# INVESTIGATING ILLUSORY MOTION PARALLAX VIA VISUAL CUE DECOUPLING AND FORWARD MODELS IN VR

DANIAL KORDMODANLOU

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

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING & COMPUTER SCIENCE
YORK UNIVERSITY
TORONTO, ONTARIO

JUNE 2026

© DANIAL KORDMODANLOU, 2026

# Abstract

Under natural viewing, stereopsis and motion parallax provide consistent information about three-dimensional structure, but stereoscopic displays can decouple them: observers moving before a stereoscopic display that lacks viewpoint-contingent updates perceive illusory parallactic motion. This thesis investigates that illusion in head-mounted virtual reality and tests whether it arises from predictive mechanisms that use stereoscopic depth to generate expectations about motion parallax. Across five two-interval forced-choice experiments, the illusion required binocular disparity, and its perceived direction aligned with negative parallax, consistent with a signed prediction error rather than an unsigned cue conflict. Its magnitude was modulated by scene context, downward viewing angle, and observer distance. These findings support a forward-model account: binocular disparity specifies depth and predicts head-coupled motion parallax, and the absent expected parallax yields a prediction error experienced as scene motion. The results inform how the visual system integrates conflicting depth cues and the design of stereoscopic displays.

# Dedication

I dedicate this thesis to the memory of the victims of Ukraine International Airlines Flight 752; to the brave souls who stood with courage and sacrifice in the Women, Life, Freedom movement; and to the victims of the January mass killings in Iran and the Minab school attack.

Their loss is beyond words, their courage lives on, and their struggle will not fade. They are remembered, and they will not be forgotten.

# Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Nikolaus F. Troje, for his guidance, mentorship, and support throughout this research, and for the intellectual freedom he gave me to explore the questions that became this thesis. His insight, patience, and willingness to engage deeply with the ideas at every stage shaped both the work and my development as a researcher.

I am also grateful to my committee member, Professor Robert S. Allison, for his thoughtful feedback and for the perspective he brought to the project from his work on stereoscopic perception and virtual reality.

I would like to thank the members of the BioMotion Lab at the Centre for Vision Research, York University, for the discussions, technical help, and collegial atmosphere that made the day-to-day work of this thesis possible.

I am thankful to the participants who volunteered their time for the studies reported here, recruited through the York University Undergraduate Research Participant Pool, partnering research labs, and the broader university community. All procedures were approved by York University's Human Participants Review Sub-Committee under REB certificate 2025-199.

Finally, I am deeply thankful to my parents and my sister for their constant support, patience, and belief in me, and to my friends whose encouragement and companionship made this journey easier and more meaningful.

# Table of Contents

# List of Tables

# List of Figures

# Abbreviations

**2AFC** Two-Alternative Forced-Choice

**2IFC** Two-Interval Forced-Choice

**ANOVA** Analysis of Variance

**BML-TUX** Experiment framework used for trial control, blocking, and data logging

**CI** Confidence Interval

**GEE** Generalized Estimating Equation

**HMD** Head-Mounted Display

**IPD** Inter-Pupillary Distance

**LMM** Linear Mixed-Effects Model

**PC** Personal Computer

**PSE** Point of Subjective Equality

**SD** Standard Deviation

**SEM** Standard Error of the Mean

**VR** Virtual Reality

# Chapter 1

# Introduction

Depth perception in human vision arises from the integration of multiple visual cues, notably stereoscopic depth (binocular disparity) and motion parallax. Binocular disparity refers to the differences in images perceived by each eye due to their horizontal separation, allowing the brain to reconstruct depth information [Rogers and Graham, 1979, 1982]. Motion parallax, on the other hand, involves the relative motion of objects across the visual field as an observer moves, providing essential depth cues through the analysis of velocity gradients in the optical flow field [Nawrot and Stroyan, 2009]. These two systems share remarkable computational similarities, as both rely on processing information from spatially separated viewpoints; either simultaneous views from the two eyes or sequential views obtained during head movement, to extract three-dimensional structure [Allison et al., 2003]. For comprehensive reviews of depth perception and stereoscopic vision, see [Howard and Rogers, 2012].

In natural environments, these cues typically complement each other, contributing to a coherent three-dimensional perception of depth and spatial layout. However, virtual reality (VR) systems can artificially manipulate these cues, presenting unique opportunities to study their individual contributions to depth perception. For instance, we can present stereoscopic images without the corresponding motion parallax, producing a cue conflict and perceptual illusions.

When stereoscopic depth is presented without corresponding motion parallax during head movement, observers can perceive motion in physically stationary display content. Closely related effects have previously been described as *illusory concomitant motion* or *stereoillusory motion*, with apparent motion varying systematically with binocular disparity, perceived depth, head-movement amplitude, and viewing geometry [Gogel and Tietz, 1974, Peterson, 1986, Shimono et al., 2002, Ono and Ujike, 2005]. In stereoscopic display contexts, this cue conflict has also been described as *illusory motion parallax* [Benzeroual and Allison, 2013]. In sparse or simple scenes, this percept can nevertheless constitute a locally valid geometric interpretation of the available cues, because the visual system can reconcile the input by attributing the mismatch to scene motion rather than to a violation of the inferred depth structure [Gogel, 1990, Shimono et al., 2007]. Investigating this illusion can therefore provide insights into predictive mechanisms in the visual system and into how those mechanisms depend on integrated sensory information and scene context.

## 1.1 Theoretical Framework: The Forward Model

One possible explanation for this illusion is based on internal forward models, which are predictive mechanisms that estimate the expected sensory consequences of self-generated movements [Miall and Wolpert, 1996]. When the predicted and measured sensory signals do not match, the resulting prediction error can give rise to a perceptual discrepancy that may be experienced as motion. This account assumes that the visual system operates under strong priors of scene rigidity and environmental stability, such that discrepancies between predicted and measured sensory input are more likely to be attributed to object or scene motion than to violations of these assumptions. These priors are generally well justified in natural viewing, but can lead to non-veridical percepts when depth cues are artificially decoupled, as in virtual reality.

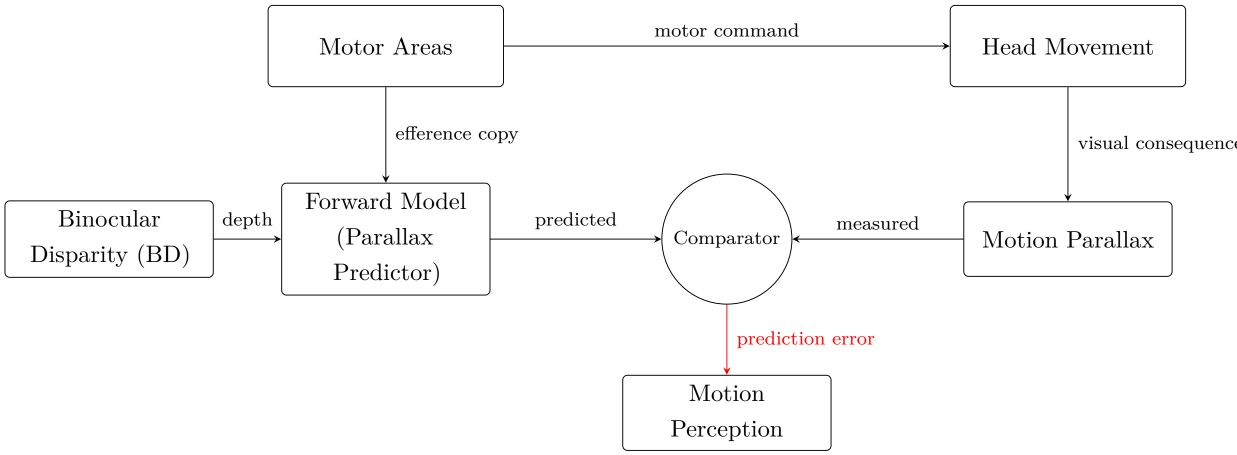


Figure 1.1: Conceptual forward-model framework for illusory parallactic motion. A motor command for head movement generates an efference copy that drives a forward model predicting the expected motion parallax. Predicted and measured parallax are compared, and a mismatch can be experienced as illusory motion. Binocular disparity, shown here as an input to the forward model, constrains the expected magnitude and direction of parallax; zero disparity specifies a flat surface and predicts zero parallax, whereas monocular viewing provides insufficient information to constrain three-dimensional structure and biases interpretation toward picture mode [Troje, 2019].

Figure 1.1 provides a conceptual illustration of how such a mechanism may contribute to the illusion. The distinction between *picture mode* and *presence mode*, introduced by Troje [2019], helps describe when these predictive processes are likely to be engaged. In picture mode, the visual system treats a scene as a flat representation, either because depth cues are weak or absent, or because the stimulus context discourages interpretation of sensorimotor contingencies as reflecting a stable three-dimensional environment. In presence mode, however, sensorimotor contingencies are interpreted as cues to an immersive three-dimensional environment, and the visual system expects the sensory input to update consistently with head motion. Thus, cue availability and perceptual interpretation jointly determine whether predictive mechanisms associated with presence mode are engaged.

When the perceptual system operates in a context that encourages presence mode, forward-model predictions about motion parallax become relevant. If the visual input does not follow these predicted changes, the resulting mismatch can be interpreted as movement within the scene, potentially giving rise to illusory parallactic motion. Together, these elements provide a principled account of how predictive mechanisms, informed by depth cues and prior assumptions about the environment, can give rise to the perceptual effects investigated in this thesis when those cues are decoupled.

Building on this framework, the present work investigates not only when illusory parallactic motion arises, but also how its strength is modulated by the broader visual context in which it is embedded. While prior research has primarily treated such illusions as arising from local conflicts between binocular disparity and motion parallax, this thesis demonstrates that **scene context systematically modulates the perceived strength of illusory motion**.

Specifically, we show that the presence or absence of surrounding environmental structure influences how strongly predictive mechanisms are engaged, suggesting that expectations about motion parallax are not determined solely by local depth cues, but are shaped by the global organization of the scene. By combining controlled virtual reality experiments with a forward-model interpretation, this work provides new evidence that predictive processes in vision are context-dependent, and that perceptual inference operates at the level of entire scene representations rather than isolated cues.

# Chapter 2

# Background

## 2.1 Depth from Binocular Disparity and Motion Parallax

Binocular disparity and motion parallax are two fundamental cues to depth perception. The visual system normally uses them interactively to infer 3D structure [Rogers and Graham, 1982, Bradshaw and Rogers, 1996]. Binocular disparity arises from the difference between the two eyes' views, while motion parallax comes from an observer's head movements, which produce relative image motion. Early work demonstrated that motion parallax can serve as an independent depth cue much like stereopsis. For instance, Rogers and Graham [1982] showed that human depth perception from motion parallax closely parallels that from binocular disparity. Subsequent research revealed that these cues do not operate in isolation but interact in the brain's depth computations. Bradshaw and Rogers [1996] found evidence of cross-cue interactions: adapting to depth defined by one cue could raise thresholds for detecting depth with the other, and combining disparity and parallax yielded higher sensitivity than either cue alone. These results suggest that the combination of binocular disparity and motion parallax is not a simple linear addition of independent estimates, but instead reflects interactions between cues at multiple stages of depth processing [Bradshaw et al., 1998]. However, evidence from cross-adaptation and cue-combination paradigms does not unequivocally determine the level or mechanism of integration, leaving open whether these interactions arise from shared representations, decision-level processes, or predictive mechanisms operating across cues.

From a normative perspective, such cross-cue benefits and interactions can be formalized as probabilistic cue integration: cues are weighted by their relative reliability to minimize uncertainty in the final estimate. Empirically, humans often approximate statistically optimal (maximum-likelihood/Bayesian) integration in multisensory estimation tasks [Ernst and Banks, 2002], and broader Bayesian accounts argue that uncertainty is incorporated into neural computation [Knill and Pouget, 2004]. Neurophysiological and psychophysical studies likewise support early integration; for example, adding stereo information can resolve directional ambiguities in a rotating 3D motion display [Nawrot and Blake, 1989]. Nawrot and Blake [1989] reported that when structure-from-motion displays were supplemented with disparity cues, the previously bistable or ambiguous depth perception becomes stable, implying a joint processing of motion and stereoscopic signals. In summary, depth from disparity and from motion parallax share many properties and are processed in a coordinated manner, and comprehensive treatments further detail how these cues normally work in harmony to yield a consistent sense of depth. Debates about whether explicit binocular matching is always necessary for stereopsis further illustrate that the effective contribution of disparity can depend on viewing conditions and competing interpretations [Gillam, 1995]. Disrupting the balance or correspondence between them can lead to perceptual anomalies, as discussed next.

## 2.2 Depth Cue Conflicts in Virtual Reality

In an ideal VR setup with head tracking and stereoscopic displays, both disparity and motion parallax cues are updated in synchrony to preserve a stable 3D perception. However, when these cues are not properly synchronized, the user can experience perceptual conflicts [Allison et al., 2003, Benzeroual and Allison, 2013]. One common conflict arises when a stereoscopic 3D display is used without head tracking (or with incomplete motion coupling). In such cases, moving one's head does not yield the expected change in viewpoint, resulting in **illusory motion parallax** of the virtual scene. Essentially, the scene appears distorted as if it was to "wobble" or slosh around. Benzeroual and Allison [2013] describe this effect as the world appearing sheared or "rubbery" when the observer's lateral head motion isn't accompanied by corresponding image shifts. The visual system, detecting stereoscopic depth but no appropriate head-coupled generated motion parallax, interprets the discrepancy as the scene itself moving relative to the observer.

Such depth cue conflicts can undermine the sense of stability and immersion in VR. Ono and Ujike [2005] systematically investigated conditions for visual stability versus illusory motion when observers made head movements. They identified the parameters under which motion parallax driven by head movements yielded a stable world, and conversely when it led to perceived depth changes or concomitant motion in the scene [Ono and Ujike, 2005]. Their work highlights that a lack of congruence between physical head motion and virtual camera motion is a key trigger for these illusions. In practical terms, VR systems must update the view in real-time with head movement; otherwise the user's brain, expecting parallax, will generate a false motion percept (as if the world is moving). Indeed, questions often arise regarding the limits of this illusion; for example, does closing one eye abolish it? It does, since binocular disparity is then absent, consistent with the dependence of such effects on binocular disparity [Rogers and Graham, 1982]. Shimono et al. [2002] reported that the strength of illusory motion increases with both disparity magnitude and head-movement amplitude, although the precise thresholds depend on stimulus geometry and viewing conditions. Beyond illusory motion, other depth cue conflicts in VR can also cause perceptual issues. A well-known example is the vergence–accommodation conflict, in which stereoscopic displays specify depth while accommodation remains fixed at the display plane, often leading to visual discomfort [Hoffman et al., 2008]. Likewise, mismatches in scaling of motion parallax (e.g. an incorrect "gain" between real head movement and virtual camera movement) can distort depth perception [Allison et al., 2003, Benzeroual and Allison, 2013]. Researchers have started to link these conflicts to user experience issues: for instance, inconsistent parallax cues have been hypothesized to contribute to simulator or cybersickness. Benzeroual and Allison [2013] conjectured that the non-rigid shearing illusion caused by untracked head movements might induce discomfort or motion sickness in stereoscopic games, although their experimental results found only limited increases in symptoms under those conditions.

Recent work has operationalized these cue combinations using controlled viewing setups that independently manipulate binocular disparity and motion parallax to create picture, window, parallax-only, and stereo-only viewing conditions [Wang and Troje, 2024, 2023]. These configurations make it possible to isolate the contribution of each cue and to study what happens when one cue is present in the absence of the other, providing a controlled experimental analogue of the cue-conflict situations described above.

