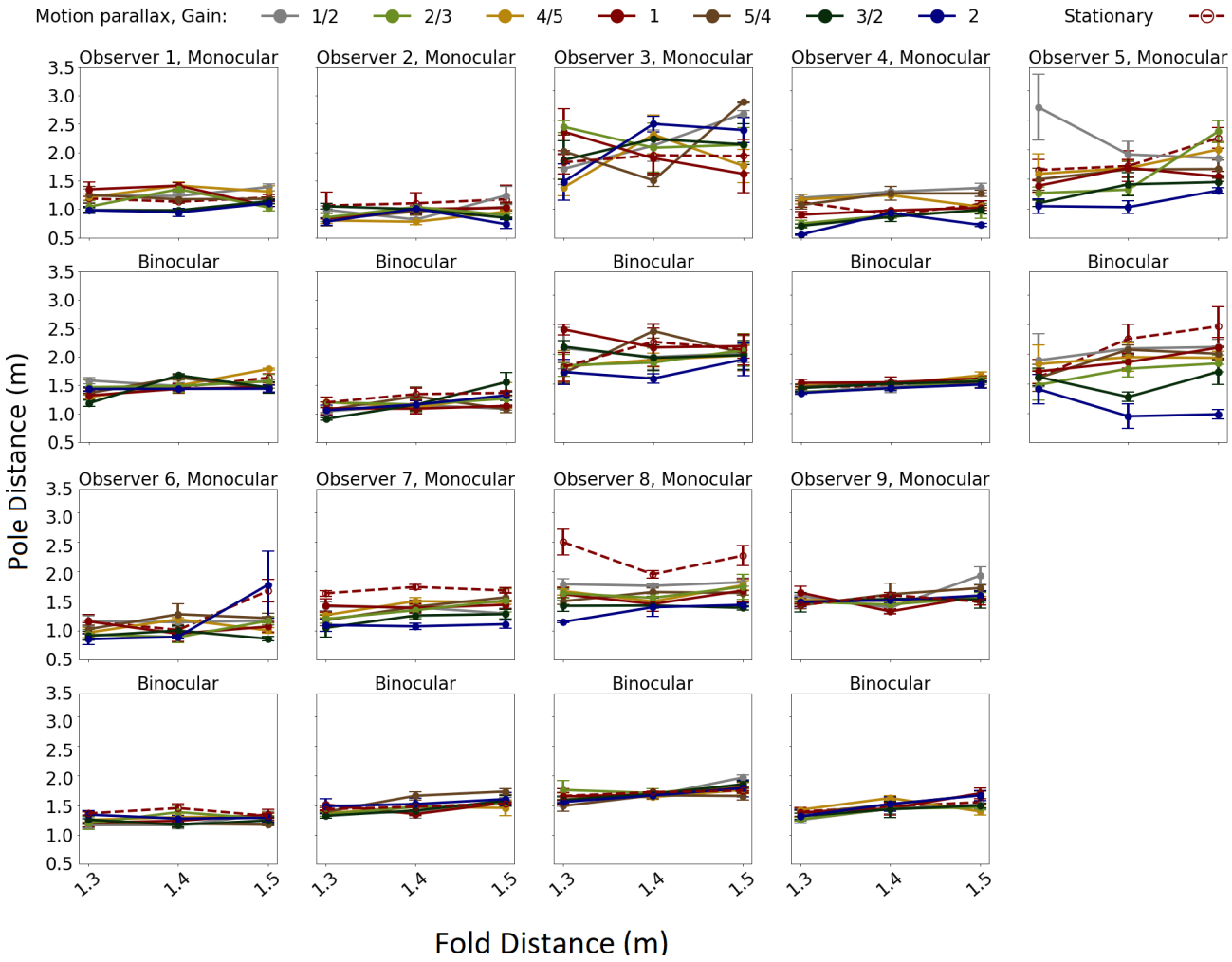


(b) Individual Depth settings as a function of Fold Distance in Experiment 2.

Figure B.8: Individual Depth settings in Experiment 2. Error bars and regions around dashed lines represent ± 1 SEM.



(a) Individual pole distance settings as a function of Gain in Experiment 2. Stationary settings are plotted as a line across the x-axis for easy visualization.



(b) Individual Pole Distance settings as a function of Fold Distance in Experiment 2.

Figure B.9: Individual Pole Distance settings in Experiment 2. Error bars and regions around dashed lines represent ± 1 SEM.

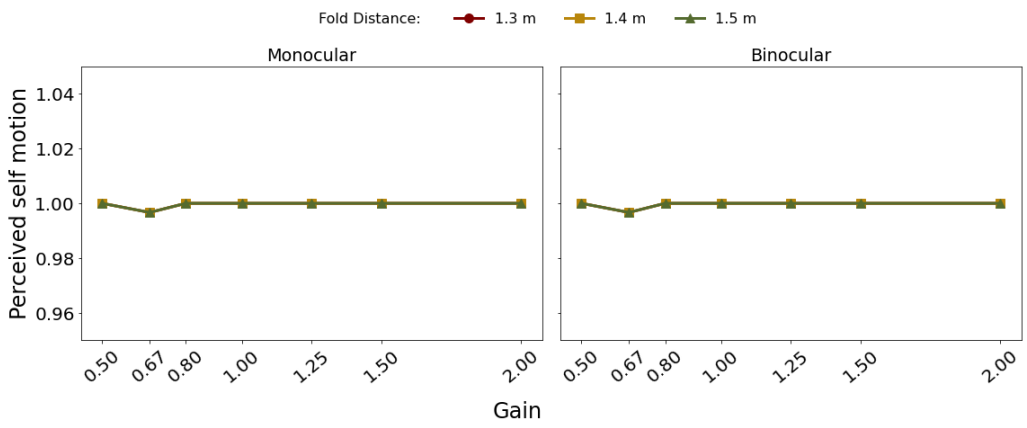


Figure B.10: Predictions of perceived self motion as a function of Gain averaged across all observers in Experiment 2. Error bars (too small to be seen in this plot) represent ± 1 SEM.

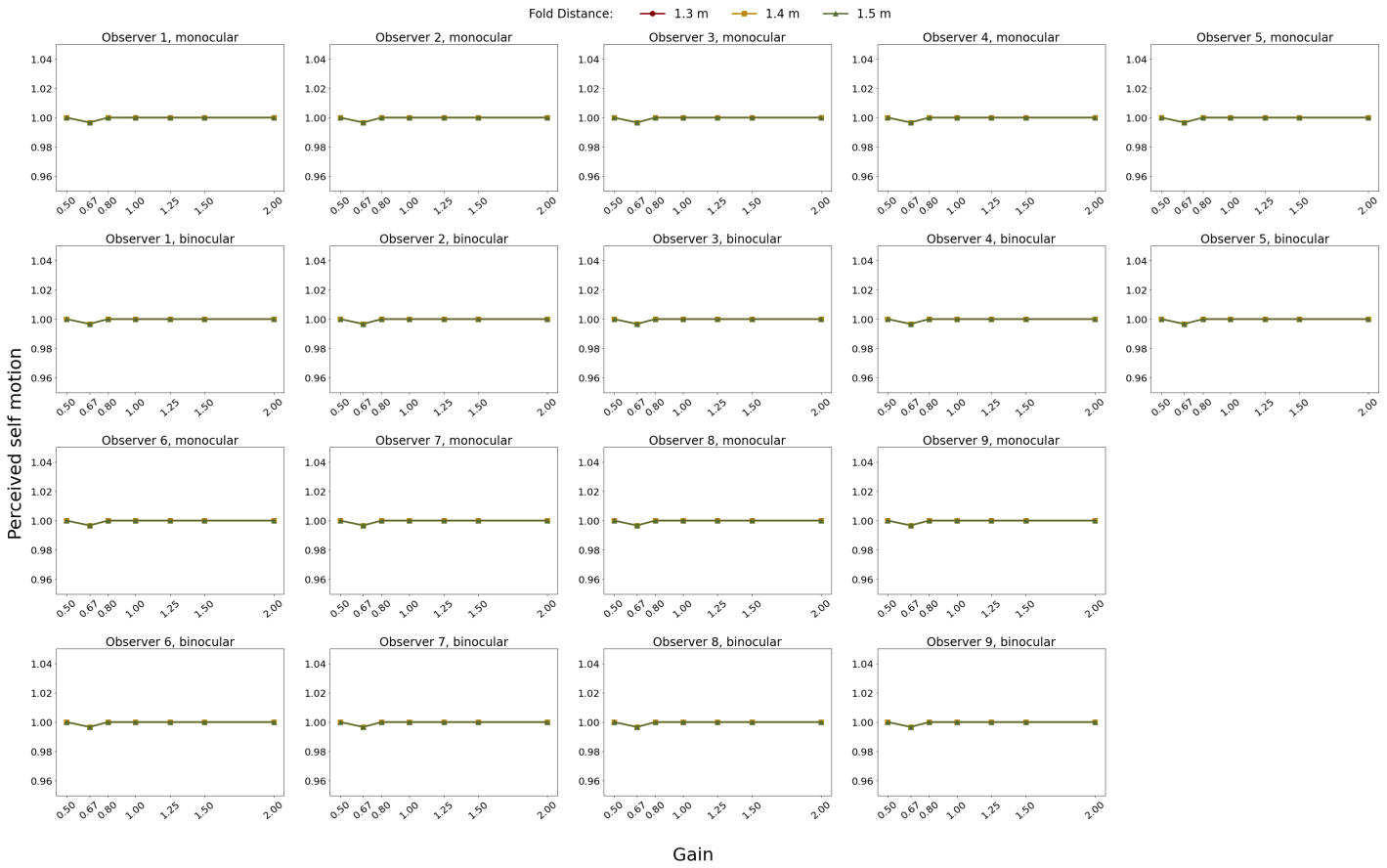


Figure B.11: Individual predictions of perceived self motion as a function of Gain in Experiment 2. Error bars (too small to be seen in this plot) represent ±1 SEM.