Finally, it is worth noting that analogous depth-cue conflicts and resulting illusions are not unique to VR but also occur in real-world illusion setups. Papathomas and Bono [2004] demonstrated dramatic **depth inversion illusions**, for example, a concave mask that appears convex (the "hollow-face" illusion) or a reversed-perspective sculpture where conflicting cues make the object look real but with flipped depth. As observers move, such stimuli can appear to rotate or deform in ways that are inconsistent with their physical structure, reflecting the visual system's preference for a stable depth interpretation over strict geometric accuracy. These effects underscore how strongly the visual system will favor a consistent depth interpretation, even at the cost of introducing illusory motion. Papathomas and Bono [2004] showed that certain conditions (like increased viewing distance) can strengthen the illusion, indicating the role of cue reliability and higher-level expectations, consistent with Bayesian models of perceptual inference in which sensory evidence is weighted by reliability and combined with prior expectations [Ernst and Banks, 2002, Knill and Pouget, 2004]. Such illusions provide insight into the same principles at work in VR: the brain resolves conflicting depth signals by selecting an interpretation (sometimes an incorrect one), which can yield compelling but illusory motion. By drawing on the knowledge from these studies, from classic concomitant motion experiments [Gogel, 1990, Peterson, 1986] to modern VR investigations [Allison et al., 2003, Ono and Ujike, 2005, Benzeroual and Allison, 2013], researchers and developers can better predict and mitigate when illusory motion parallax will occur. Ensuring consistency between binocular disparity and motion parallax in VR is key to preventing unintended scene motion illusions and maintaining a stable, comfortable depth perception experience. These phenomena highlight the role of strong priors for rigidity, stationarity, and coherent three-dimensional structure in resolving conflicting depth cues, even when doing so produces compelling but non-veridical motion percepts. This interpretation is consistent with constructivist accounts of perception [Gregory, 1997], in which the visual system relies on prior knowledge and hypothesis-like interpretations to resolve ambiguous sensory input, and with Bayesian accounts [Kersten et al., 2004], which formalize this idea in terms of probabilistic inference over scene structure. Within these frameworks, visual illusions are understood as systematic consequences of normally adaptive assumptions about scene structure, stability, and object coherence.

Taken together, prior research has established that conflicts between binocular disparity and motion parallax can give rise to compelling illusions of motion, both in laboratory settings and in virtual reality environments. These studies have primarily focused on the conditions under which such cue conflicts occur, the role of disparity magnitude and head movement, and the resulting perceptual consequences in terms of motion and stability.

However, existing work has largely treated these illusions as local cue conflicts, without considering how broader **scene context** influences their strength and interpretation. In particular, it remains unclear how the presence or absence of surrounding environmental structure, spatial layout, and contextual depth cues modulates the perceptual outcome of cue conflicts. Classic work on layout perception emphasizes that depth is redundantly specified and that observers exploit multiple cues in a context-dependent manner, rather than relying on a fixed hierarchy across viewing situations [Cutting and Vishton, 1995]. In addition, formal and empirical work on pictorial versus visual space highlights that perceived spatial structure depends on both available cues and the viewing situation, thereby motivating explicit manipulations of scene structure and display framing in experimental settings [Koenderink and van Doorn, 2008, Koenderink et al., 2011]. More broadly, work on perceptual reference frames shows that performance can depend on whether stimuli are encoded relative to egocentric, environmental, or gravitational coordinates, underscoring that context can include global reference-frame structure in addition to local visual cues [Troje, 2003, Chang et al., 2010].

Furthermore, while several accounts implicitly invoke expectations about scene stability and coherence, these effects have not been systematically examined within a **predictive processing or forward-model framework**. A natural theoretical lens for these expectation-driven effects is predictive processing, in which perception is cast as inference under a generative model and mismatches between predicted and observed sensory signals (prediction errors) drive updates to perceptual hypotheses. Classic formulations include predictive coding accounts in early vision and hierarchical inference and free-energy formulations of perception and action [Rao and Ballard, 1999, Friston, 2010, Clark, 2013].

Within this broader framework, Bayesian accounts emphasize that sensory cues are weighted according to their reliability and uncertainty, predicting context-dependent reweighting when signals conflict [Knill and Pouget, 2004]. Taken together, these perspectives suggest that perceptual outcomes in cue-conflict situations should depend not only on the local signals themselves, but also on the broader scene context that shapes prior expectations and cue weighting.

Concretely, we interpret scene context as a latent variable that influences the inferred generative model of the display (e.g., picture-like versus window-like viewing). By shifting the assumed relationship between head movement and retinal motion, scene context alters predictions about motion parallax and thereby modulates the magnitude of the resulting prediction error.

The present work addresses this gap, namely the lack of a systematic, predictive-processing account of how broader scene context shapes illusory parallactic motion, by investigating how scene context modulates illusory parallactic motion in virtual reality, and by interpreting these effects within a predictive framework in which expectations about motion parallax are shaped not only by local depth cues, but also by the global structure of the visual environment.

# Chapter 3

# Objectives, Questions, and Approach

The primary objective of this research is to develop a robust method for measuring illusory motion that arises when visual depth cues are decoupled in virtual reality (VR). Specifically, this project focuses on the phenomenon that occurs when stereoscopic depth is presented without accompanying motion parallax, creating a mismatch that leads to the perception of motion in otherwise static scenes. Before exploring theoretical explanations for this illusion, it is necessary to establish a reliable and replicable experimental procedure for quantifying its perceptual strength.

The chapter is organized around three testable hypotheses and a sequence of five experiments. Hypothesis 1 asks whether binocular disparity is necessary for the illusion. Hypothesis 2 asks whether the perceived motion has a signed direction predicted by a forward model of motion parallax. Hypothesis 3 asks whether the strength of the illusion depends on contextual and geometric features of the surrounding scene. The five experiments are introduced in Section 3.5.

Building upon this, the following research objectives guide the proposed work:

## 3.1 Development and Validation of Quantification Method of the Illusion (Measurement Paradigm)

The objective is to **design and validate** an experimental paradigm that allows participants to compare motion in two types of displays: one that presents binocular disparity without motion parallax (stereo-only), and one that provides head-coupled motion parallax without binocular disparity (parallax-only). The parallax-only display serves as a reference stimulus, in which the magnitude and direction of motion parallax can be systematically varied. By adjusting the parallax gain of this reference stimulus, we aim to determine the Point of Subjective Equality (PSE) at which the perceived motion in both intervals is matched.

## 3.2 Testing the Role of Binocular Disparity

We aim to investigate the role of binocular disparity in driving the observed illusory motion. This leads to the first testable Hypothesis:

### 3.2.1 Hypothesis 1: Necessity of Binocular Disparity

The illusion is expected to vanish when the stereo-only display is viewed monocularly, since no binocular depth information is available to generate and constrain predictions of motion parallax.

## 3.3 Testing Direction Predictions from Forward Models

We aim to investigate whether the observed illusory motion can be explained by the brain's internal predictions of motion parallax based on stereoscopic depth. Under a forward-model account, the visual system uses an efference copy of the head movement together with the inferred depth structure of the scene to predict the retinal motion that should accompany that movement. When the display provides stereoscopic depth but no head-coupled motion parallax, the predicted parallax is not delivered by the input. The unconsumed prediction is then experienced as scene motion, and because it represents the parallax that *should have* occurred but did not, the percept should be opposite in direction to that physical parallax (i.e., the parallax that would normally result from the observer's head movement). This theory, rooted in forward-model mechanisms, leads to the second testable Hypothesis:

### 3.3.1 Hypothesis 2: Direction of Illusory Motion

The perceived motion in the stereo-only condition is expected to be opposite in direction to the real motion observed in parallax-only displays with positive parallax gain (i.e., parallax consistent with natural viewing geometry), consistent with the interpretation of the percept as a prediction error: the parallax that the forward model expected but the display did not deliver.

## 3.4 Testing Contextual and Geometric Modulators of the Illusion

We designed experiments to examine how the strength of illusory motion parallax changes as a function of the available visual references, the distance of those references from the judged display, and the surrounding scene context, as well as changes in viewing angle and observer distance from the display.

### 3.4.1 Hypothesis 3: Contextual and Geometric Modulators

The strength of the illusion is expected to be modulated by the visual references surrounding the judged stereoscopic display and by the viewing geometry. The contextual prediction is that more numerous environmental references, or references spatially closer to the judged display, should strengthen the illusion. The geometric prediction is that the viewing angle and the observer's distance from the display should also modulate the illusion.

The shared rationale is that surrounding references produce a stronger parallax signal during head movement, which provides a more sensitive reference against which the missing parallax of the stereo-only display is compared, increasing the apparent motion attributed to the display content. Viewing angle and observer distance alter how effectively those references can be used to interpret the cue conflict, and are therefore expected to shift the strength of the illusion in the same way.

## 3.5 Experimental Sequence

The hypotheses and objectives above are tested through a sequence of five experiments, described in detail in the Methodology chapter. The intuition behind the sequence is to first establish that the illusion can be measured reliably, then determine whether it has the *signed* structure predicted by a forward model (a percept whose direction, not just magnitude, matches the missing parallax), and finally ask how strongly that illusion depends on the surrounding scene and viewing geometry:

- **Experiment 1: Viewing Eye Manipulation** establishes the basic measurement paradigm by comparing a stereo-only illusory parallax display with a parallax-only measurement display across a range of parallax gains. The intuition is that an illusion must first be placed on the same scale as real motion parallax before its strength can be compared across conditions. This experiment also tests the Hypothesis that binocular disparity is necessary for the illusion by removing disparity from the illusory display in control conditions.

- **Experiment 2: Direction of Illusory Motion Parallax** tests the predicted direction of the illusion. If the illusion reflects a forward-model mismatch, one expects it not simply to produce generic apparent motion but to align with a specific parallax direction implied by the stereoscopic depth structure. This experiment therefore asks whether the perceived motion in the stereo-only display behaves like positive or negative motion parallax.

- **Experiment 3: Scene Context Variations** examines whether the surrounding scene context shapes the illusion. At a high level, this experiment asks whether nearby versus farther visual references change how observers interpret the stereo-parallax cue conflict.

- **Experiment 4: Viewing Angle Manipulation** asks whether the viewing angle, specifically, the vertical angular offset of the display relative to the observer's eye level, modulates the strength of the illusion. By shifting the display so the observer looks down at it rather than straight ahead, this manipulation tests whether the spatial relation between the observer and the display changes how effectively surrounding visual references are used to judge motion parallax. The Hypothesis is that this viewing geometry may alter the effective strength of the mismatch between expected and absent parallax.

- **Experiment 5: Observer Distance Manipulation** tests whether observer distance from the display changes the illusion. The intuition is that distance changes both the geometry of expected parallax and the effective salience of the display within its environment. This experiment therefore examines whether the same stereo-parallax mismatch is perceived differently when the observer is closer to or farther from the display.

Together, these experiments provide the methodological structure for the remainder of the thesis.

# Chapter 4

# Methodology

This study investigates the phenomenon of illusory parallactic motion in virtual reality (VR) by systematically manipulating depth cues using a head-mounted display (HMD).

## 4.1 General Methods

The general methodology defines the participant protocol, apparatus, stimuli, design logic, and procedure that were reused across the experiments. Every experiment was **within-subject**: each participant experienced all levels of every factor manipulated in that experiment, with trial repetition and counterbalancing distributing observations across conditions. Experiment-specific sample sizes, stimulus manipulations, factor levels, and response tasks are reported within the corresponding experiment sections below.

### 4.1.1 Participants

Participants were recruited from three sources: (1) the York University Undergraduate Research Participant Pool (URPP), (2) other York University research labs, and (3) volunteers external to the university. Inclusion criteria required participants to have normal stereoscopic vision. Participants were eligible if they could complete the VR task and reported normal or corrected-to-normal vision. Participants received either 1 course credit (if enrolled in PSYC1010) or a monetary honorarium of $20. Each session lasted approximately 40 minutes. Participants were recruited from multiple sources to broaden the participant base and to increase the generalizability of the findings beyond a single participant pool.

Across the full study, we recruited **165 participants** overall. This total includes participants who took part in additional pilot sessions, test experiments, and exploratory variants that were used during development of the paradigm. The final thesis reports only the experiments retained for the present analysis; pilot/test experiments and participants whose sessions belonged only to those excluded variants were not included in the reported datasets. Among the available self-identification records ($N = 93$), participants had a mean age of **20.1** years ($SD = 2.7$) and a mean height of **168.3 cm** ($SD = 11$). Reported biological sex was **56 female** (**60.2%**) and **37 male** (**39.8%**). In addition, **42 participants** (**45.2%**) reported that their vision was corrected to normal with glasses or contact lenses, and **17.2%** reported other vision problems (e.g., astigmatism).

All procedures were approved by York University's Human Participants Review Sub-Committee / Ethics Review Board under REB certificate **2025-199** for the project *Illusory Parallactic Motion in Virtual Reality*; the approval period is July 11, 2025–July 11, 2026. For each participant, the session began with informed consent and a short demographics form. Participants then completed a stereoscopic screening test before entering VR. We used the **Randot Stereopsis Test**, specifically the **10 graded circle items** probing stereo acuity between **400** and **20 seconds of arc**. Performance was recorded on a **1–10 scale**, corresponding to the number of graded circle judgments completed correctly. A score above 5, corresponding to a stereo acuity better than 70 seconds of arc, was treated as indicating normal stereoscopic vision for the purposes of inclusion in the study. Across the available session-level stereopsis records, the mean Randot score was **8.1** out of **10** ($SD = 1.7$, $n = 139$).

### 4.1.2 Apparatus

The apparatus consisted of a head-mounted display setup, a Unity rendering pipeline, and a virtual frame that we refer to as the *AlbertiFrame* throughout the remainder of the thesis. The mechanism of this virtual frame is described in the Stimuli subsection below.

Experiments were run on a **Meta Quest 3** head-mounted display (2064×2208 pixels per eye, up to 120 Hz). Scenes were authored in **Unity** (C#) and rendered on a connected **PC**; imagery was streamed to the headset over a **cabled** link to limit wireless latency. Head pose was tracked by the headset and fed each frame into the rendering pipeline. Trial order, blocking, and data logging were managed with the **BML-TUX** experiment framework [Bebko and Troje, 2020]. Where monocular viewing was required, either the left or right eye was randomly selected for digital occlusion in the rendering path for the relevant interval or display. The occluded eye received a blank black image rather than the experimental scene, while the unoccluded eye received the normal rendered view.

### 4.1.3 Stimuli

The general stimuli were defined by the AlbertiFrame rendering configurations and by a default three-cube scene geometry that later experiments modified only when their specific manipulation required it.

Participants stood in a **main** virtual room. The critical stimulus was not the full HMD framebuffer alone but the *AlbertiFrame*: a **virtual picture frame** (a mesh in the scene) showing an **embedded** rendering of a secondary 3D scene. Conceptually, the frame behaves as a window into another viewpoint: each eye sees a synthetic view of scene geometry computed as if observed from **controlled virtual cameras**, then mapped onto the frame surface. This decoupling permits **binocular disparity** in the embedded view to be presented **without** natural head-coupled motion parallax in that view, and conversely permits **motion parallax** to be driven by real HMD motion with an adjustable **gain**, independently of how disparity is configured.

On the first rendering frame of a new trial, the system performs an initial **picture-taking** step after headset recentering. That is, the system samples the current HMD pose and stores it as the reference viewpoint from which the embedded scene's stereo geometry is computed for the remainder of the trial. This records a **reference pose** for the frame and establishes a stable **picture location**: the viewpoint from which the embedded scene is nominally captured, expressed relative to the frame so that it remains coherent if the frame is moved in the world. An **anchor** aligned with the frame at picture time fixes the direction in which monocular virtual cameras "look" for subsequent updates, so geometry stays consistent with the framed view.

Operationally, motion-parallax gain was implemented as a world-space scaling of HMD translation around the picture location. Let $\mathbf{p}(t_0)$ be the recorded picture-time HMD position (captured at the initial picture time $t_0$ and expressed in world coordinates), $\mathbf{h}(t)$ be the current HMD position, $k$ be the motion-parallax gain, and $\mathbf{c}(t)$ be the embedded camera location. When motion parallax was enabled, the camera position was updated as

$$\mathbf{c}(t) = \mathbf{p}(t_0) + k\left[\mathbf{h}(t) - \mathbf{p}(t_0)\right].$$

Thus, for a fixed frame, the embedded camera translation from the reference pose was $k$ times the participant's head translation: $k = 1$ produced natural window-coupled parallax, $k = 0$ kept the embedded camera at the picture location, positive gains produced same-direction camera motion with scaled magnitude, and negative gains moved the embedded camera in the opposite direction. When motion parallax was disabled, the embedded camera remained at $\mathbf{p}(t_0)$ and reused the stored picture-time view.