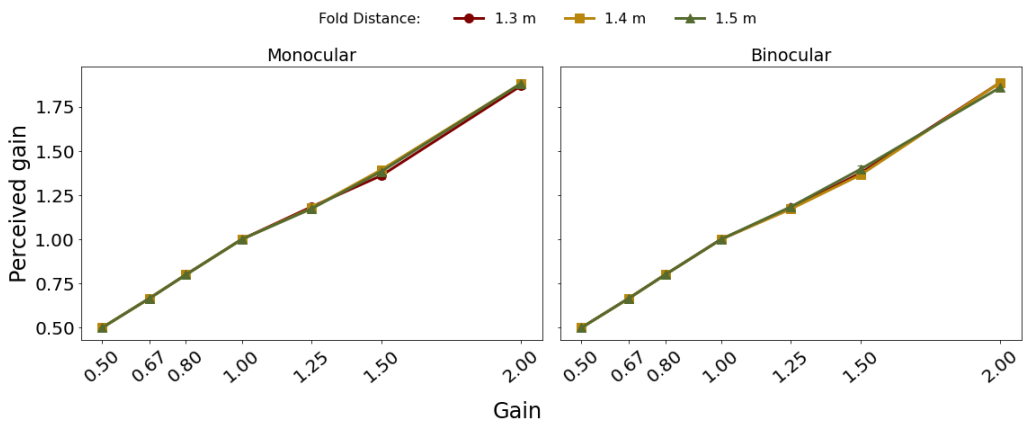


Figure B.12: Predictions of perceived gain as a function of Gain averaged across all observers in Experiment 2. Error bars (too small to be seen in this plot) represent ± 1 SEM.

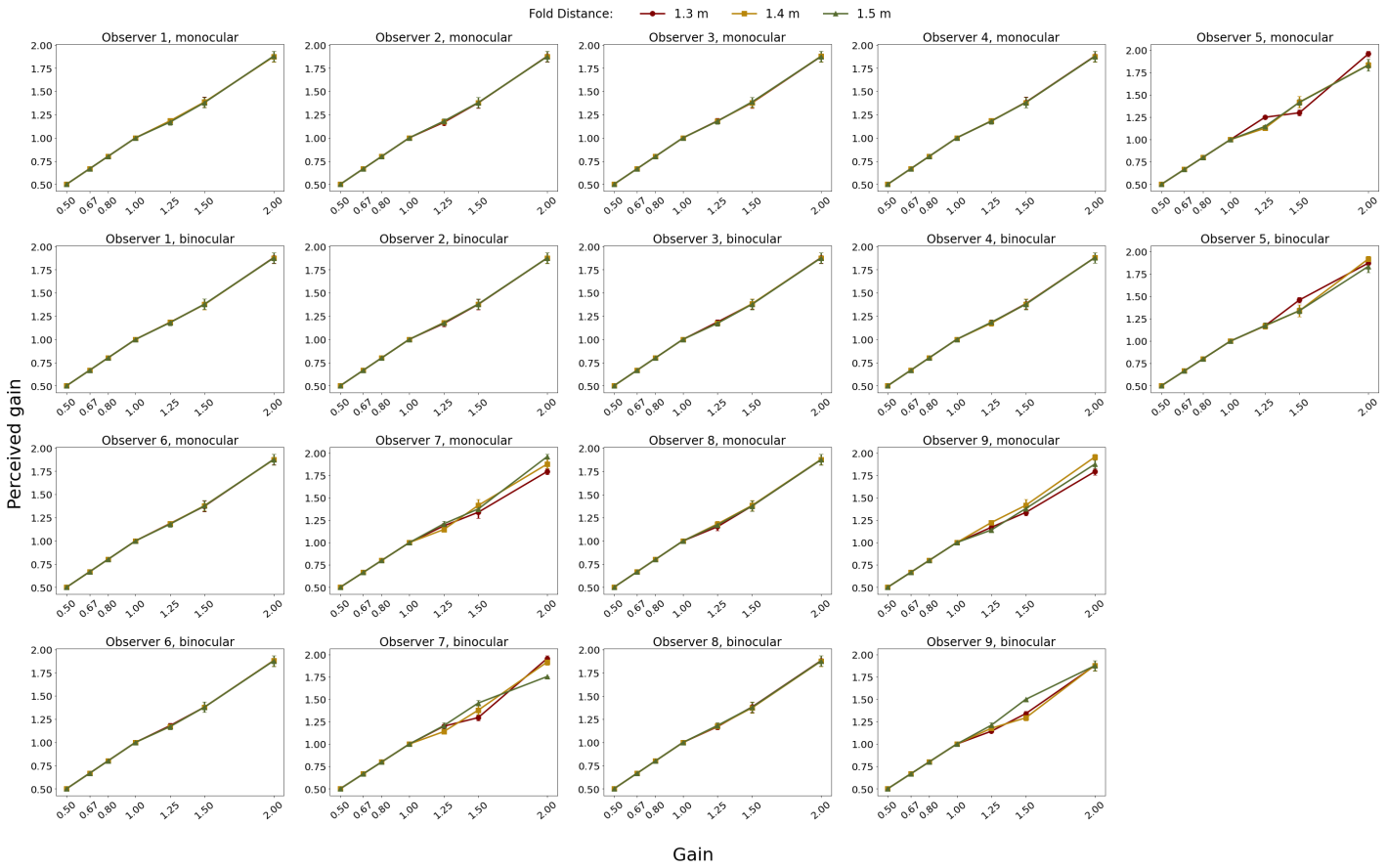


Figure B.13: Individual predictions of perceived gain as a function of Gain in Experiment 2. Error bars (some are too small to be seen) represent ±1 SEM.

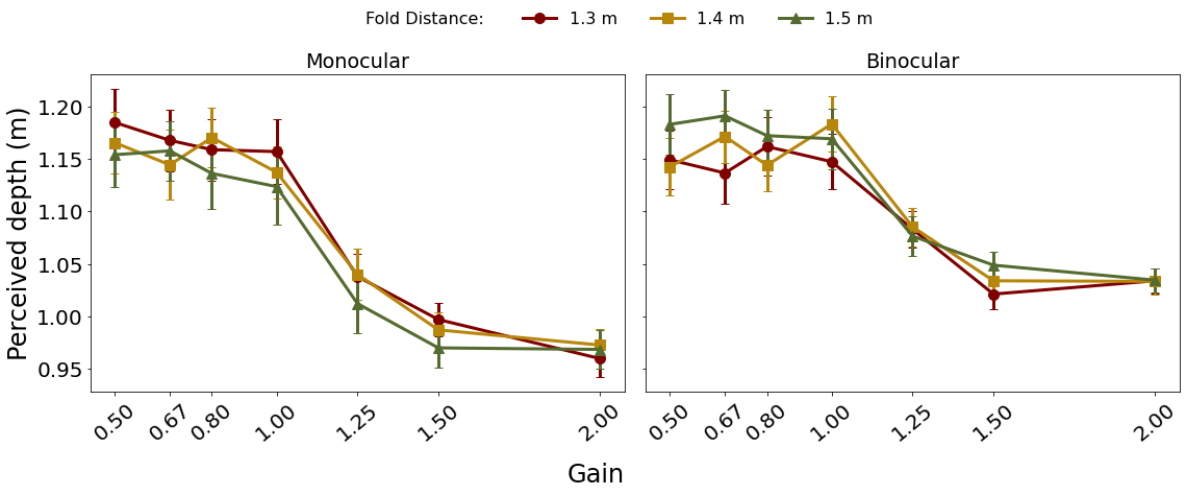


Figure B.14: Predictions of perceived depth as a function of Gain as a function of Gain averaged across all observers in Experiment 2. Error bars represent ± 1 SEM.