Each frame uses a **camera location** transform (logical viewpoint) that is copied each update to a real-time **render camera**. That camera draws the embedded scene into an off-screen **render texture**. A **projector** component places the resulting image on the frame mesh (texture projected onto the canvas). Stereo presentation uses separate left- and right-eye camera locations with an inter-ocular offset consistent with the headset's inter-pupillary distance (with optional scaling where manipulated in specific designs). The loop (update camera and projector poses, render, update projected material) runs continuously while the interval is **live**.

The AlbertiFrame is controlled by two independent binary settings: whether **motion parallax** is enabled or disabled in the embedded view, and whether **stereo** (binocular disparity in the embedded view) is enabled or disabled, yielding **four** distinct rendering configurations. Independently of that choice, **two scalar gains** may be set in software: **motion parallax gain** scales the extent to which the embedded virtual viewpoint follows head translation relative to a pivot, and **inter-ocular distance gain** (IPD gain) scales the effective separation between the left and right embedded virtual cameras when stereo is active. The experiments reported in this thesis exclusively manipulated motion parallax gain; IPD gain was held at its natural default value. Holding the IPD gain constant preserves the stereoscopically specified depth of the illusory display, ensuring that any changes in the perceived illusion across conditions arise from our intended manipulations rather than an altered stereoscopic structure.

The four configurations are as follows:

- **Motion parallax on, stereo on:** The frame behaves like a **window** into the embedded scene: binocular disparity and head-coupled parallax are both consistent with a natural viewpoint. This configuration was available in the implementation but was not the focus of the illusory-parallax experiments reported here.

- **Motion parallax off, stereo on:** The embedded view carries **binocular disparity** but the virtual cameras do **not** move with the head in the parallax law; head movement therefore fails to produce the motion parallax predicted from the specified depth. This is the **illusory motion parallax display** (stereo-only interval) used in the thesis.

- **Motion parallax on, stereo off:** The embedded rendering is **monocular** (no binocular offset in the AlbertiFrame cameras), but the viewpoint **does** follow head motion with the chosen parallax gain. This is the **measurement (parallax-only) display**. The participant may view the HMD with **both eyes open** or with **one eye occluded** in software for the rest of the scene, as in Experiment 1; regardless, the content *inside* the AlbertiFrame is always the monocular parallax view.
- **Motion parallax off, stereo off:** Neither disparity nor head-coupled parallax updates the embedded view in the usual way; the result is a **static picture**, that is, a frozen embedded appearance suited to calibration or control displays.

The cue **decoupling** central to the thesis happens primarily on the second and third configurations above: stereoscopic structure without expected parallax in the illusory interval, and motion parallax without stereoscopic depth in the embedded measurement interval, with motion parallax gain varied to titrate perceived strength. In all experiments, parallax gain was applied only to the objects rendered inside the AlbertiFrame, while the frame itself remained fixed in world coordinates; any apparent motion was therefore perceived relative to the stationary frame.

Unless otherwise noted, the same base spatial arrangement of the measurement scene was retained in the later experiments, with only the experiment-specific manipulations changing from one study to the next. The base experiment used a fixed spatial layout that also served as the default **measurement-scene configuration** for the later experiments unless a specific manipulation altered it (for example, scene context, viewing angle, or observer distance). Along the viewing axis, the participant/reference origin was defined at $z = 1.62$. The virtual display on which either the illusory display or the measurement display was shown was positioned at $\mathbf{(}-0.509\mathbf{,}\ 1.565\mathbf{,}\ 2.42\mathbf{)}$ and had physical dimensions of $0.629 \times 0.353$ **m**. In the default setup, the **center of the frame** was aligned with the participant's **eye level**. Consequently, the height of both the frame and the supporting table surface was adjusted per participant so that the table surface remained exactly **half the frame height** (0.177 **m**) below eye level.

A horizontal support surface (table surface) was placed behind the display at $y = 1.565$, while the room floor remained at $y = 0$. In this configuration, the lower edge of the frame was aligned with the table surface, so the display extended upward by 0.353 **m**. The table surface began only behind the frame rather than extending toward the viewer. Three cubes were placed on that surface behind the frame as the primary scene objects. Their nominal positions were **Cube 1**: $(-0.509, 1.565, 2.8)$, **Cube 2**: $(-0.641, 1.565, 3.1)$, and **Cube 3**: $(-0.361, 1.565, 3.1)$. Cube 2 and Cube 3 were each randomly relocated within a circular region in the $x$–$z$ plane centered on their nominal positions with radius 0.15 **m**, and both cubes also received a random rotation about the vertical ($y$) axis.

For the structured-background version of this base scene, a checkerboard wall was placed at $z = 9.84$. The wall was approximately **8.22 m** from the participant and approximately **7.42 m** behind the frame center. This arrangement provided a stable far-background reference while keeping the display, the table surface, and the three cubes fixed in a common world-centered layout. Figure 4.1 and Figure 4.2 show the same layout in top-down and side-view schematic form, respectively.

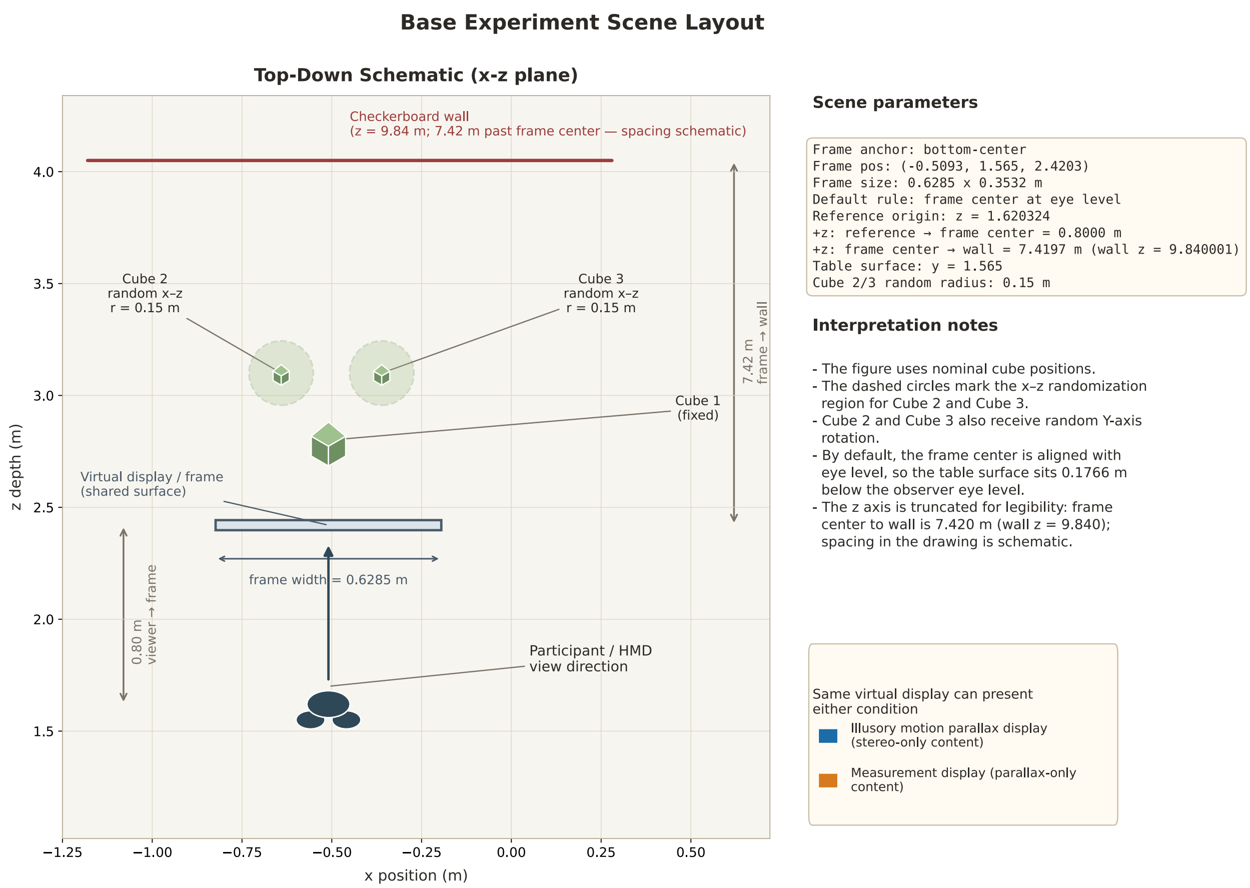


Figure 4.1: Top-down schematic of the general base scene layout in the $x$–$z$ plane. This geometry is used in the base experiment and retained as the default measurement-scene layout in later experiments unless explicitly manipulated. The top-down view shows the virtual display position, nominal cube locations, randomized $x$–$z$ placement regions for Cube 2 and Cube 3, and the far checkerboard wall, located approximately 8.22 m from the participant/reference origin.

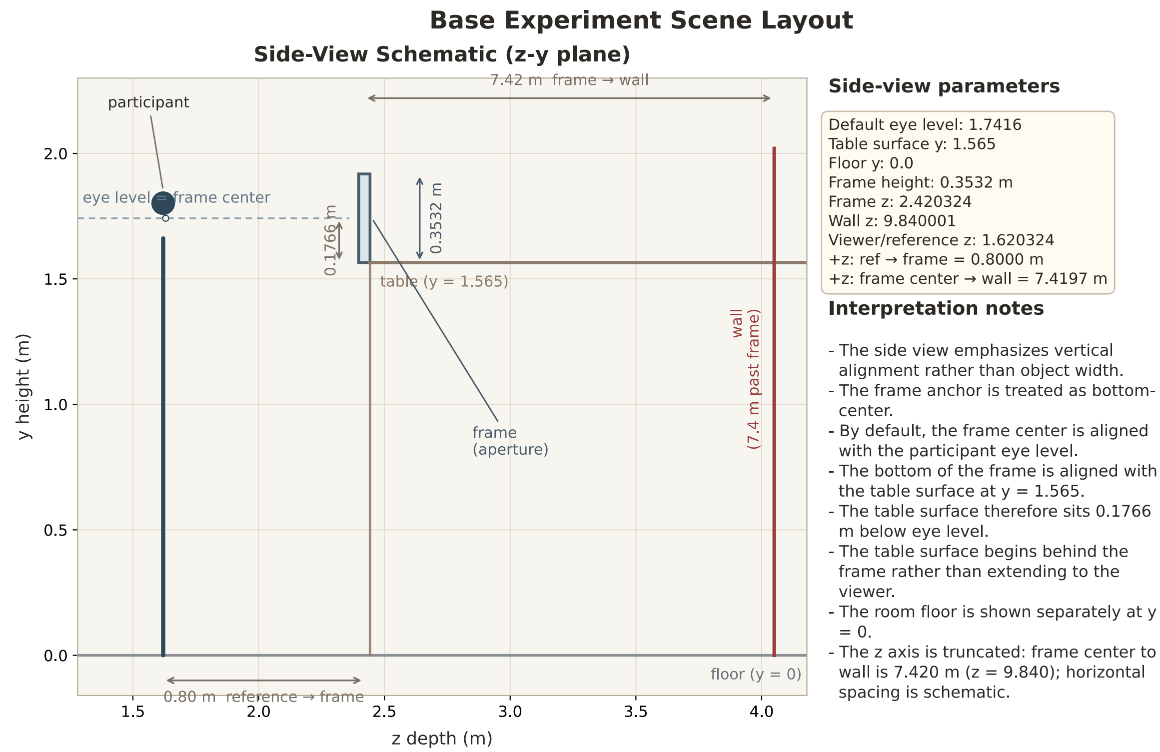


Figure 4.2: Side-view schematic in the $z$–$y$ plane, showing the table surface at $y = 1.565$ and the room floor at $y = 0$. The side view shows the participant, the frame aligned with the table surface behind it, the checkerboard wall approximately 7.42 m behind the frame center, and the separate room floor.

The layout described above defines the **default positioning setup** of the scene. In addition, before each session the entire relevant scene assembly received a participant-specific **vertical ($y$) offset** based on measured height so that the center of the AlbertiFrame aligned with the center of the participant's eye level. This same offset was applied to the frame, the eye-level reference, the table, and other coupled reference objects, preserving the spatial relations within the scene while standardizing the effective viewpoint across participants. As a result, although participants differed in physical height, they all viewed the display from the same intended eye-centered reference geometry.

### 4.1.4 Design

Across the experiments, the same cue-decoupling logic defined the two display types being compared. The **illusory display** refers to the **stereo-only** configuration (**parallax off, stereo on**), in which binocular disparity specifies depth but the embedded view does not update with head-coupled motion parallax. The **measurement display** refers to the **parallax-only** configuration (**parallax on, stereo off**), in which the embedded viewpoint follows head motion according to the assigned parallax gain but the embedded view itself contains no binocular disparity.

In the core magnitude-comparison design, what varies trial-by-trial is the head-coupled **parallax gain** applied to the **measurement display**, which remains in the parallax-only configuration throughout. The **illusory motion parallax display** stays in the stereo-only configuration (embedded disparity on, embedded motion parallax off), so the cue structure on the illusory side does not change when measurement gains are adjusted. Participants judge which interval contains stronger apparent motion; further spatial layouts and task formats are described in the experiment-specific sections. We then fitted a psychometric function to the response data to determine the point of subjective equality (PSE), which represents the applied gain in the measurement display that matched the perceived motion in the illusory parallax display. The specific psychometric model, fitting procedure, and PSE extraction are described in the Dependent Measures and Statistical Analysis section later in this chapter (Section 4.7).

Across experiments, parallax gain defined the scaling applied to head-coupled motion in the measurement display. By convention, positive values indicate parallax in the expected direction (consistent with physical scene geometry), while negative values reverse the direction of motion, creating a perceptually unnatural effect. Zero gain represents the absence of motion parallax. Each experiment specifies its own gain set and repetition count; when a dense base set was needed, gains were sampled more densely near zero to resolve the psychometric transition around equal perceived motion.

When an experiment manipulated **HMD viewing configuration**, the relevant display could be shown either *binocularly* (both headset displays active) or *monocularly* (one eye digitally occluded). It is important to distinguish this physical viewing state from the visual cues actually rendered within the display. For example, viewing a scene binocularly simply means both eyes receive an image; it does not automatically add stereoscopic depth to the content. Therefore, even when viewed binocularly with both eyes active, the parallax-only measurement display still presents the exact same 2D projection to each eye, providing motion parallax but zero stereoscopic disparity. Conversely, applying monocular viewing to the stereo-only display removes the observer's physical ability to perceive its rendered disparity, allowing us to test whether the illusion relies on binocular depth.

For the magnitude-comparison experiments, the default task format was a **single-frame two-interval forced-choice (2IFC)** procedure. Participants viewed one interval containing the illusory display and one interval containing the measurement display, with interval order counterbalanced, and then judged which interval appeared to contain stronger motion. Experiments that changed this default format or added an additional manipulation retained the same display definitions but specify their own layout, gain values, viewing mode, and response task in the sections below.

### 4.1.5 Procedure

Before trials, participant **standing height** was applied as a vertical offset to the frame, table, and related reference objects so that the **center of the AlbertiFrame** aligned with that participant's eye level. Because the table surface was coupled to the frame geometry, it was always positioned **half the frame height** below eye level, unless manipulated in specific designs. After height adjustment and recentering, an initial picture-taking pass aligned the AlbertiFrame reference geometry for that participant. The same block-level pattern (VR mode, frame activation, height, recenter, initial picture) was reused across experiment code paths, while experiment-specific manipulations were handled at the trial or task level as described in the corresponding experiment sections. At the start of each trial, the headset was recentered to restore the intended observer/reference origin before each individual trial. This trial-level recentering limited unintended depth drift by keeping the observer-to-frame distance matched to the current trial's geometry; in Experiment 5, the same recentering rule used the trial's assigned distance condition.

Wherever we used a two-interval procedure, the order of the intervals is randomized. During each interval, participants are instructed to generate modest self-induced motion (specifically, slightly leaning left and right on their feet) to engage parallax when available. Because motion parallax is inherently head-coupled, this active movement is essential for observers to experience the parallax cues in the measurement display and the dynamic cue conflict in the illusory display. After viewing both intervals, participants are asked to indicate which interval appears to contain more motion. Participants were instructed to judge the apparent motion of the scene relative to the display, rather than their own head movement. Each trial takes approximately 10 seconds to complete, including two 3-second viewing intervals separated by a 0.5-second transition, as well as short inter-trial transitions. All trials are presented in randomized order to control for sequence effects and ensure unbiased responses when participants compare motion across viewing modes.