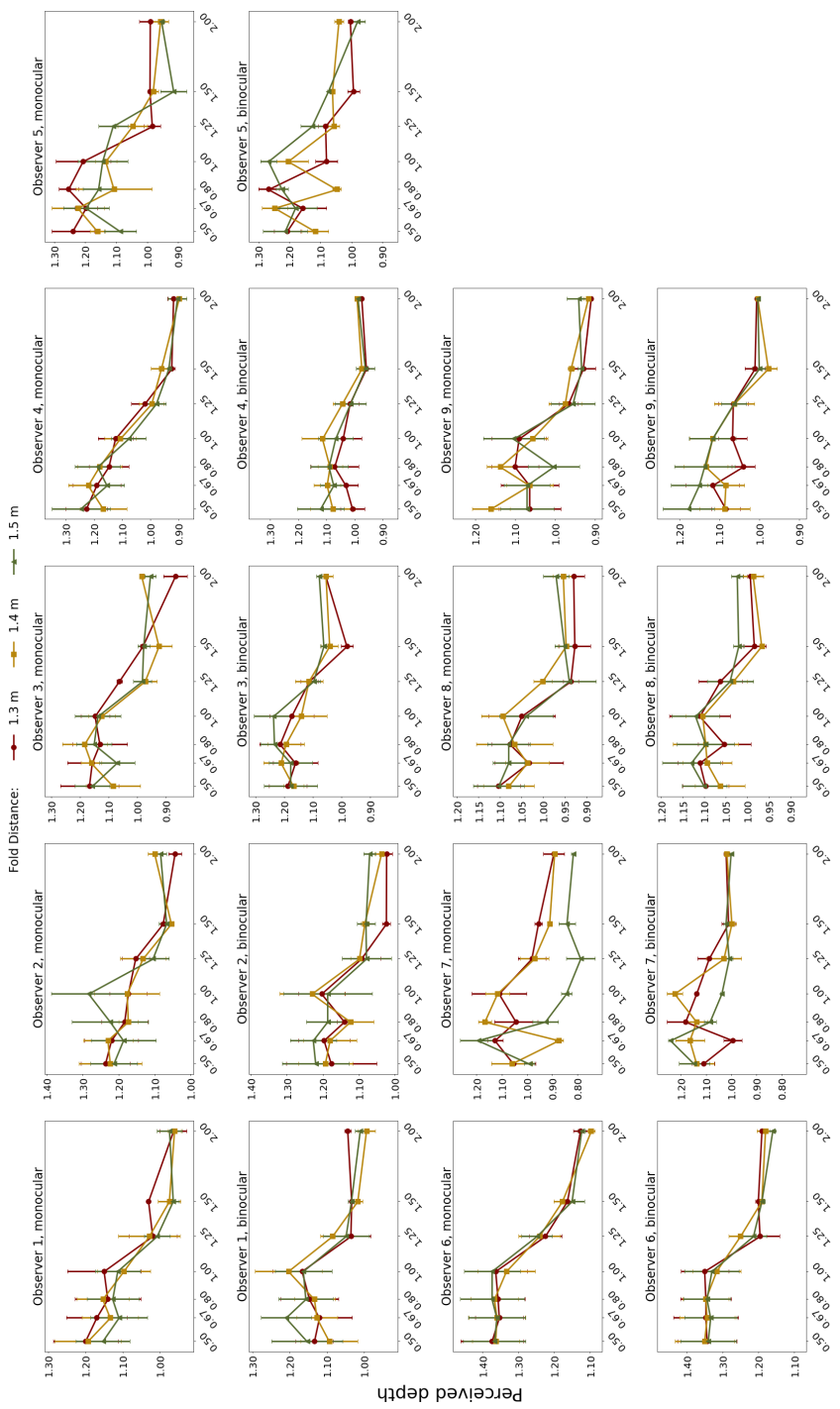


Figure B.15: Individual predictions of perceived depth as a function of Gain in Experiment 2. Error bars represent $\pm$ 1 SEM.

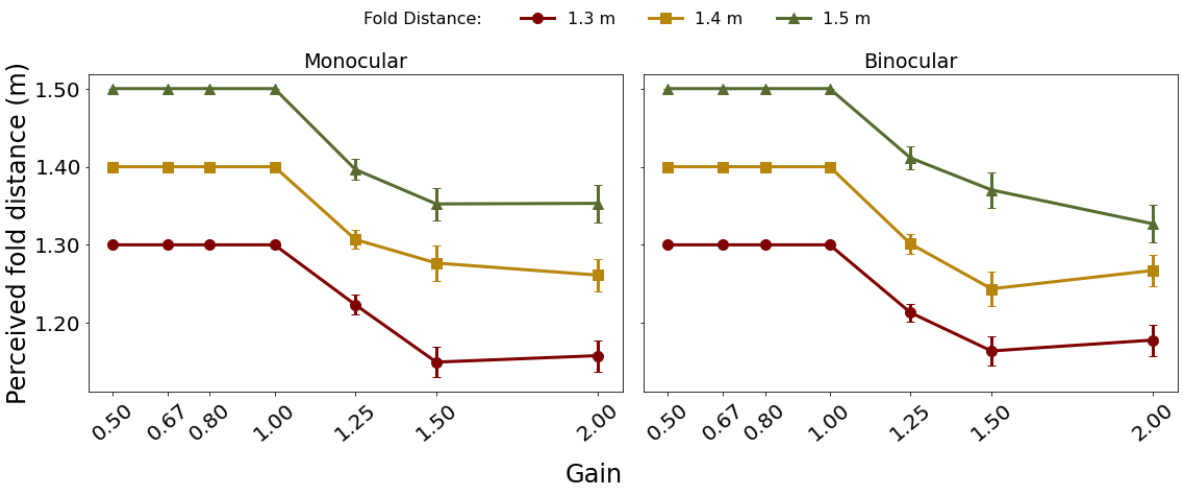


Figure B.16: Averaged predictions of perceived fold distance as a function of Gain averaged across all observers in Experiment 2. Error bars represent ±1 SEM.

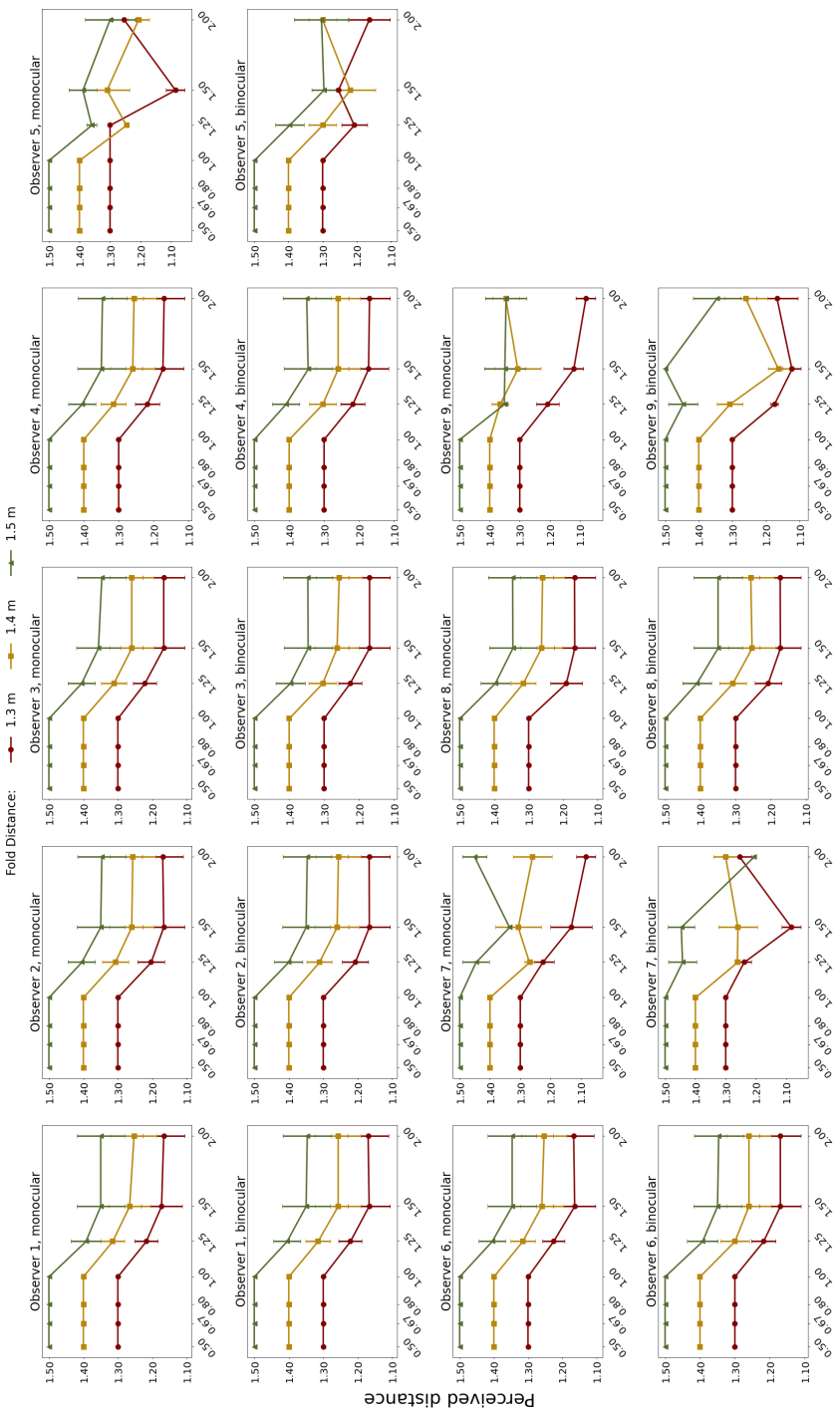


Figure B.17: Individual predictions of perceived fold distance as a function of Gain in Experiment 2.