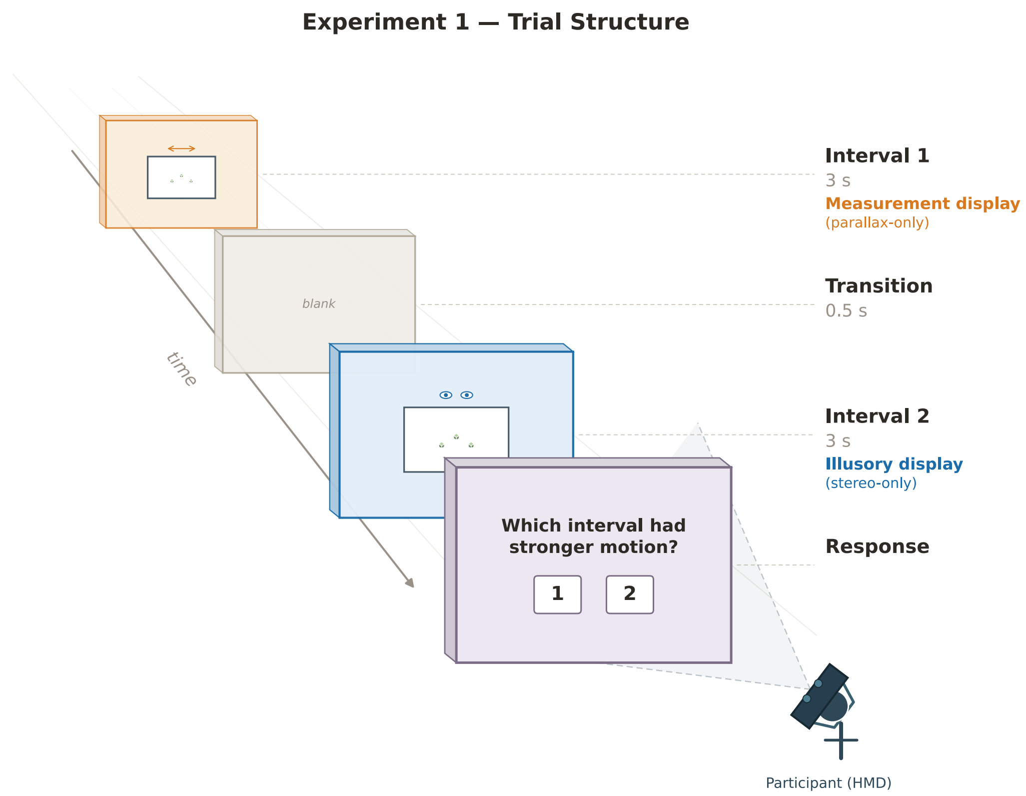


Figure 4.3: Two-interval magnitude-comparison trial structure used in the base paradigm. Each trial presents the stereo-only illusory parallax display and the parallax-only measurement display in randomised order, separated by a 0.5 s blank transition. After both intervals, the participant indicates which interval appeared to contain stronger motion. Panel size recedes with depth to convey the temporal sequence from the participant's viewpoint.

## 4.2 Experiment 1: Viewing Eye Manipulation

### 4.2.1 Participants

Experiment 1 included **20 participants**. Age and height records were available for all 20 participants. Mean age was 20.5 years ($SD$ = 2.42, range = 18–27), and mean height was 166.24 cm ($SD$ = 9.53, range = 152–188). Randot stereopsis scores were available for **17 participants** ($M$ = 8.59/10, $SD$ = 1.18, range = 6–10). Recruitment, consent, compensation, and stereo-screening procedures followed the general participant protocol described in Section 4.1.1.

### 4.2.2 Stimuli

Experiment 1 used the general **single-frame base scene** described above: the AlbertiFrame positioned at eye level, the three-cube scene behind the frame, and the structured far background. The **illusory interval** used the **stereo-only** display (*motion parallax off, stereo on*), whereas the **measurement interval** used the **parallax-only** display (*motion parallax on, stereo off*).

The critical stimulus manipulation in this experiment was the **HMD eye state**. Binocular viewing activated both headset displays, whereas monocular viewing digitally occluded one eye (randomly left or right). In the standard conditions, the illusory display remained binocular and only the measurement display was shown under binocular versus monocular headset viewing. In the extended conditions, the illusory display itself was also shown monocularly to test the Hypothesis that binocular disparity is necessary for the illusion.

### 4.2.3 Design

Experiment 1 was organized into two sets. In **Set A**, the illusory display remained binocular and the measurement display was presented under **binocular** and **monocular** HMD viewing. The measurement display used **11 parallax gain values**, {-1, -0.5, -0.25, -0.125, -0.0625, 0, 0.0625, 0.125, 0.25, 0.5, 1}. Each combination of gain and measurement-display viewing mode was repeated **six** times, with whether the illusory or measurement interval appeared **first** counterbalanced across those repetitions, yielding **132 trials** ($11 \times 2 \times 6$).

In **Set B**, the illusory display itself was rendered monocularly, while the measurement display was again tested under binocular and monocular HMD viewing. This set used **five gain values**, {-1, -0.125, 0, 0.125, 1}, with six repetitions per gain-by-viewing-mode combination (again with counterbalanced interval order), yielding **60 additional trials** ($5 \times 2 \times 6$). Across both sets, Experiment 1 therefore comprised **192 trials per participant**.

### 4.2.4 Procedure

Each trial followed the general **2IFC** procedure. Participants viewed one illusory interval and one measurement interval in randomized order, generating modest self-induced motion to engage parallax as described in the general procedure (Section 4.1.5). After the second interval, they indicated **which interval appeared to contain stronger motion**. All **192 trials** were presented in a single fully randomized order rather than as two separately experienced sets.

## 4.3 Experiment 2: Direction of Illusory Motion Parallax

### 4.3.1 Participants

Experiment 2 included **41 participants**. Age records were not available in the updated demographics table for this experiment. Height records were available for all 41 participants, with a mean height of 162.62 cm ($SD = 8.96$, range = 143–182). Randot stereopsis scores were available for **39 participants**, with a mean score of 8/10 ($SD = 1.95$, range = 0–10). Recruitment and screening followed the common procedures described earlier in the chapter.

### 4.3.2 Stimuli

Experiment 2 used a **dual-frame spatial layout** that was specific to the direction-judgment task. Instead of presenting the illusory and measurement displays in two temporal intervals, this experiment displayed **two AlbertiFrames simultaneously**, one to the left and one to the right of the participant. One frame always contained the **stereo-only illusory parallax display**, and the other contained the **parallax-only measurement display** at the trial's assigned motion-parallax gain. Presenting the displays side by side allowed participants to judge the relative **direction** of the apparent motion under the same bout of self-induced head movement, rather than comparing motion strength across remembered intervals. The side assigned to the illusory display was counterbalanced across trials so that any left–right bias could be separated from the perceived motion direction.

### 4.3.3 Design

Experiment 2 used a **direction-judgment design**. The translation gain applied to the measurement display was drawn from the set $\{-1, -0.5, -0.1, 0.1, 0.5, 1\}$. Positive and negative gains produced parallax in opposite directions, enabling us to test which direction participants perceived as matching the illusory motion. The side containing the stereo-only illusory parallax display was counterbalanced across trials. Each gain level was repeated **six** times in total, with **three trials** placing the stereo display on the right frame and **three trials** placing it on the left frame, yielding **36 trials per participant**.

### 4.3.4 Procedure

Rather than making a magnitude comparison, participants performed a **two-alternative forced-choice (2AFC) direction judgment**. On each trial, they viewed the two displays and used self-induced motion to reveal the relation between the stereo-driven illusory motion and the measurement-display parallax. They then indicated whether the perceived direction of motion in the illusory motion parallax display was the *same as* or *opposite to* the direction of motion in the measurement display. This binary judgment, together with response time, was recorded on every trial. The design allowed us to determine whether the illusory motion behaved like positive or negative parallax, thereby addressing Hypothesis 2 outlined in Section 3.3.1.

## 4.4 Experiment 3: Scene Context Variations

### 4.4.1 Participants

Twenty-four participants were recruited for Experiment 3. Six were excluded prior to analysis: some had incomplete sessions in which not all scene-layout $\times$ frame conditions were completed, and others did not yield interpretable participant-level psychometric fits across the required set of conditions. The final analyzed sample was **18 participants**. Age, height, and Randot stereopsis records were available for all 18 participants. Mean age was 20 years ($SD = 2.401$, range = 18–27), mean height was 169.269 cm ($SD = 8.003$, range = 155–185), and the mean Randot score was 7.722/10 ($SD = 1.447$, range = 5–10). Recruitment and screening followed the common procedures described earlier in the chapter.

### 4.4.2 Stimuli

Experiment 3 used the general **single-frame two-interval setup** but varied the **scene context** surrounding the judged display. The illusory interval was always the **binocular stereo-only** display, and the comparison interval was a **binocular parallax-only** measurement display using negative gains. No eye was digitally occluded in this experiment so that the scene-context manipulation could be tested without adding the headset eye-state factor used in Experiment 1, and so that the measurement display remained visually comparable across scene layouts.

Three scene layouts were used:

- **Closed environment**: the same cube-on-table scene was shown with a structured background approximately 8.22 m from the participant/reference origin, or 7.42 m behind the frame center, with no external environment visible beyond the frame.

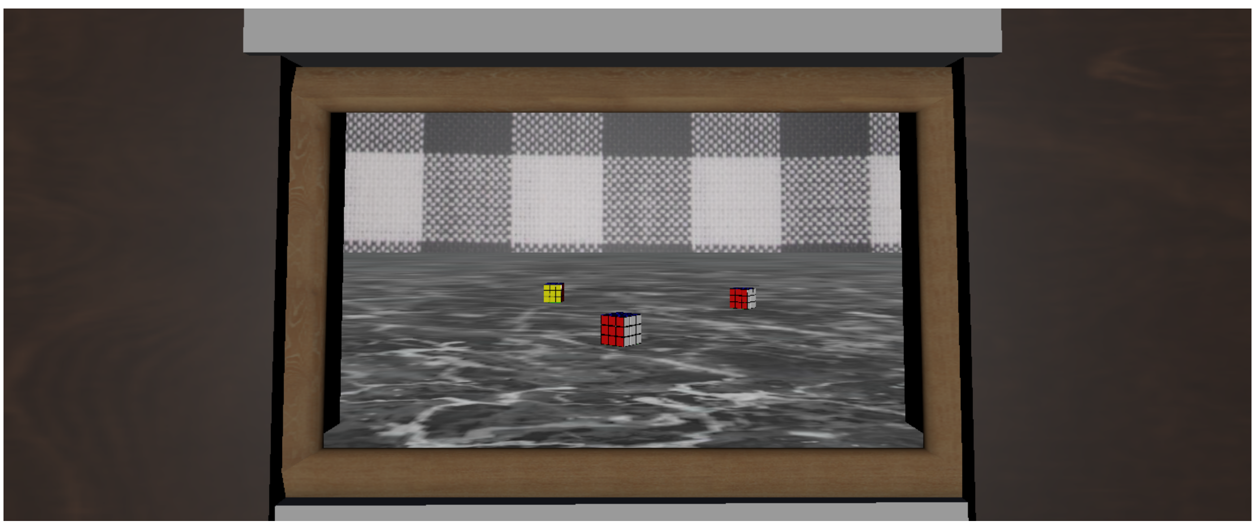

Figure 4.4: Closed environment condition in Experiment 3. The cube-on-table scene is shown with a structured background inside the display, with no broader virtual room visible beyond the frame; this condition provides local scene structure while limiting external environmental references.

- **Open environment**: the cubes and table remained visible inside the frame, but participants could also see the surrounding virtual environment outside the frame, providing additional peripheral context.

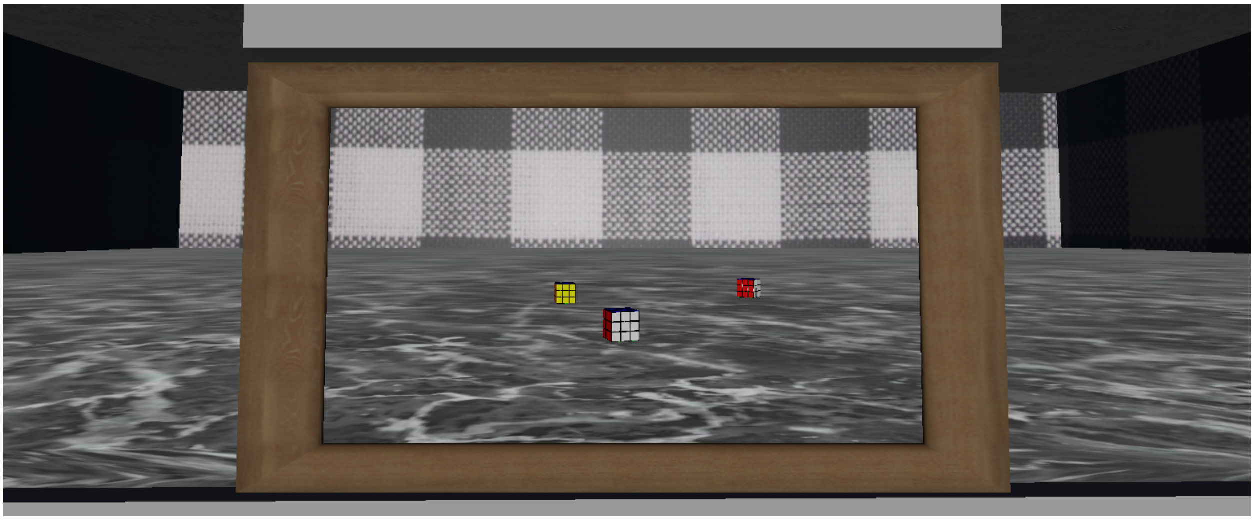

Figure 4.5: Open environment condition in Experiment 3. The cubes and table remain the judged scene inside the frame, but the surrounding virtual environment is also visible outside the frame, adding peripheral spatial context while preserving the same central comparison display.

- **No environment**: the judged scene was presented without the broader environmental structure used in the other two layouts, isolating the relation between the frame and the scene content itself.

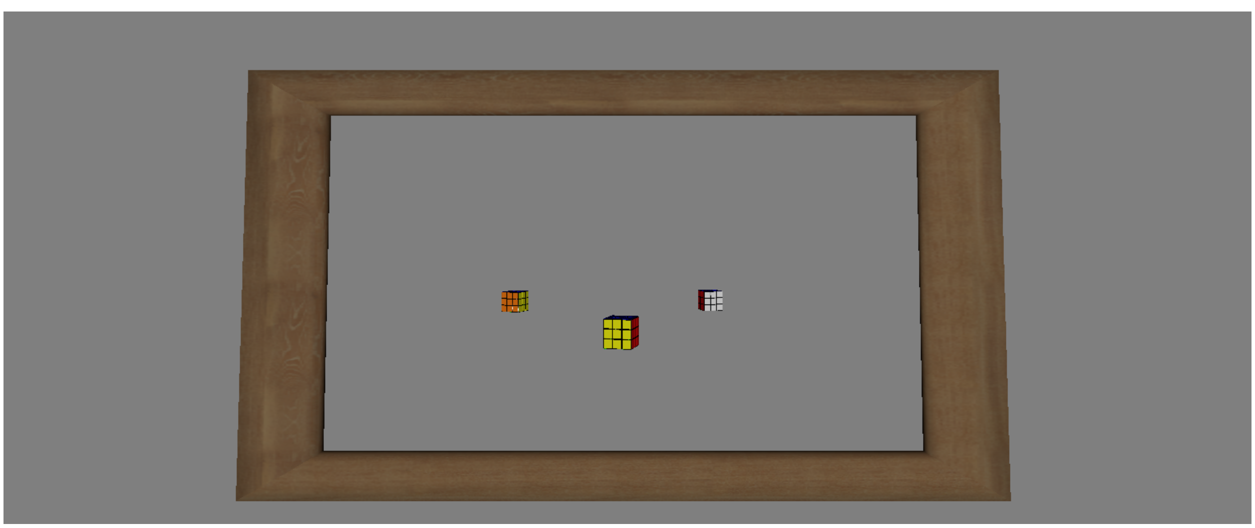

Figure 4.6: No-environment condition in Experiment 3. The broader environmental structure is removed so that the judged display is dominated by the relation between the AlbertiFrame and the rendered scene content, isolating the local frame-content reference.

Frame visibility was manipulated independently of scene layout. When the frame was **visible**, a border surrounded the judged display and served as an explicit reference for relative motion. When the frame was **hidden**, the same scene content extended to the display edges without that border.

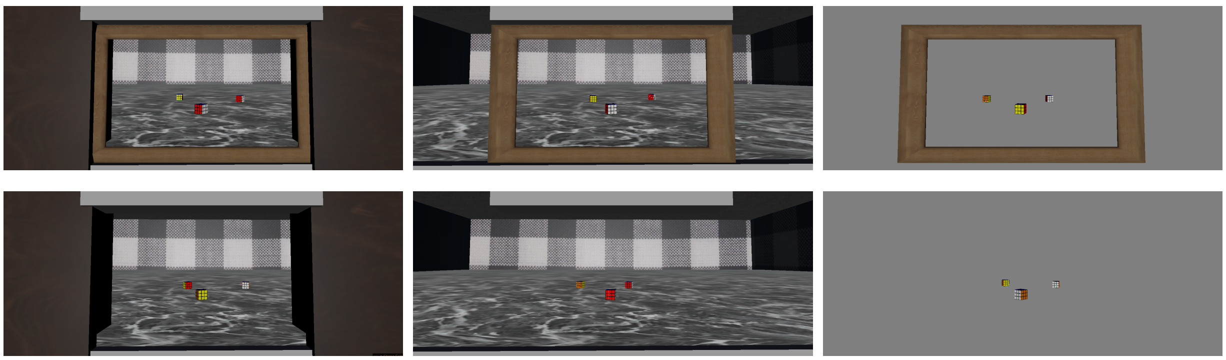

Figure 4.7: Frame-visibility manipulation across all Experiment 3 scene layouts. Columns show the closed environment, open environment, and no-context scene layouts; the top row shows the frame-visible version of each layout, and the bottom row shows the corresponding frame-hidden version. The comparison illustrates that frame visibility was crossed with scene layout, allowing the experiment to test whether a visible boundary alters the judged strength of illusory motion within each scene context.

### 4.4.3 Design

Experiment 3 used a **factorial design** with two primary factors: **scene layout** (*closed environment*, *open environment*, *no context*) and **frame visibility** (*visible*, *hidden*). The measurement display used **four negative parallax gain values**, {-0.0625, -0.125, -0.25, -0.5}, chosen to compare the stereo-driven illusion against reversed-parallax motion in the same general perceptual direction. Each combination of scene layout, frame visibility, and gain value was repeated **six** times, with presentation order counterbalanced across repetitions, yielding **144 trials per participant** ($3 \times 2 \times 4 \times 6$).

### 4.4.4 Procedure

Each trial followed the general **2IFC** procedure. Participants viewed one illusory interval and one measurement interval in counterbalanced order while engaging in self-induced motion. After both intervals, they judged **which interval appeared to contain stronger motion**. The scene-layout manipulation changed the amount of surrounding environmental structure, whereas the frame-visibility manipulation changed whether a visible border framed the judged display. This allowed the experiment to test whether broader scene context and the explicit frame boundary modulate the perceived strength of illusory parallactic motion.

## 4.5 Experiment 4: Viewing Angle Manipulation

### 4.5.1 Participants

Experiment 4 included **12 participants**. Age records were available for all 12 participants ($M = 20.33$ years, $SD = 1.97$, range = 18–24), and height records were also available for all 12 participants ($M = 175.07$ cm, $SD = 12.99$, range = 157–204). Randot stereopsis scores were available for **11 participants** ($M = 7.73/10$, $SD = 2$, range = 3–10). Recruitment and screening followed the same procedures described for the broader study.

### 4.5.2 Stimuli

Experiment 4 used the general **single-frame magnitude-comparison setup** but manipulated the **vertical placement** of the frame-and-scene group so that both intervals were presented at different angular offsets relative to the observer. The motivation for this angle set was geometric: vertical frame placement was the primary manipulated factor, and the corresponding viewing angles were the values produced by the feasible vertical positions of the AlbertiFrame within the constraints of the scene layout. This manipulation used the same kind of vertical repositioning used during participant-specific height calibration at the start of the experiments, but applied it systematically on each trial according to the assigned viewing-angle condition. For a given trial, the virtual frame was translated vertically, and the table surface and cubes behind the frame were translated with it, preserving their relative positions, sizes, depth separations, and randomization regions. The observer position, frame dimensions, horizontal/depth distance to the frame, cube geometry, and gain values remained fixed. Viewing angle was computed as the absolute vertical angular offset between the participant's eye-height line and the table top position, $\theta = \tan^{-1}(|y_{\text{table}} - y_{\text{eye}}|/d)$, where $d$ was the horizontal $x$–$z$ distance from the eye/reference position to the table top position. The tested viewing angles were 2.83°, 13.14°, 22.90°, 31.60°, 38.80°, 44.84°, 50.00°, and 54.23°.

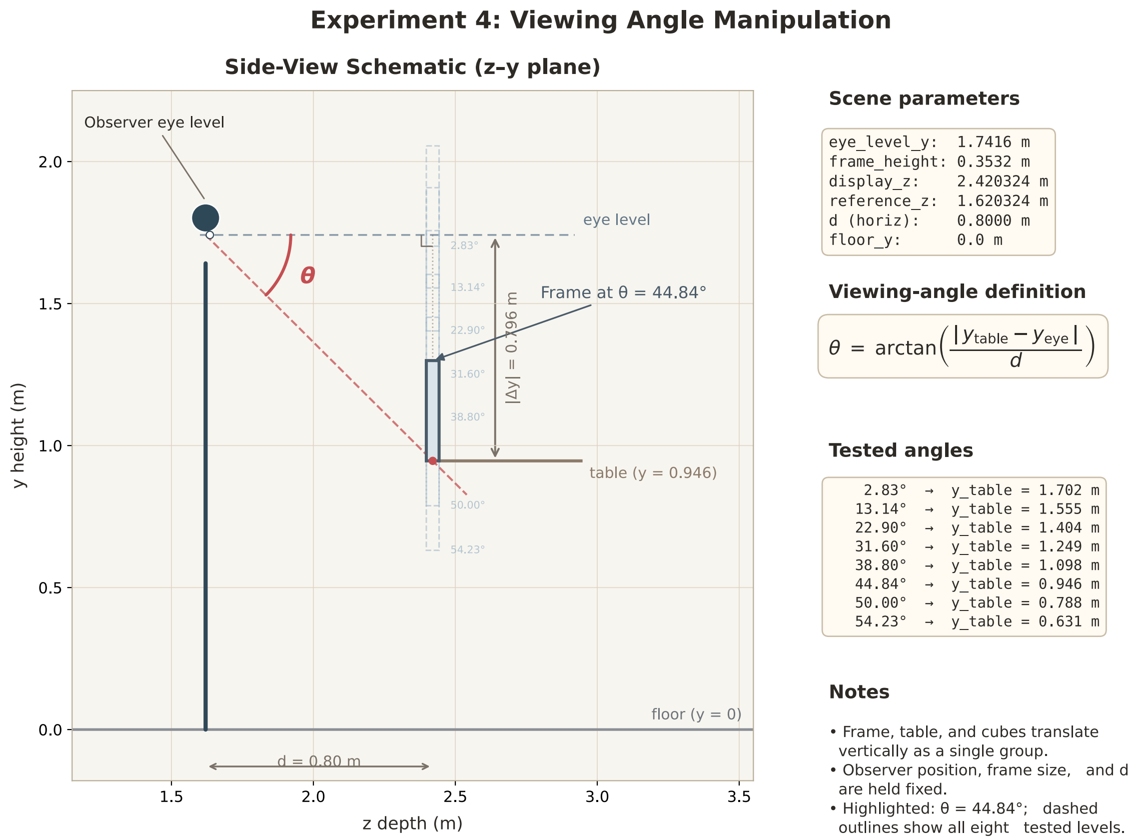


Figure 4.8: Side-view schematic of the viewing-angle manipulation in Experiment 4. The frame, table surface, and cubes were translated vertically as a single group to create the assigned angular offset $\theta$ from eye level, where $\theta = \tan^{-1}(|y_{\text{table}} - y_{\text{eye}}|/d)$ and $d$ is the horizontal distance from the observer to the display. Dashed outlines show all eight tested frame positions; the solid frame illustrates $\theta = 44.84°$. Display size, horizontal/depth distance, object layout, and gain values were held fixed across conditions.

The illusory interval remained the **binocular stereo-only** display. The measurement interval remained the **parallax-only** display and was presented under **binocular viewing** using the negative gain set {-0.0625, -0.125, -0.25, -0.5}. Within each trial, both intervals used the same repositioned frame-and-scene geometry so that the comparison varied viewing angle rather than the relative layout of the illusory and measurement displays.

### 4.5.3 Design

Experiment 4 crossed **viewing angle** (8 levels) with **measurement-display gain** (4 negative levels). Each angle-by-gain combination was repeated **six** times, with interval order balanced across repetitions, yielding **192 trials per participant** ($8 \times 4 \times 6$). On every trial, both intervals were presented at the **same** viewing angle so that the comparison isolated the effect of angle rather than a positional mismatch between the illusory and measurement displays.

### 4.5.4 Procedure

The procedure followed the same **2IFC** format used in the earlier magnitude-comparison experiments. On each trial, participants viewed one interval at the assigned viewing angle, generated brief self-induced motion during the presentation, then viewed the second interval at that same angle. After both intervals, they indicated **which interval appeared to contain stronger motion**. Interval order was counterbalanced, with three trials per condition presenting the measurement display first and three presenting the illusory display first.

## 4.6 Experiment 5: Observer Distance Manipulation

### 4.6.1 Participants

Experiment 5 included **19 participants** from the same recruitment cohort as Experiment 3. For these participants, the order of Experiments 3 and 5 was randomized across individuals, such that some completed Experiment 3 first and others completed Experiment 5 first. Age, height, and Randot stereopsis records were available for all 19 participants. Mean age was 19.79 years ($SD = 2.37$, range = 18–27), mean height was 168.6 cm ($SD = 10.35$, range = 151–195.58), and the mean Randot score was 7.32/10 ($SD = 1.77$, range = 4–10). Recruitment and screening were the same as in the other experiments.

### 4.6.2 Stimuli

Experiment 5 used the general **single-frame magnitude-comparison setup** but manipulated the **observer's distance from the frame**. The only geometric variable changed across distance conditions was the observer-to-frame separation along the viewing axis. For each trial, the frame-and-scene group was placed at the assigned distance from the observer/reference origin, while the frame dimensions, table and cube sizes, cube placement regions, height alignment, and AlbertiFrame rendering logic were left unchanged. No compensatory scaling or angle normalization was applied. As a result, changing distance also changed the visual angle subtended by the otherwise identical frame-and-scene arrangement. The tested viewing distances were **0.4 m**, **0.8 m**, and **1.2 m**.

The illusory interval was always a **binocular stereo-only** display so that disparity-defined depth was available in every condition. The measurement interval remained a **parallax-only** display. Its HMD viewing mode was manipulated between **monocular** and **binocular** presentation, and its motion-parallax gain was drawn from the positive set $\{0.125, 0.25, 0.5, 1\}$. Positive gains were used because they provided a more sensitive regime for revealing separation from the illusory display across conditions. The inclusion of both monocular and binocular measurement-display viewing modes also preserved comparability with Experiment 1. Apart from the calibrated observer distance, the general scene geometry and display implementation were unchanged.

### 4.6.3 Design

Experiment 5 crossed **observer distance** (3 levels), **measurement-display gain** (4 positive levels), and **measurement-display viewing mode** (2 levels: monocular, binocular). Each condition was repeated **six** times with counterbalanced interval order, yielding **144 trials per participant** ($3 \times 4 \times 2 \times 6$). The design therefore tested whether the gain required to match the stereo-only illusion changed systematically as the observer moved closer to or farther from the frame.

### 4.6.4 Procedure

Each trial followed the general **2IFC** procedure. Both the illusory interval and the measurement interval were presented at the **same assigned observer distance** so that the comparison isolated the effect of distance. Participants engaged in self-induced motion during each presentation and then reported **which interval appeared to contain stronger motion**. Trial order was randomized, and interval order was balanced across repetitions within every distance-by-gain-by-viewing-mode condition.

## 4.7 Dependent Measures and Statistical Analysis

The primary dependent measure was the participant's forced-choice response on each trial. In the magnitude-comparison experiments, raw responses were first normalized for interval order: when the stereo-only interval was presented second, the two response labels were flipped before binary coding. After this normalization, **1** indicated that the participant judged the **stereo-only/illusory interval** as containing stronger motion, and **0** indicated that the participant judged the **parallax-only measurement interval** as containing stronger motion. In Experiment 2, which used a direction-judgment task rather than a magnitude comparison, the binary response variable indicated whether the illusory display appeared to move in the **same direction** as the measurement display. Response times were recorded where available and treated as supplementary descriptive evidence rather than as the primary endpoint.

Before response coding and model fitting, trial logs were passed through a general preprocessing pipeline to reduce the influence of acclimation trials, timing artifacts, and response-time outliers. The first two trials of each participant session were removed as initial familiarization trials. Trials with non-positive response times were removed as timing-system errors, and trials with response times greater than 60 s were removed as likely inattention or stalled-interface events. For response-time summaries, remaining response times were capped within participant at the median plus three scaled median absolute deviations (median $+3 \times 1.4826 \times \text{MAD}$), limiting the influence of extreme but finite response-time values. The analysis code also allowed file-specific block/trial exclusions when a trial was documented as unusable; no additional manual block/trial exclusions were specified unless reported in an experiment-specific analysis. Valid-trial counts and valid-participant counts in the Results therefore reflect these preprocessing filters together with the fit-validity criteria described below.

For participant-level PSE estimation in the magnitude-comparison experiments (Experiments 1, 3, 4, and 5), responses were grouped by participant, condition, and measurement-display gain, and the mean binary response at each gain was fit with a single tanh psychometric function:

$$f_{\tanh}(g) = \frac{\tanh(s(g-o)) + 1}{2},$$

where $g$ is measurement-display gain, $s$ is slope, and $o$ is the horizontal offset. Because the tanh midpoint occurs at $f(g) = 0.5$ when $g = o$, the PSE was defined directly as the fitted offset $o$. No separate fitting branches or axis reversals were required: a positive slope ($s > 0$) produces the rising transition characteristic of negative-gain conditions, while a negative slope ($s < 0$) produces the falling transition characteristic of positive-gain conditions. The sign of the fitted slope therefore adapts automatically to the gain polarity of each condition, and the PSE is always the offset $o$ on the original measurement-gain axis. When participant-level fits did not converge or did not provide an interpretable 0.5 crossing within the sampled gain range, that participant-by-condition estimate was omitted from PSE summaries, while valid trial-level responses remained available for response-level analyses.

Experiment 2 was analyzed as a direction-judgment experiment rather than as a primary PSE experiment. Its binary response measured whether the illusory display appeared to move in the same or opposite direction as the measurement display, and the main analyses therefore used same-direction response proportions and trial-level logistic models rather than magnitude-matching PSEs.

Inferential analyses were chosen to match the response structure of each experiment. Participant-wise PSEs were analyzed with paired or one-sample tests when the design involved direct within-participant contrasts; normality checks and non-parametric or permutation-based tests were reported where appropriate. Trial-level binary responses were analyzed with binomial logistic generalized estimating equation (GEE) models when the analysis needed to retain the repeated-measures structure of individual trials. For continuous experimental manipulations, such as viewing angle, PSE-based trend analyses used mixed-effects or repeated-measures models as reported in the corresponding Results sections. All experiment-specific model formulas, valid-sample counts, and robustness checks are reported with the relevant results.

## 4.8 Proposed Model

Figure 1.1 illustrates our interpretation of illusory parallactic motion using the conceptual framework introduced by Troje [2019]. In his work, Troje distinguishes between two perceptual modes that determine how the retinal image is processed and interpreted: picture mode that is applied to projected images, and presence mode that is applied to tangible objects and surrounding scenes. For clarity, the term *illusory motion parallax display* is used consistently throughout to refer to stereoscopic displays lacking motion parallax.