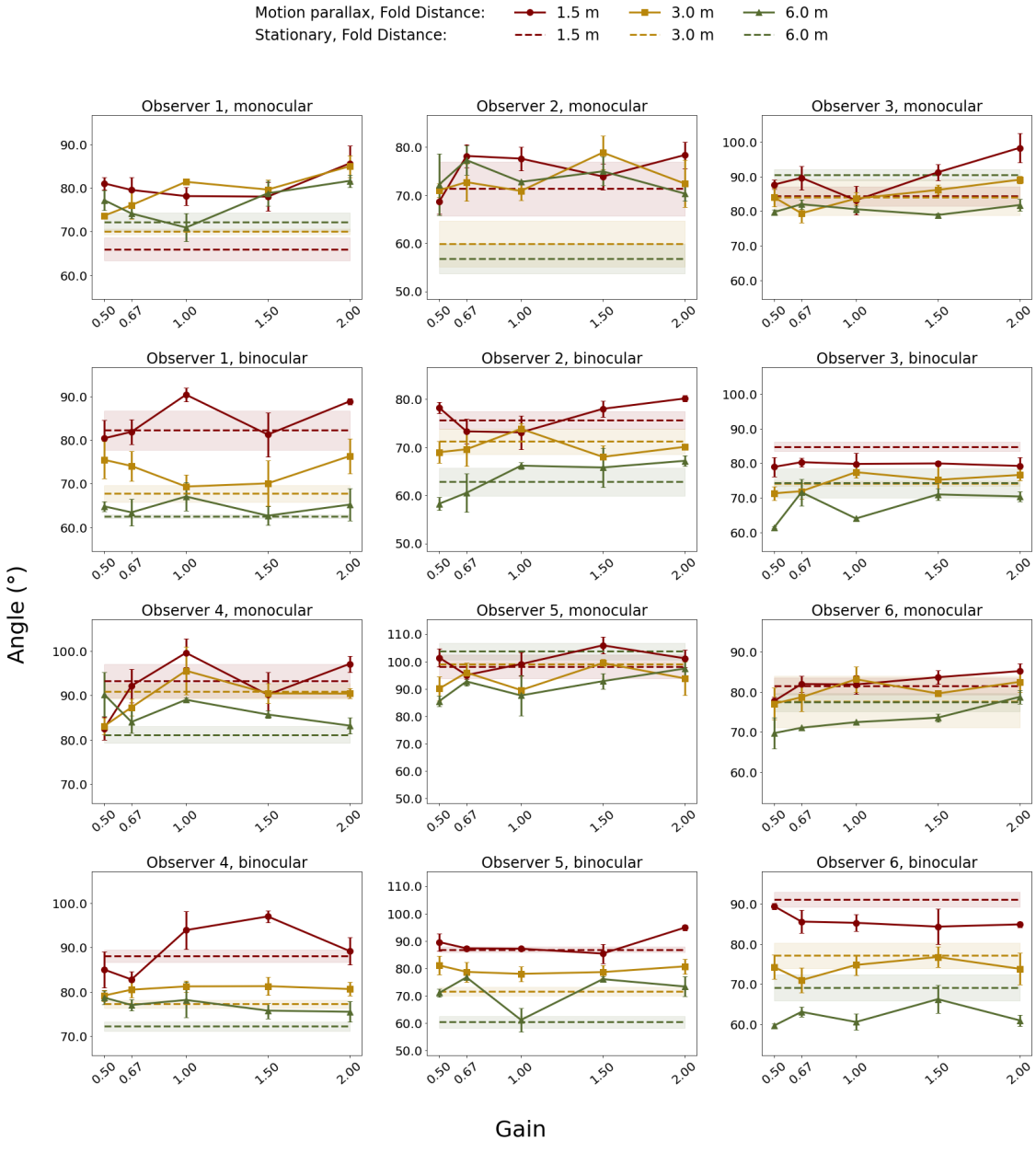
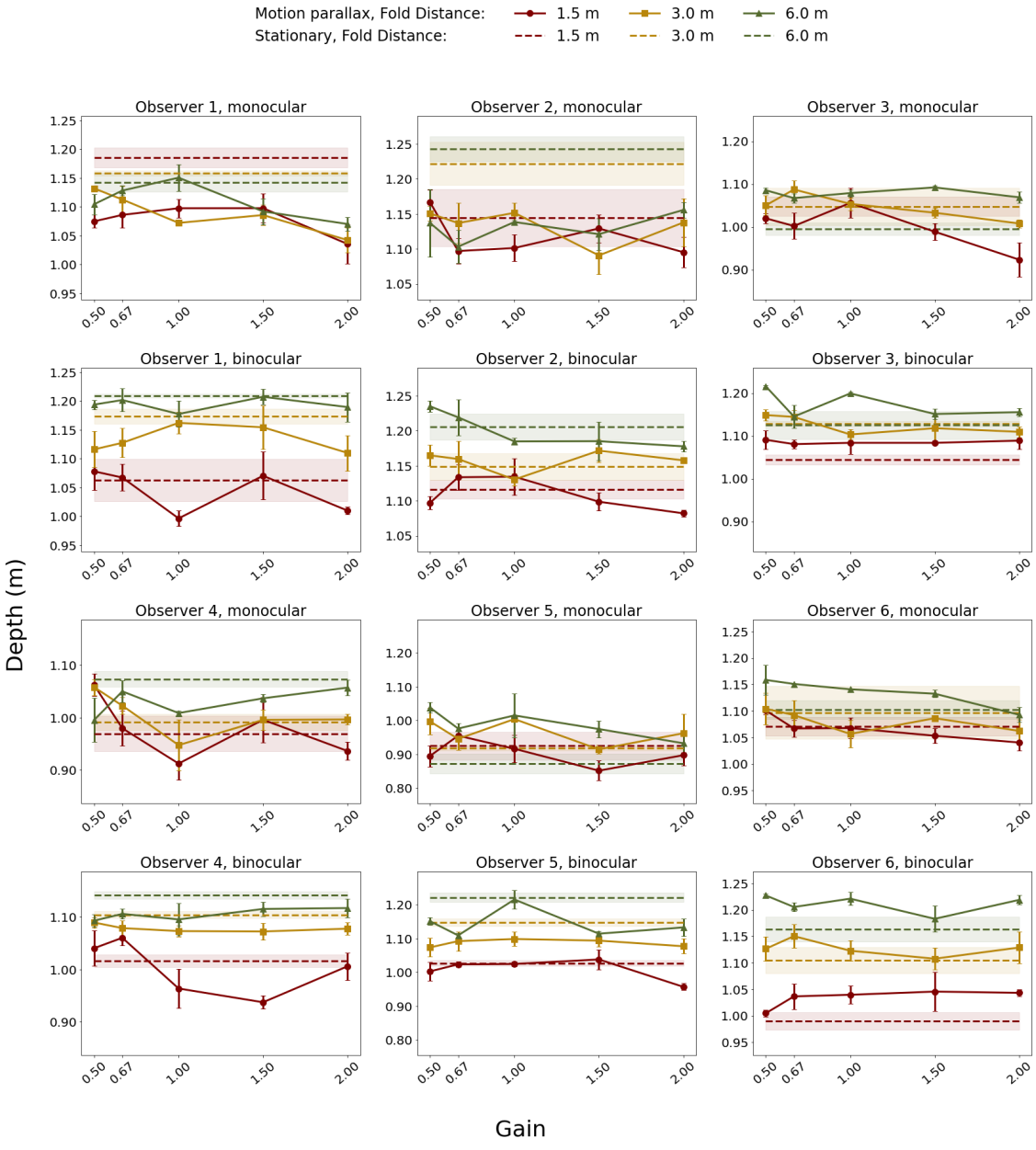
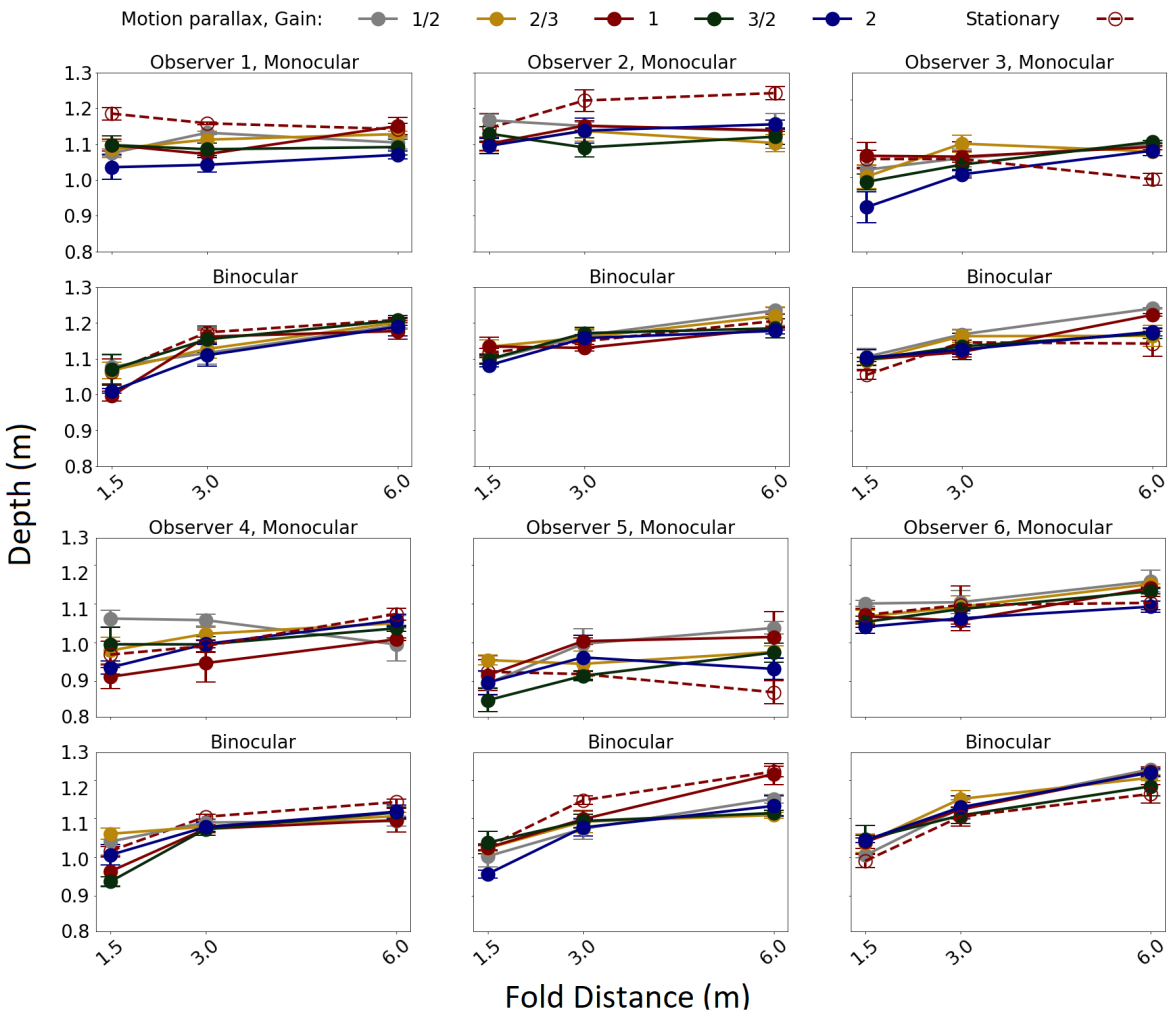