In picture mode, even though real motion parallax is still physically available, the brain does not expect the scene to update in a geometrically consistent way with head movement. In this mode, sensorimotor contingencies are not interpreted to reflect the three-dimensional layout of the environment, and therefore no prediction is generated about how the scene would change under a new viewpoint.

In presence mode, sensorimotor contingencies matter: the observer expects motion parallax that is consistent with head movement and with the perceived three-dimensional structure of the scene.

We apply this framework to model the behavior of the illusory display. When the illusory motion parallax display is viewed binocularly, stereoscopic depth encourages presence mode, leading the brain to predict motion parallax. Because the illusory display provides no actual parallax, the mismatch between expected and received input is interpreted as scene motion, producing the illusion.

This model provides a simple explanatory structure for the binocular dependency of the illusion and guides how we interpret the behavior of the illusory display in our experiments.

# Chapter 5

# Results

## 5.1 Experiment 1: Viewing Eye Manipulation

### 5.1.1 Design and Descriptive Results

Experiment 1 was designed as a method-development study. Its main goal was to establish a procedure for measuring the strength of the illusion so that the same method could then be used in the later experiments. We collected data from 20 participants in a single experiment that included 11 different parallax gain values (both positive and negative) in the measurement display, allowing us to densely sample the psychometric function for each condition. Figure 5.1 first shows the general pattern of results for the standard illusion-producing case, in which the illusory display retained binocular stereoscopic depth. At this stage, the analysis is not yet focused on the viewing manipulation within the measurement display; instead, it provides an overall comparison of negative versus positive measurement gains against the binocular stereo-only illusory parallax display. These results are presented to motivate and validate the proposed measurement approach, rather than as a definitive test of the underlying theoretical account.

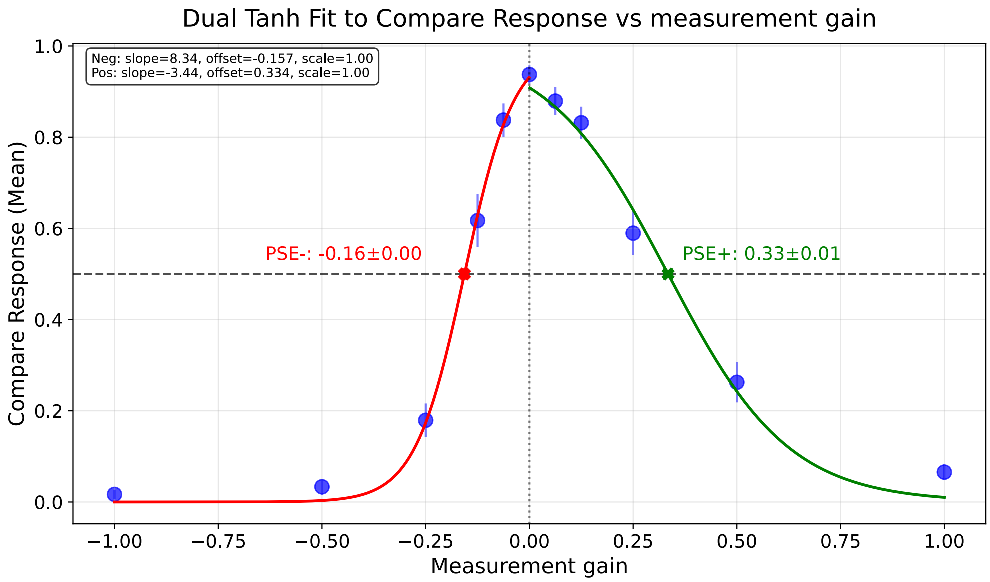


Figure 5.1: Percentage of trials where the illusory display was perceived as moving faster than the measurement display as a function of the parallax gain applied to the measurement display. The horizontal axis represents the parallax gain values in the measurement display. The vertical axis shows the proportion of trials (on a 0–1 scale) where participants perceived the illusory display as moving faster than the measurement display. A value of 0.5 indicates equal perceived motion between the two. This overview shows the standard binocular illusory display and emphasizes the contrast between negative and positive measurement gains before the measurement-viewing conditions are separated in the later Set A analysis.

To summarize the psychometric transition in this method-development experiment, we used the tanh PSE estimator defined in Section 4.7. Each condition was fit with a single tanh function whose slope sign adapts to the gain polarity, yielding the PSE directly as the fitted offset. These fitted curves are used to visualize response trends and to motivate the condition-specific PSE comparisons reported below, rather than to introduce a separate estimator.

The overview fit in Figure 5.1 suggests an asymmetry between gain polarities. When the measurement display used negative gains, participants judged it as stronger than the illusion across a wider range of values, whereas positive gains had to become larger before the measurement display was judged as stronger. This broad pattern motivated the condition-level analyses below, where gain polarity and HMD viewing configuration are separated explicitly.

To unpack this overall pattern, we next separated the data into the full set of experimental conditions. The eight conditions in Experiment 1 came from crossing the two viewing conditions of the measurement display (monocular, binocular) with the two gain polarities (positive, negative), and doing so for two versions of the illusory display itself. In **Set A**, the illusory display remained binocular, which is the standard condition in which the illusion is present. In **Set B**, the same four comparisons were repeated with the illusory display shown monocularly, allowing us to test whether the illusion depends on binocular disparity.

Thus, Figure 5.1 serves as a general summary of the binocular-illusory case, focusing only on the broad negative-versus-positive pattern. We now turn to Set A, where that same case is broken down by measurement-display viewing condition and shown in more detail across the four underlying condition combinations:

### 5.1.2 Set A - Binocular Illusory Display (standard illusion present):

1. **BB+**: Binocular illusory, binocular measurement, positive gains
2. **BM+**: Binocular illusory, monocular measurement, positive gains
3. **BB−**: Binocular illusory, binocular measurement, negative gains
4. **BM−**: Binocular illusory, monocular measurement, negative gains

In the tables, figures, and text below we refer to these conditions by compact three-character codes. The first letter indicates the illusory-display viewing mode (**B** = binocular, **M** = monocular), the second letter indicates the measurement-display viewing mode (**B** = binocular, **M** = monocular), and the trailing sign indicates the polarity of the measurement-display gain (+ = positive, − = negative).

### Psychometric Fits and PSE Comparisons

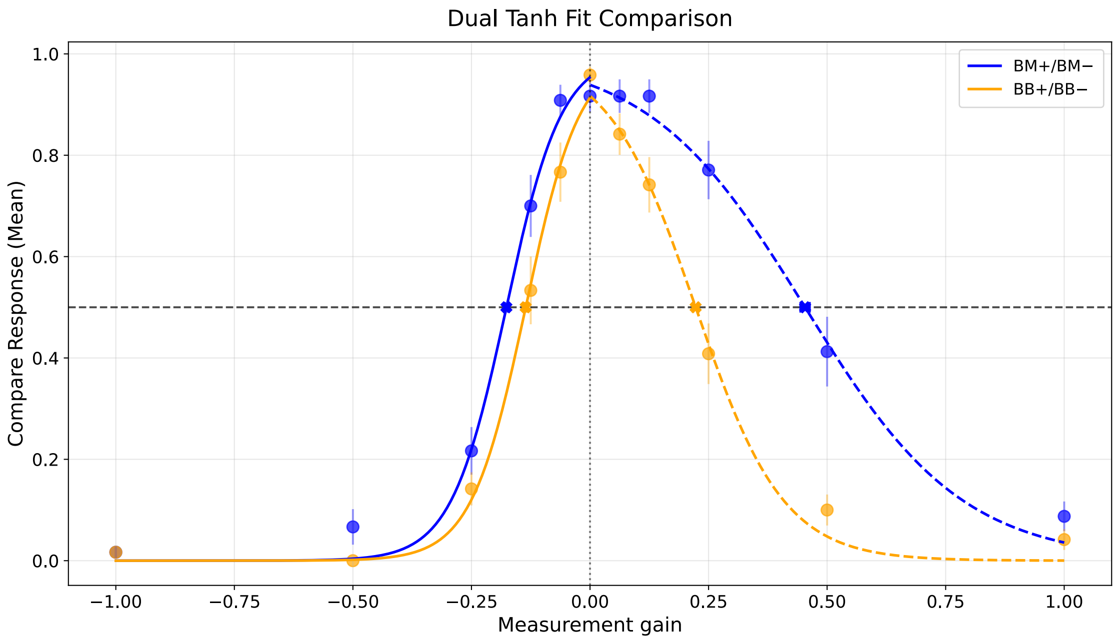


Figure 5.2: Psychometric comparison for the four Set A conditions in Experiment 1, where the illusory display was viewed binocularly. The fitted curves show the response pattern across positive and negative measurement-display gains for monocular and binocular measurement-viewing conditions.

Each condition yielded a separate fit, providing the corresponding PSE (i.e., the fitted offset). Statistical comparisons of these PSEs (Figure 5.3; Tables 5.1) revealed:

- **Positive gains:** A significant difference was observed between the monocular and binocular measurement-display conditions ($t(19) = 5.93$, $p < 0.0001$, Cohen's $d = 1.33$). The monocular measurement-display condition produced a larger positive PSE ($M = 0.471$, SEM $= 0.044$) than the binocular measurement-display condition ($M = 0.243$, SEM $= 0.023$), consistent with the interpretation that positive parallax gains are strongly modulated by whether disparity is available in the measurement display.
- **Negative gains:** A significant difference was also observed between the monocular and binocular measurement-display conditions ($t(19) = -3.05$, $p = 0.007$, Cohen's $d = -0.68$). The monocular measurement-display condition yielded a more negative PSE ($M = -0.188$, SEM $= 0.018$) than the binocular measurement-display condition ($M = -0.137$, SEM $= 0.014$), indicating that reversed parallax motion also differed

between the two viewing modes. Since the Shapiro-Wilk test indicated non-normality ($p < 0.001$), a Wilcoxon signed-rank test was also conducted and confirmed the result ($p < 0.001$).

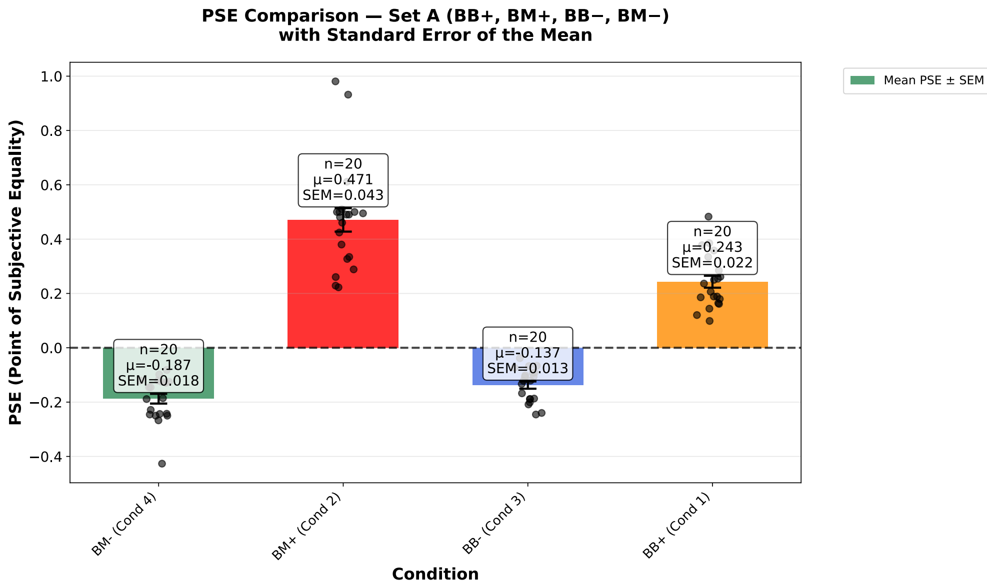


Figure 5.3: Group-level comparison of illusory motion parallax perception across Set A conditions ($N = 20$). Points of Subjective Equality (PSEs ± SEM) derived from the fitted psychometric functions for all four viewing conditions. Both the positive-gain and the negative-gain paired differences between monocular and binocular measurement-display conditions were statistically significant, with the monocular condition yielding larger absolute PSEs in both cases.

Table 5.1 summarizes the paired comparisons between the monocular and binocular measurement-display conditions, and Table 5.2 shows that each of the four binocular-illusory Set A conditions differed reliably from zero, confirming that the illusion produced measurable non-zero PSEs in every condition regardless of the measurement-display viewing mode.

Table 5.1: Paired comparisons of fitted PSEs between monocular and binocular measurement-display conditions in Experiment 1. Positive-gain and negative-gain contrasts are shown separately for the binocular-illusory Set A conditions, using the code labels (BB+, BM+, BB−, BM−) defined above. Wilcoxon $p$ is reported when the Shapiro-Wilk test indicates non-normal paired differences.

| **Contrast** | **N** | **Mean A** | **Mean B** | **Mean Diff** | **SD Diff** | **Cohen's d** | **t** | **df** | **p ttest** | **Shapiro p** | **Wilcoxon p** | **Sig[1]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BM+ vs BB+ | 20 | 0.4709 | 0.2432 | 0.2277 | 0.1717 | 1.3264 | 5.932 | 19 | < 0.0001 | 0.262 | | *** |
| BM− vs BB− | 20 | −0.1875 | −0.1373 | −0.0502 | 0.0736 | −0.6818 | −3.049 | 19 | 0.007 | < 0.001 | < 0.001 | ** |

Table 5.2: One-sample tests of fitted PSEs against zero for the four binocular-illusory Set A conditions in Experiment 1. All four conditions differed significantly from zero, indicating reliable non-zero PSEs when the illusory display retained stereoscopic disparity. Conditions use the Set A code labels defined above (BB+, BM+, BB−, BM−). Wilcoxon $p$ is reported when the Shapiro-Wilk test indicates non-normal PSE distributions.

| **Condition** | **N** | **Mean** | **SD** | **SEM** | **t** | **df** | **p ttest** | **Shapiro p** | **Wilcoxon p** | **Sig** |
|---|---|---|---|---|---|---|---|---|---|---|
| BB+ | 20 | 0.2432 | 0.1007 | 0.0225 | 10.795 | 19 | < 0.000001 | 0.2431 | | *** |
| BM+ | 20 | 0.4709 | 0.1981 | 0.0443 | 10.631 | 19 | < 0.000001 | 0.0035 | < 0.0001 | *** |
| BB− | 20 | −0.1373 | 0.0615 | 0.0137 | −9.988 | 19 | < 0.000001 | 0.2795 | | *** |
| BM− | 20 | −0.1875 | 0.0824 | 0.0184 | −10.179 | 19 | < 0.000001 | 0.0114 | < 0.0001 | *** |

### Reference Condition Selection

Under the working assumption that the illusory motion reflects the discounting of predicted parallax in the opposite direction, negative parallax conditions provide a useful reference for further analysis. Given that negative binocular and monocular conditions remained relatively close in absolute PSE magnitude, we selected the **negative binocular condition (BB−)** as the most appropriate reference for further analysis (This choice reflects a working assumption that is further examined in Experiment 2.). In the coding introduced above, this corresponds to retaining the standard stereo (binocular) illusory display together with binocular measurement-display viewing, paired with negative measurement gains. This choice reflects both the ecological relevance of binocular viewing and the practical advantage of using binocular stimuli over monocular occlusion in participant testing.