Figure B.18: Individual Angle settings as a function of Gain in Experiment 3. Stationary settings are plotted as a line across the x-axis for easy visualization. Error bars and regions around dashed lines represent ± 1 SEM.

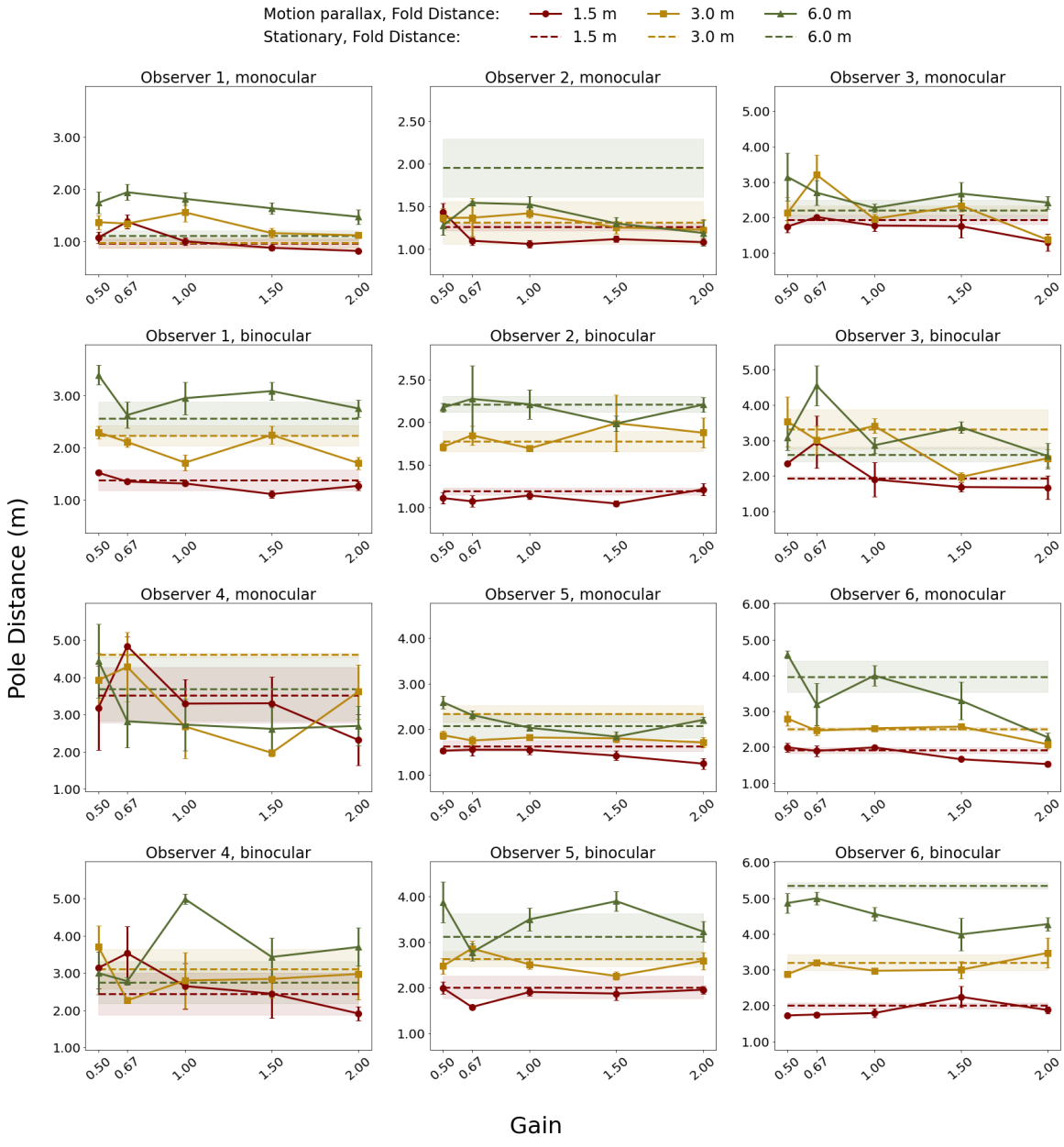


(a) Individual Depth settings as a function of Gain in Experiment 3. Stationary settings are plotted as a line across the x-axis for easy visualization.

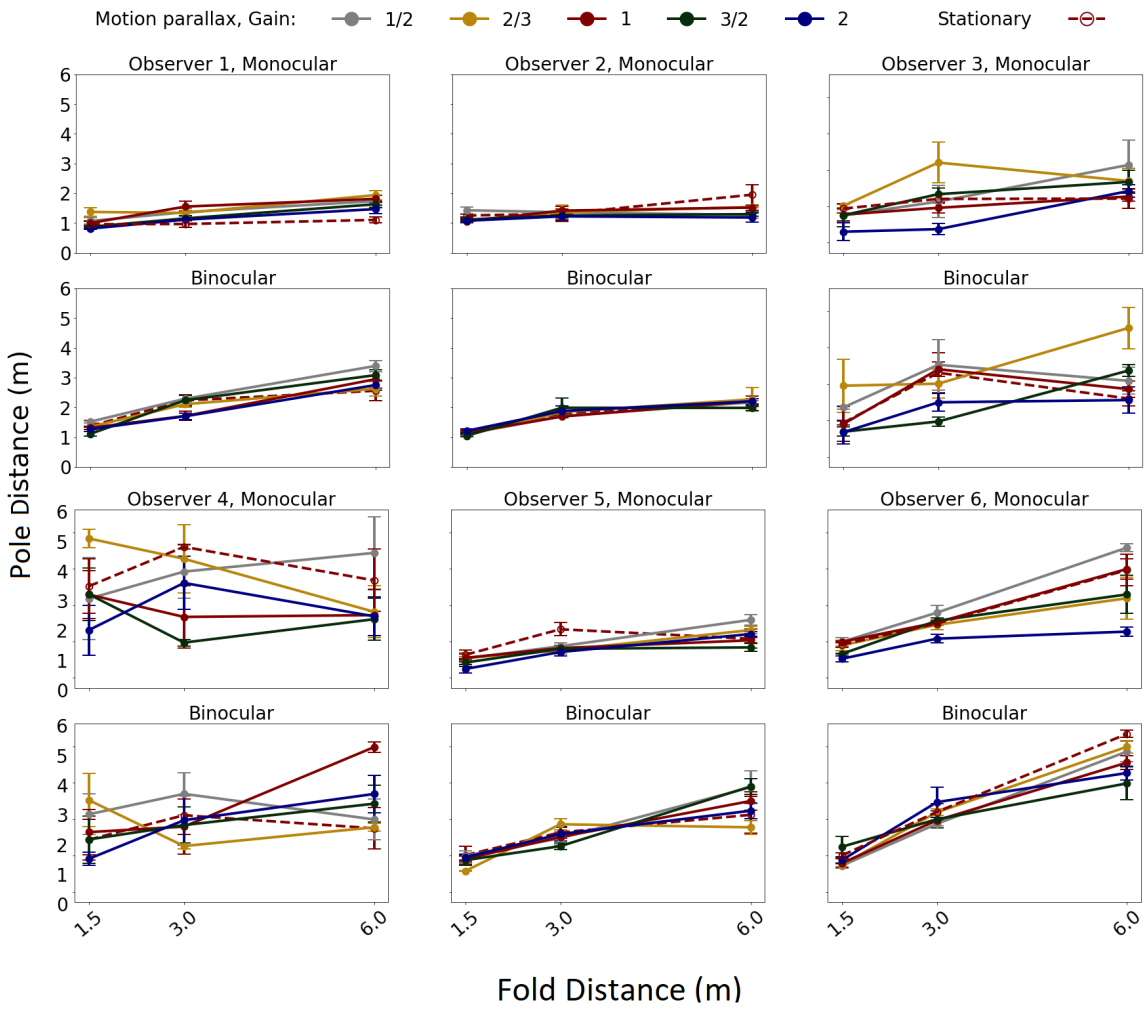


(b) Individual Depth settings as a function of Fold Distance in Experiment 3.

Figure B.19: Individual Depth settings in Experiment 3. Error bars and regions around dashed lines represent ± 1 SEM.



(a) Individual Pole Distance settings as a function of Gain in Experiment 3. Stationary settings are plotted as a line across the x-axis for easy visualization.



(b) Individual Pole Distance settings as a function of Fold Distance in Experiment 3.

Figure B.20: Individual pole distance settings in Experiment 3. Error bars and regions around dashed lines represent ± 1 SEM.

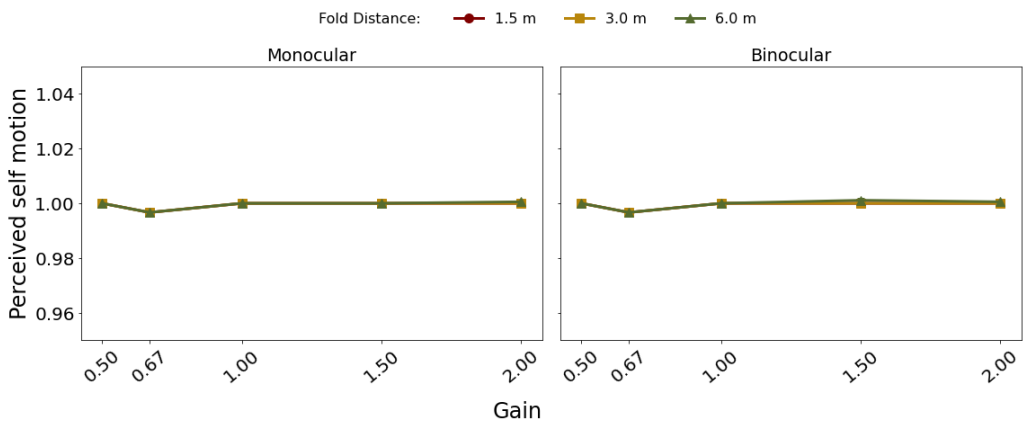


Figure B.21: Predictions of perceived self motion as a function of Gain averaged across all observers in Experiment 3. Error bars (too small to be seen in this plot) represent ± 1 SEM.

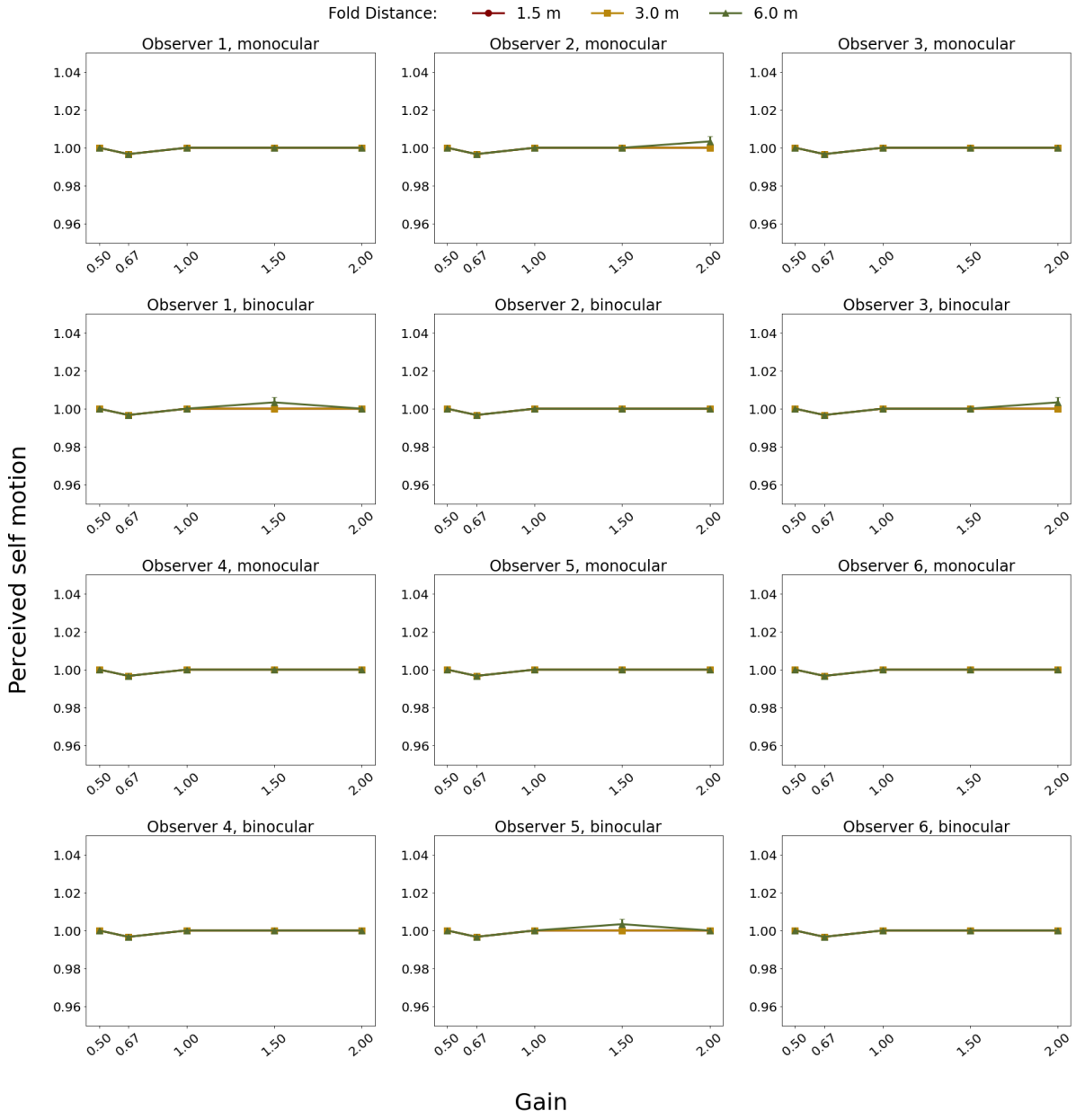


Figure B.22: Individual predictions of perceived self motion as a function of Gain in Experiment 3. Error bars (some are too small to be seen in this plot) represent ± 1 SEM.