[1]Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

### 5.1.3 Set B - Monocular Illusory Display (testing if illusion requires disparity):

Set B follows directly from the same logic as Set A: after analyzing the standard binocular-illusory case in detail, we repeated the same four comparisons with the illusory display itself shown monocularly. This isolates the contribution of binocular disparity in the illusory display and tests whether the illusion disappears when that cue is removed:

5. **MB+**: Monocular illusory, binocular measurement, positive gains

6. **MM+**: Monocular illusory, monocular measurement, positive gains

7. **MB−**: Monocular illusory, binocular measurement, negative gains

8. **MM−**: Monocular illusory, monocular measurement, negative gains

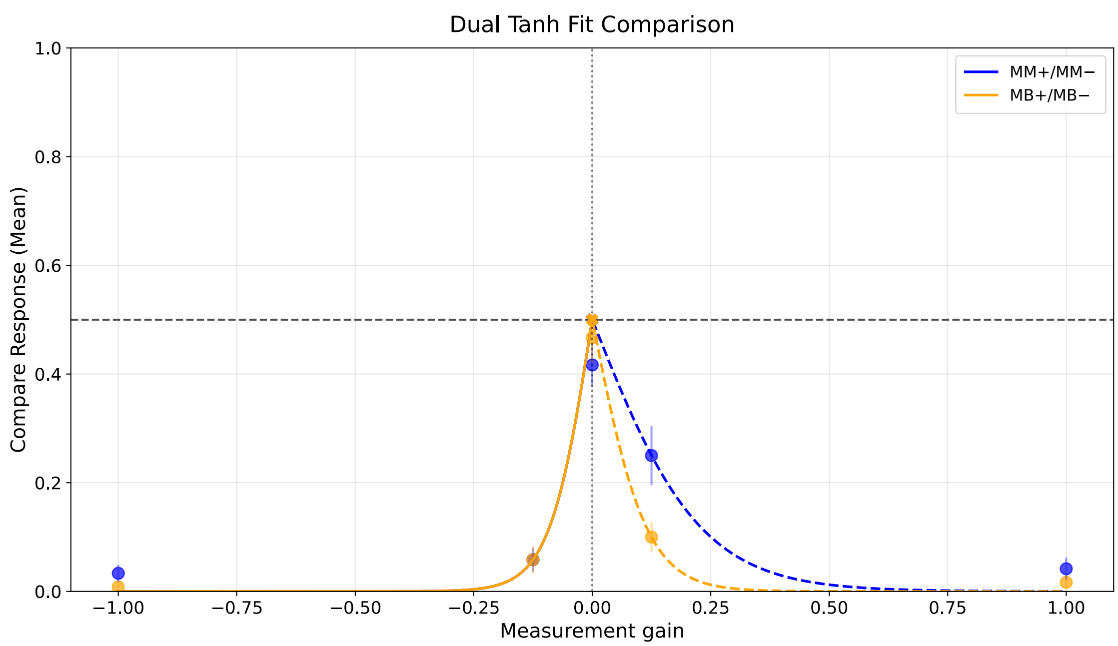


Figure 5.4: Psychometric comparison for the four Set B conditions in Experiment 1, where the illusory display was viewed monocularly. The fitted curves show the response pattern across positive and negative measurement-display gains for monocular and binocular measurement-viewing conditions.

As shown in Figure 5.4, participants never perceived the monocular illusory display as containing more motion than the measurement display. The average responses remained consistently below 0.5 across all gain types and viewing modes, confirming that binocular disparity is a necessary condition for the illusion to occur. The fitted PSEs for the four monocular-illusory conditions also clustered near zero (Figure 5.5): $M = -0.001$ for MM− ($N = 20$), $M = -0.002$ for MB− ($N = 20$), $M = 0.025$ for MM+ ($N = 20$), and $M = 0.007$ for MB+ ($N = 20$). These values were close to zero and substantially smaller than the corresponding Set A PSEs, consistent with the absence of the illusion when binocular disparity is removed. For a combined psychometric overview showing all eight conditions from Set A and Set B compared together in a single plot, see Figure A.3 in Appendix A.2.

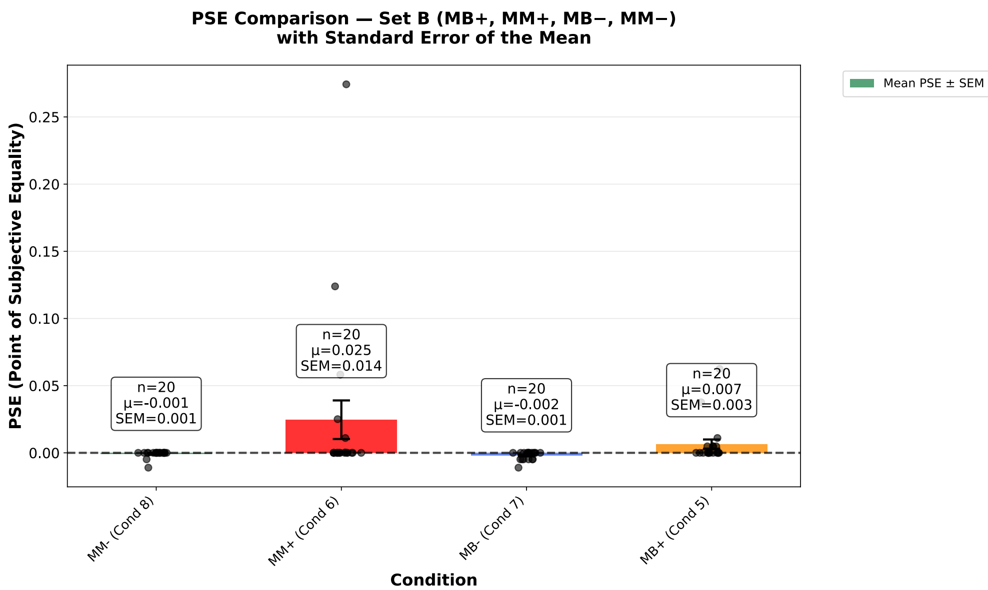


Figure 5.5: Group-level comparison of illusory motion parallax perception for Set B conditions ($N = 20$). Points of Subjective Equality (PSEs ± SEM) derived from the fitted psychometric functions for the four monocular-illusory conditions. The PSEs cluster near zero, confirming the absence of the illusion when stereoscopic depth is removed.

These analyses address Hypothesis 1 (see Section 3.2.1), concerning the necessity of binocular disparity for the perception of illusory parallactic motion.

### 5.1.4 Forward Model Interpretation

Building on the framework presented in Figure 1.1, we interpret these results through the lens of forward model mechanisms [Miall and Wolpert, 1996, Wolpert and Kawato, 1998]. In binocular viewing of illusory motion parallax display, the brain predicted motion parallax based on head movement and stereoscopic depth. The absence of actual parallax created a mismatch between predicted and received sensory input, which was experienced as illusory motion. In monocular viewing, the illusion disappeared, suggesting that no prediction error was generated. This contrast between monocular and binocular viewing supports the idea that forward models underlie illusory parallactic motion.

The observed asymmetry, as shown in Figure 5.1, between positive and negative parallax gains further reinforces this interpretation: forward models are calibrated for natural (positive) parallax experiences, but they struggle with unnatural negative parallax, leading to stronger perceived motion in those conditions. Together, these preliminary results motivate the forward-model interpretation developed in this thesis.

## 5.2 Experiment 2: Direction of Illusory Motion Parallax

### 5.2.1 Design and Descriptive Results

These analyses address Hypothesis 2 (see Section 3.3.1), concerning the predicted direction of illusory parallactic motion relative to real motion parallax.

Across 36 trials (N = 41), each participant judged whether the illusory motion in the stereo display moved in the same or opposite direction as the parallax in the measurement display. Whether the stereo-only illusory parallax display appeared in the observer's left or right frame was counterbalanced, so that any bias toward one side of the dual-frame layout could not masquerade as a directional effect. Six gain levels ($\{-1, -0.5, -0.1, 0.1, 0.5, 1\}$) were used to present measurement-display parallax in both directions.

Figure 5.6 shows the group-average proportion of "same direction" responses as a function of measurement-display gain. The transition between the negative and positive regimes was sharp and centered near gain = 0.

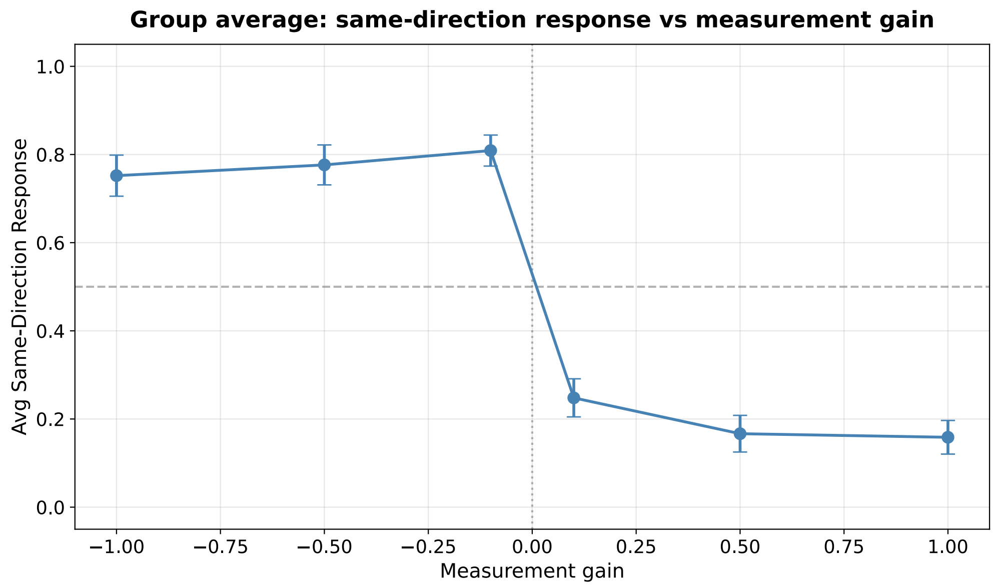


Figure 5.6: Group-average proportion of "same direction" responses as a function of measurement-display gain (N=41). Error bars indicate ±1 SEM. The dashed line marks chance level (0.5). Negative gains produced high same-direction judgments, while positive gains produced low same-direction judgments, indicating that the illusory motion aligns with the direction of negative parallax. For negative gains, the mean proportions were $M = 0.752$ (gain $= -1$), $M = 0.776$ (gain $= -0.5$), and $M = 0.809$ (gain $= -0.1$). For positive gains, the mean proportions were $M = 0.248$ (gain $= 0.1$), $M = 0.167$ (gain $= 0.5$), and $M = 0.159$ (gain $= 1$), showing a sharp transition centered near gain $= 0$.

Supplementary participant-level visualizations, including the full per-participant panel figure and consistency summary, are provided in Appendix A.3.

### 5.2.2 Collapsed Negative Versus Positive Gains

To provide a single summary statistic, we collapsed responses across all negative gain levels and across all positive gain levels for each participant. The mean same-direction proportion was substantially higher for negative gains ($M = 0.779$) than for positive gains ($M = 0.191$). This large separation confirms Hypothesis 2 (see Section 3.3.1): the illusory motion is opposite in direction to the parallax that would be expected from natural head-coupled viewing of a scene at the stereoscopically specified depth. In other words, the illusory motion behaves as though it were driven by *negative* parallax.

Table 5.3: Collapsed comparison of the proportion of "same direction" responses for negative versus positive gains (N=41). The table summarizes the paired $t$-test and Wilcoxon signed-rank analyses for the participant-wise collapsed negative-versus-positive gain comparison.

| **Metric** | **Value** |
|---|---|
| N | 41 |
| Mean (negative gains) | 0.7791 |
| Mean (positive gains) | 0.1911 |
| Mean difference (pos-neg) | -0.5881 |
| Shapiro-Wilk W | 0.8768 |
| Shapiro-Wilk p | 0.0004 |
| t-statistic | 9.5380 |
| df | 40 |
| p-value | $< 0.000001$ |
| Cohen's d | -1.4896 |
| Wilcoxon W | 24.5000 |
| Wilcoxon p | $< 0.000001$ |
| Result ($\alpha = 0.05$) | *** |

Normality of the paired differences was violated (Shapiro-Wilk $W = 0.8768$, $p = 0.0004$). A paired t-test on the collapsed means nevertheless confirmed a strong difference between negative and positive gains, $t(40) = 9.538$, $p < 0.000001$, Cohen's $d = -1.49$. A Wilcoxon signed-rank test corroborated this result, $W = 24.5$, $p < 0.000001$ (Table 5.3).

A participant-level consistency summary is provided with the other Experiment 2 participant-level plots in Appendix A.3.

### 5.2.3 Forward Model Interpretation

Within the forward-model framework, the brain predicts positive (natural) motion parallax based on the stereoscopically perceived depth. Because the illusory motion parallax display provides no actual parallax, the prediction error, the difference between expected and received sensory input, has the opposite sign to the predicted parallax. The perceptual system interprets this signed prediction error as scene motion in the negative-parallax direction. The alignment of participants' direction judgments with negative parallax (Figure 5.6; see also Appendix A.3) therefore provides direct evidence that the illusion arises from a signed prediction error generated by the forward model, rather than from an unsigned mismatch or a non-directional perceptual bias.

This directional finding also retroactively justifies the use of the negative binocular condition (BB−) as the reference in Experiment 1 (Section 5.1.2), since the illusory motion and negative measurement-display parallax share the same perceptual direction. Although the Experiment 1 sample shows some viewing-mode modulation even within the negative-gain comparison, the negative binocular condition remains the most practical reference because it preserves ecological binocular viewing while probing the same parallax direction that matches the illusory percept. For ease of testing, later experiments therefore used the same configuration: the illusory display was kept as a standard stereo (binocular) display and the measurement display was also viewed binocularly, so that participants did not need to wear an eye patch or alternate occlusion, while still probing the same negative-parallax regime that matches the illusory motion in direction. In the compact-code scheme introduced in Section 5.1.2, Experiments 3 and 4 therefore correspond to the **BB−** cell, holding both the illusory and measurement viewing modes fixed at **B** (binocular) and using only negative measurement gains. This motivates the design choice in Experiment 3, where only negative gains are used in the measurement display, ensuring that the measurement parallax and the illusory motion are always in the same perceptual direction and can be compared on the basis of magnitude alone.

# 5.3 Experiment 3: Scene Context Manipulation

## 5.3.1 Design and Descriptive Results

This experiment examined how scene context affects the strength of the illusory motion. Of the 24 recruited participants, 6 were excluded prior to analysis due to incomplete sessions or excessively noisy response patterns, leaving **18 participants** in the final sample. Each participant completed **144 planned trials** (3 scene layouts $\times$ 2 frame conditions $\times$ 4 gains $\times$ 6 repetitions). Before analysis, the first **2 trials** of each session were discarded as warm-up, and a per-participant response-time cap at **median** $+3\times$ **MAD** was applied. The planned design combined one of three scene layouts (**closed environment**, **open environment**, **no environment**), with the border frame either visible or hidden around the illusory display, and one of four negative gain values ($\{-0.0625, -0.125, -0.25, -0.5\}$). Trials were two-interval comparisons between the stereo (illusory motion parallax) and parallax (measurement) displays. For each layout $\times$ frame $\times$ gain cell, the design included **six repetitions** of the two-interval comparison: three with the illusory motion parallax display first and three with the measurement display first, balancing presentation order.

### 5.3.2 Effect of Scene Context

Figure 5.7 shows the **group fitted psychometric curves** for the three scene contexts on a single plot. Across all three contexts, the probability of judging the stereo interval as stronger increased as the negative gain in the measurement display approached zero. The fitted PSEs (marked by crosses) were ordered systematically: the **closed-environment** condition had the least negative PSE, the **open-environment** condition was more negative, and the **no-environment** condition was most negative. This leftward shift indicates that the illusion was strongest in the no-environment condition and weakest in the closed-environment condition.

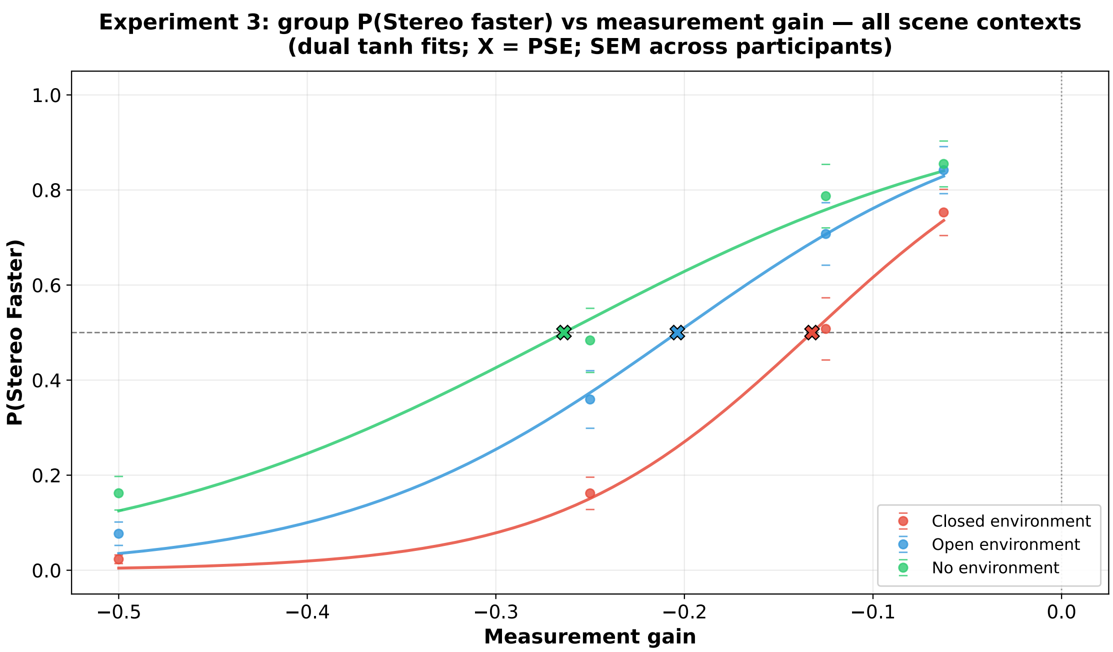


Figure 5.7: Fitted group psychometric curves for Experiment 3, shown separately for the three scene contexts (closed environment, open environment, no environment). Curves plot the probability that the stereo interval was judged stronger than the measurement interval as a function of negative gain. Crosses mark the fitted PSEs, showing that the no-environment condition is shifted furthest left and the closed-environment condition furthest right.