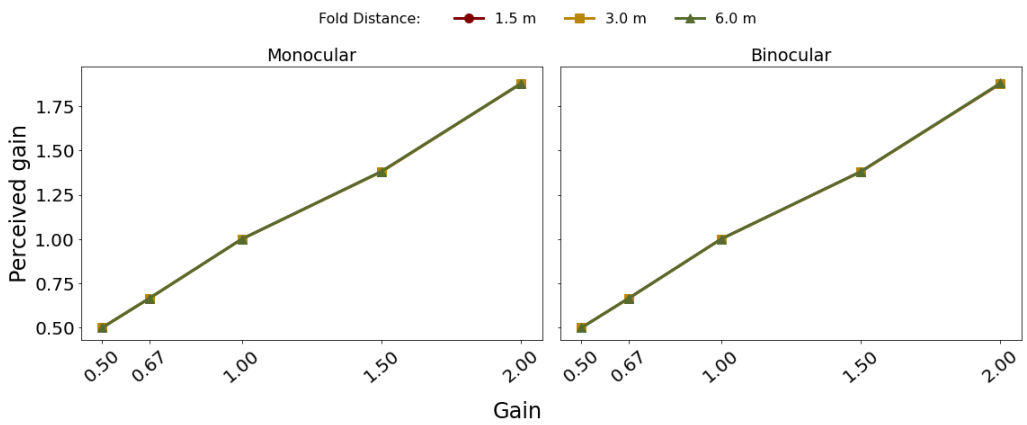


Figure B.23: Predictions of perceived gain as a function of Gain averaged across all observers in Experiment 3. Error bars (too small to be seen in this plot) represent ± 1 SEM.

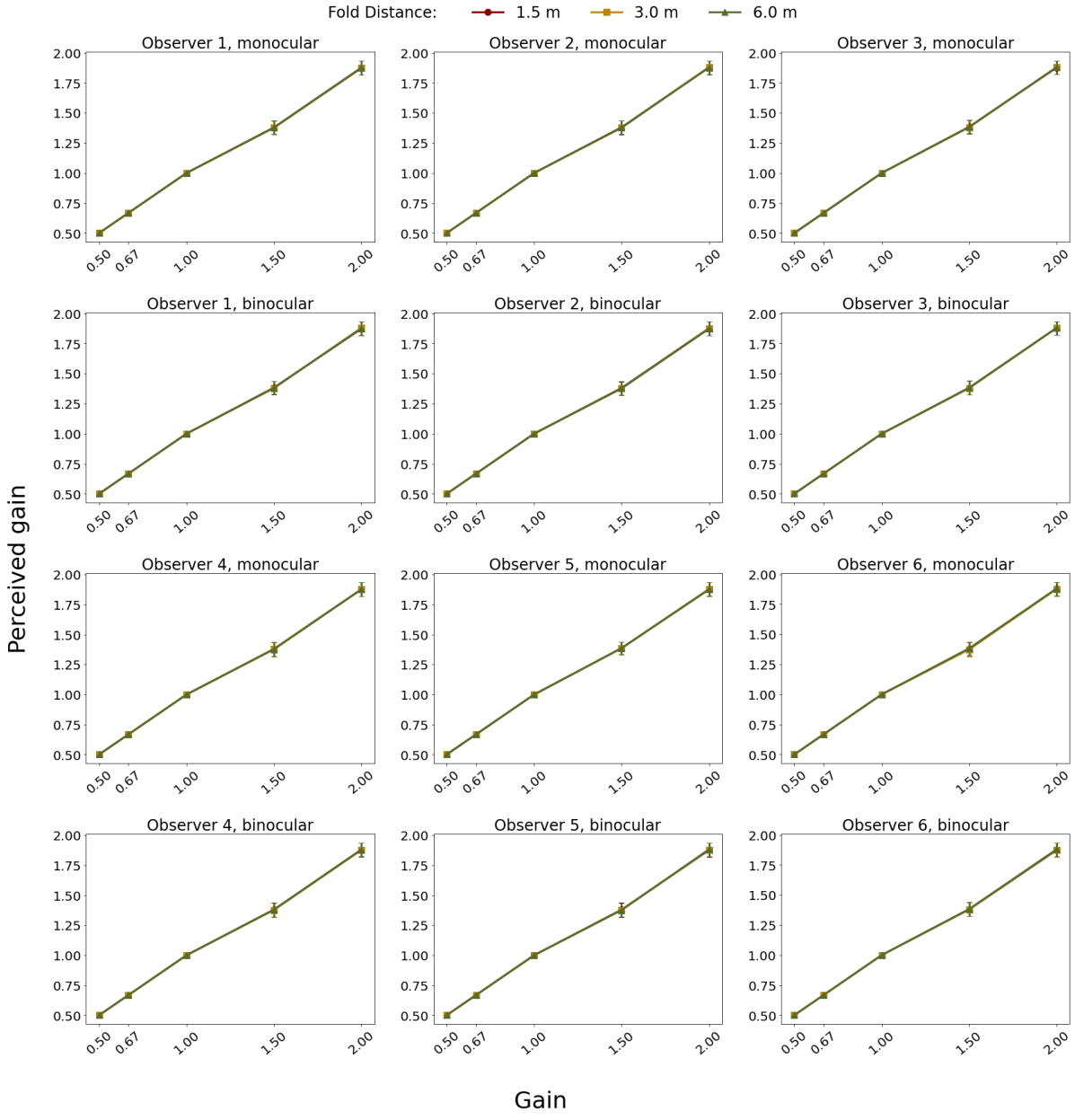


Figure B.24: Individual predictions of perceived gain as a function of Gain in Experiment 3. Error bars represent ± 1 SEM.

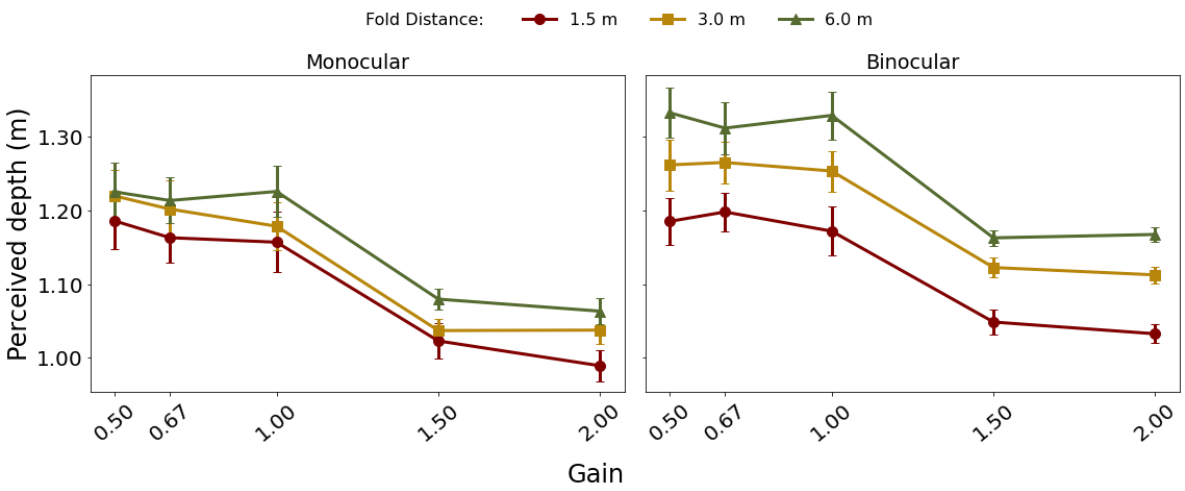


Figure B.25: Predictions of perceived depth as a function of Gain averaged across all observers in Experiment 3. Error bars represent ± 1 SEM.

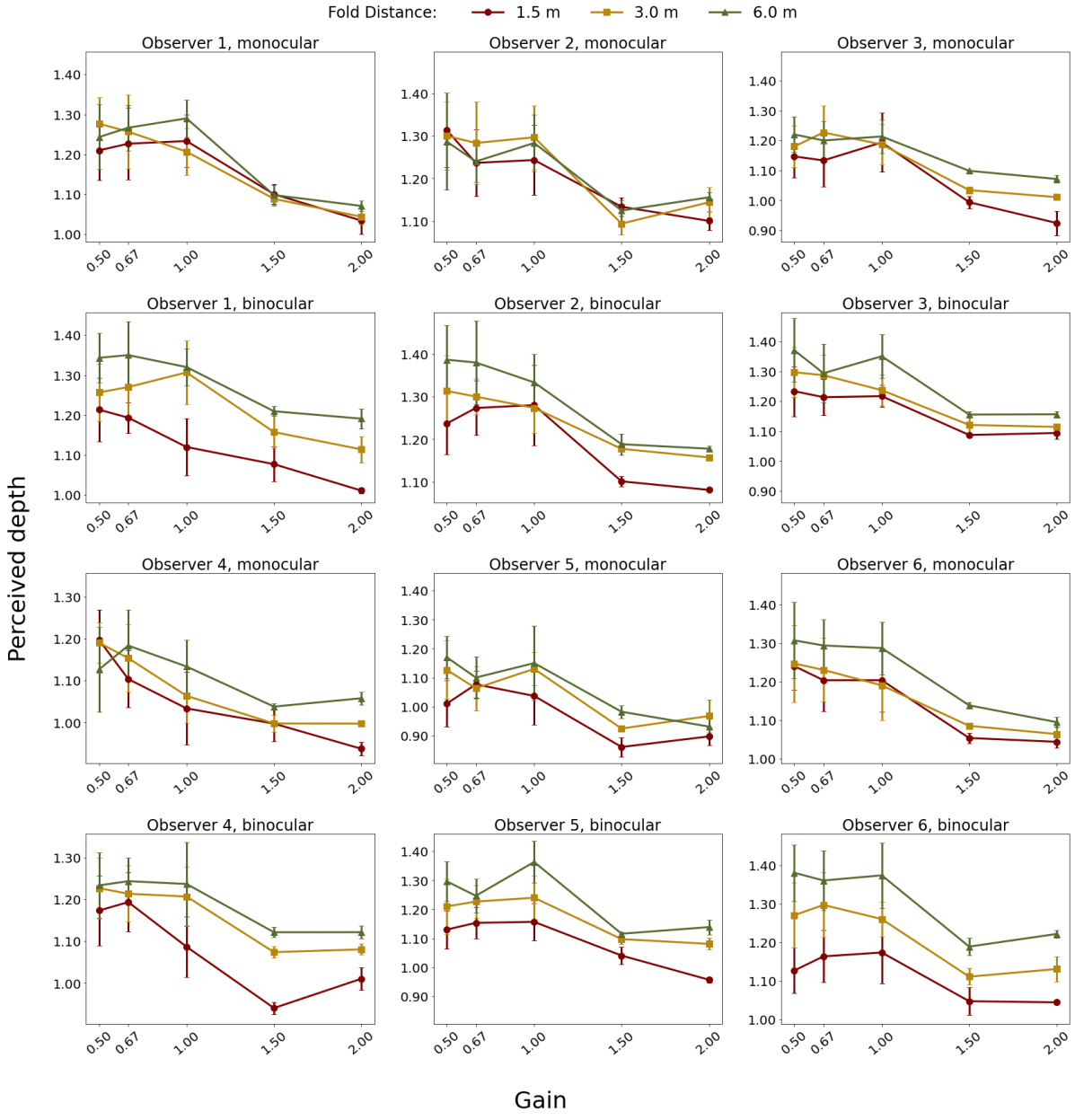


Figure B.26: Individual predictions of perceived depth as a function of Gain in Experiment 3. Error bars represent ± 1 SEM.

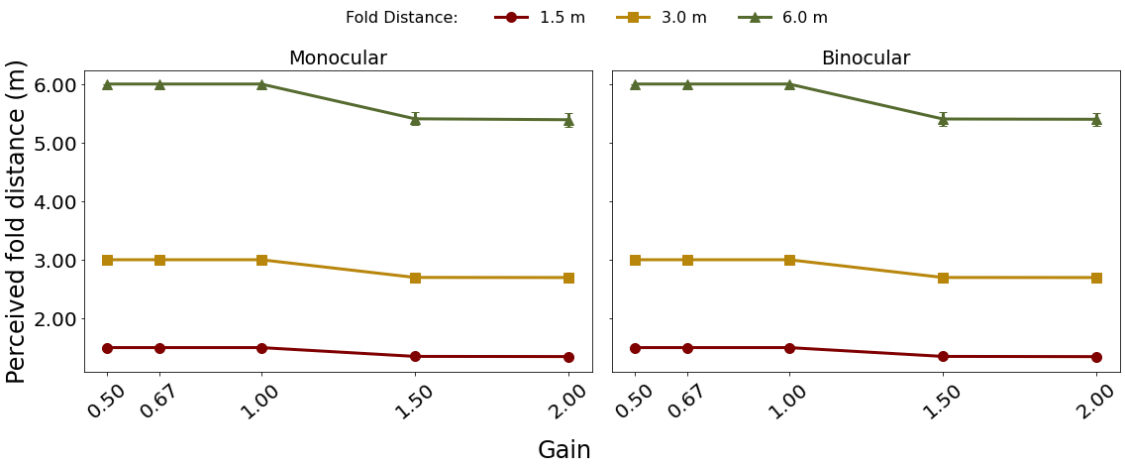


Figure B.27: Predictions of perceived fold distance as a function of Gain averaged across all observers in Experiment 3. Error bars (some are too small to be seen in this plot) represent ± 1 SEM.

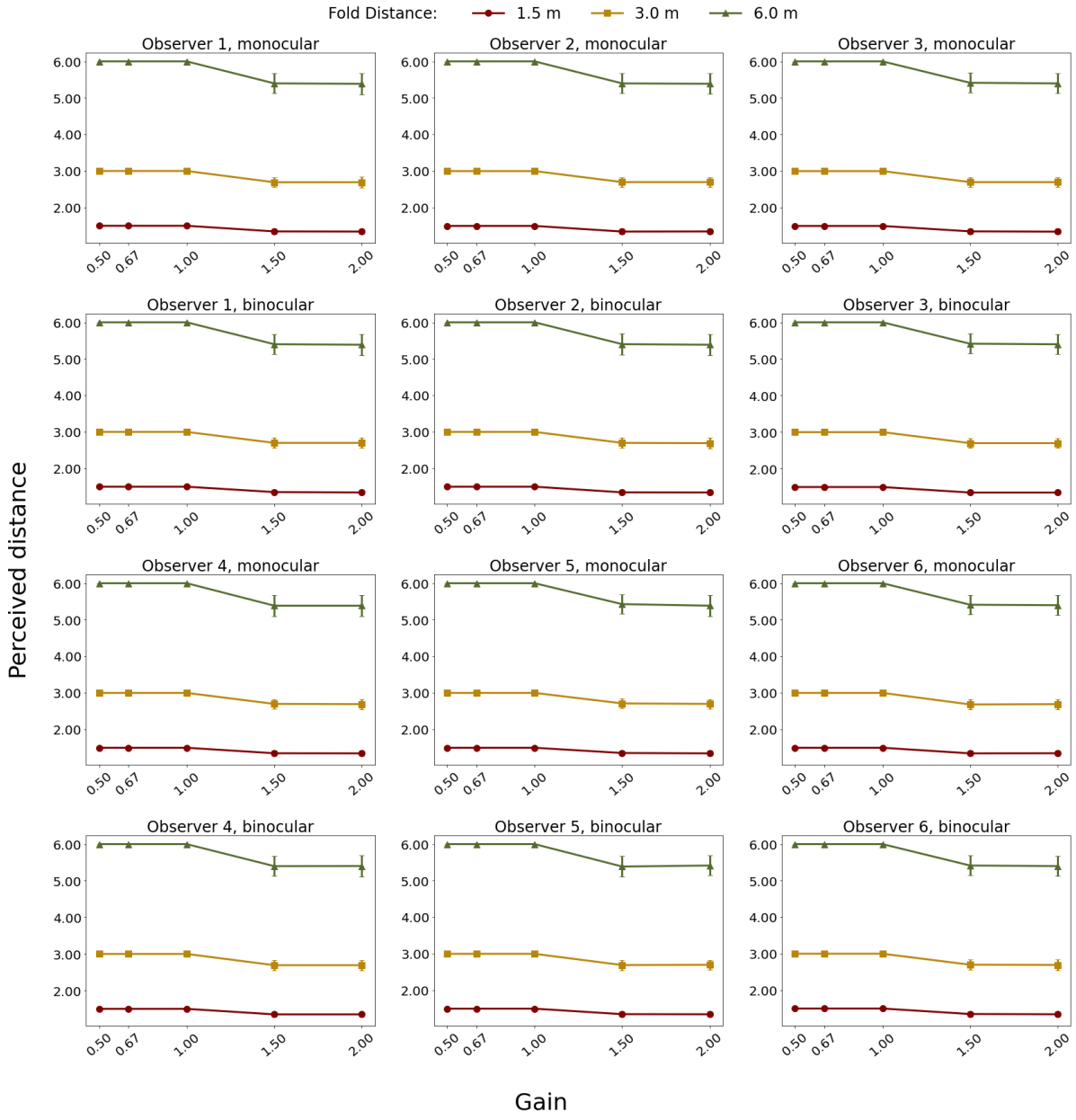


Figure B.28: Individual predictions of perceived fold distance as a function of Gain in Experiment 3. Error bars (too small to be seen in this plot) represent ± 1 SEM.