The same fitted psychometric analysis at the **participant level** is provided in Appendix A.4. Individual observers were variable, but the dominant pattern remained visible across the panel set: most participants placed the **closed-environment** condition closer to zero than the **open-environment** and **no-environment** conditions.

To summarise the context effect quantitatively, we fitted psychometric functions for each participant and scene context and extracted both the gain at which the stereo and parallax displays were judged equally strong (PSE) and the corresponding fitted slope. Table 5.4 shows the resulting summary statistics, and Figures 5.8 and 5.9 provide compact visual summaries of the same fitted parameters. Valid PSE estimates were available for **18 participants** in each scene context. Mean PSEs were least negative for **closed environment** ($M = -0.131$, SEM = 0.015), more negative for **open environment** ($M = -0.204$, SEM = 0.021), and most negative for **no environment** ($M = -0.267$, SEM = 0.031). Because more negative PSE values indicate that the illusory motion remained stronger over a broader range of negative measurement gains, this ordering suggests that the illusion was weakest in the most structured scene and strongest when surrounding context was removed. Mean fitted slopes were steeper in the **closed-environment** and **open-environment** conditions than in the **no-environment** condition, although slope variability was substantial.

Table 5.4: Context-level summary of fitted PSEs and psychometric slopes in Experiment 3. PSE is the gain at which the stereo and measurement displays were judged equally strong. More negative PSEs indicate a stronger illusory-motion percept relative to the measurement display.

| **Scene Context** | **N** | **Mean PSE** | **PSE SEM** | **Mean Slope** | **Slope SEM** |
|---|---|---|---|---|---|
| Closed environment | 18 | −0.131 | 0.015 | 13.772 | 4.029 |
| Open environment | 18 | −0.204 | 0.021 | 14.625 | 4.481 |
| No environment | 18 | −0.267 | 0.031 | 9.048 | 3.319 |

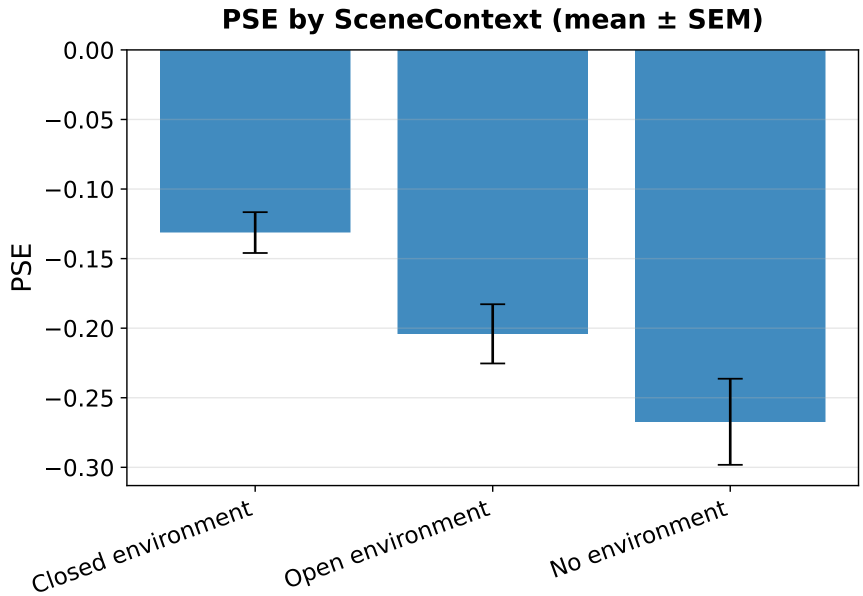


Figure 5.8: Mean fitted PSEs for the three Experiment 3 scene contexts. Error bars indicate ±1 SEM across participants with valid fits. More negative PSEs correspond to a stronger illusory-motion percept relative to the measurement display, with the no-environment condition showing the strongest effect and the closed-environment condition the weakest.

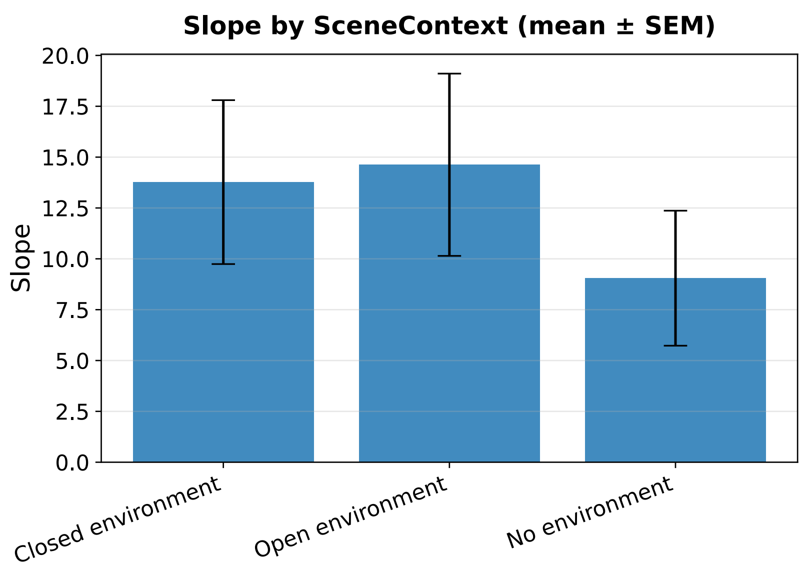


Figure 5.9: Mean fitted psychometric slopes for the three Experiment 3 scene contexts. Error bars indicate ±1 SEM. Average slopes were steeper for the closed-environment and open-environment conditions than for the no-environment condition, with substantial variability across participants.

At the participant level, the dominant direction remained consistent with the group-level summaries: 15 of 18 participants had a more negative PSE in the **no-environment** condition than in the **closed-environment** condition, and 16 of 18 had a more negative PSE in the **open-environment** condition than in the **closed-environment** condition. A histogram of individual context effects (Figure 5.10) visualizes this distribution.

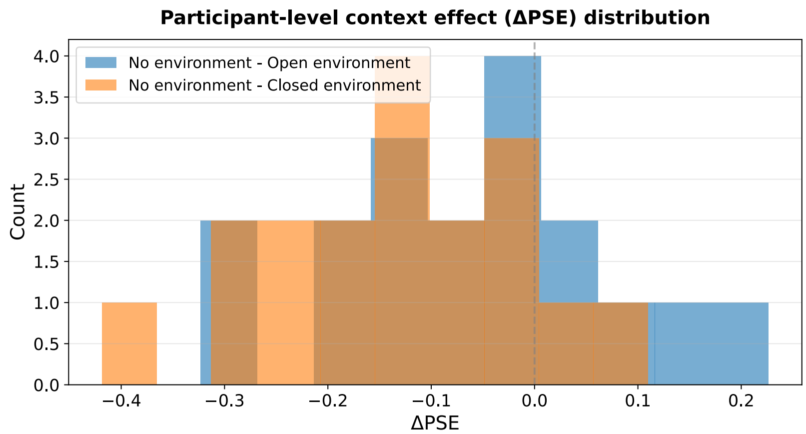


Figure 5.10: Distribution of participant-wise context effects for Experiment 3, expressed as the difference in PSE between no environment and the closed environment. Values below zero indicate a stronger illusion in the no-environment condition. The distribution is shifted in the observed direction while still showing participant-level variability.

### 5.3.3 Effect of Frame Visibility

We next examined the descriptive effect of the explicit border frame within each environment. Figure 5.11 shows the full environment × frame summary: the means shift clearly across environment levels, whereas the frame-visible and frame-hidden bars remain close within each environment.

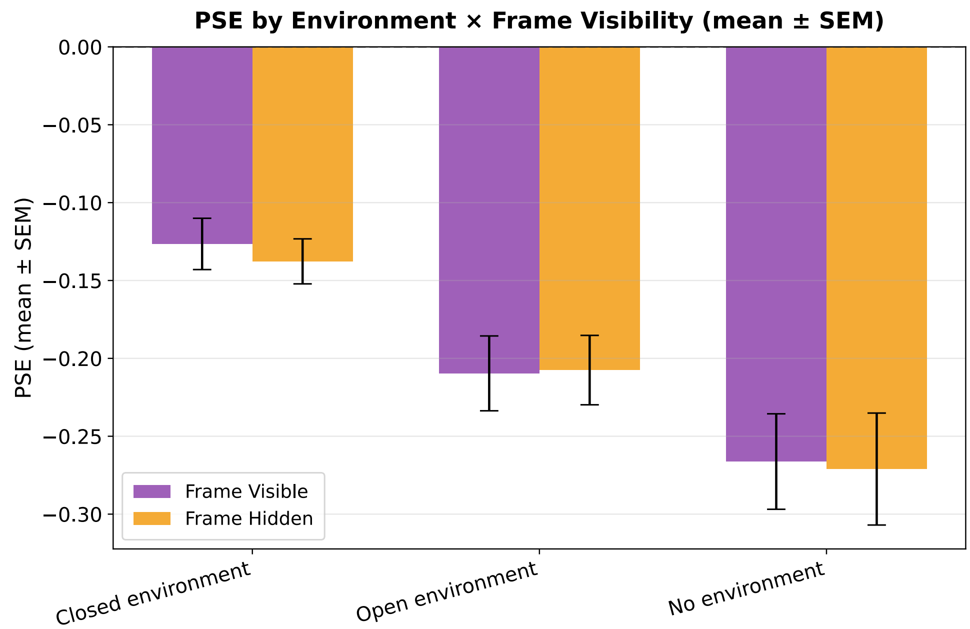


Figure 5.11: Mean fitted PSEs for Experiment 3 by environment and frame visibility. Error bars indicate ±1 SEM across participants. The dominant shift is across environment levels, with PSEs becoming more negative from closed environment to open environment to no environment, while frame-visible and frame-hidden means remain similar within each environment.

The same pattern is visible in the fitted psychometric curves. Figure 5.12 shows frame-visible and frame-hidden trials separately within each scene context; the two frame conditions remain relatively close within a given scene, especially compared with the larger between-scene shifts reported above.

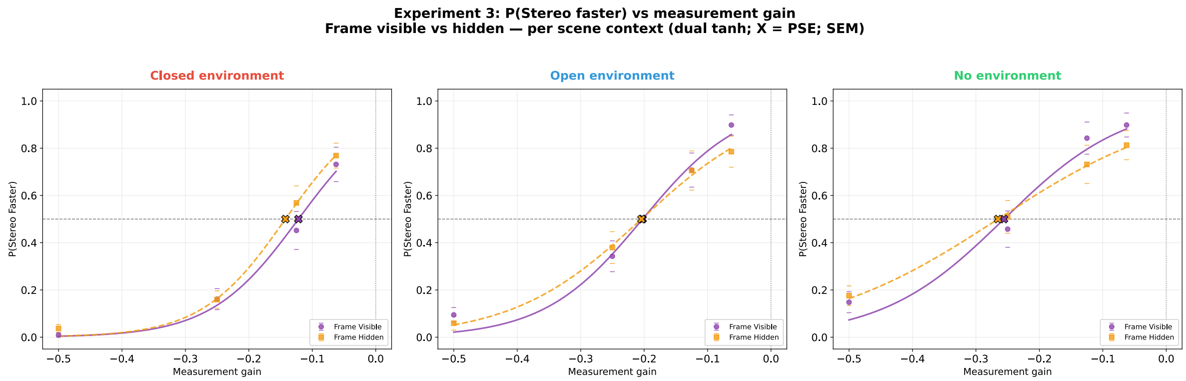


Figure 5.12: Fitted psychometric curves for Experiment 3, split by both scene context and frame visibility. Within each scene, the frame-visible and frame-hidden fits remain relatively similar, whereas the larger systematic shifts occur across scene contexts.

### 5.3.4 PSE Statistical Analysis

After summarizing the PSE patterns for both scene context and frame visibility, we tested the participant-wise PSE estimates with a two-way repeated-measures ANOVA, using environment and frame visibility as within-participant factors. This overall test asks whether the three scene layouts, the two frame-visibility states, or their interaction produced reliable shifts in PSE. As shown in Table 5.5, the main effect of environment was significant, $F(2, 34) = 13.375$, $p < 0.001$, partial $\eta^2 = 0.440$. By contrast, the main effect of frame visibility was not significant, $F(1, 17) = 0.133$, $p = 0.720$, and the environment-by-frame interaction was also not significant, $F(2, 34) = 0.126$, $p = 0.882$. Thus, the ANOVA identified environment as the reliable source of the PSE shift, while providing no evidence for a simple frame-visibility effect or a frame-dependent change in the environment effect.

Table 5.5: Two-way repeated-measures ANOVA on Experiment 3 PSE values by environment and frame visibility. The environment factor produced a reliable PSE shift, whereas the simple frame-visibility effect and the environment-by-frame interaction were not significant.

| **Source** | **F** | **Num_DF** | **Den_DF** | **p** | **partial_eta_sq** | **Sig** | **N_complete** |
|---|---|---|---|---|---|---|---|
| Environment | 13.375 | 2 | 34.0 | $< 0.001$ | 0.440 | *** | 18 |
| Frame | 0.133 | 1 | 17.0 | 0.720 | 0.008 | ns | 18 |
| Environment × Frame | 0.126 | 2 | 34.0 | 0.882 | 0.007 | ns | 18 |

Because the repeated-measures ANOVA showed a reliable effect of environment, follow-up tests were restricted to the environment factor. We used paired-samples $t$-tests on participant-wise environment marginal PSEs collapsed across frame visibility, with Bonferroni correction applied across the three environment comparisons. The largest corrected contrast was between **no environment** and the **closed environment** (mean difference = $-0.137$, Cohen's $d = -1.077$, $t(17) = -4.570$, Bonferroni-corrected $p < 0.001$), followed by **open environment** versus **closed environment** (mean difference = $-0.076$, Cohen's $d = -1.111$, $t(17) = -4.715$, Bonferroni-corrected $p < 0.001$). The **no environment** versus **open environment** contrast was smaller and was not significant after correction (mean difference = $-0.060$, Cohen's $d = -0.461$, $t(17) = -1.954$, Bonferroni-corrected $p = 0.202$). These contrasts therefore indicate that removing or opening the surrounding scene structure shifted PSEs more negative relative to the closed environment, while the additional difference between open environment and no environment was weaker. Table 5.6 summarizes the same Bonferroni-corrected follow-up tests; all contrasts include the **18 participants** with complete environment-by-frame PSE estimates. Because the ANOVA did not show a significant main effect of frame visibility or environment-by-frame interaction, we did not conduct additional frame follow-up tests.

Table 5.6: Bonferroni-corrected pairwise differences in PSE between scene layouts in Experiment 3. Follow-up paired-samples $t$-tests were performed only for the significant environment main effect from the repeated-measures ANOVA, using participant-wise environment marginal PSEs collapsed across frame visibility. Negative mean differences indicate more negative PSEs, and therefore stronger illusory motion relative to the measurement display, in the first-listed condition.

| **Contrast** | **N** | **Mean Diff in PSE** | **Cohen's $d$** | $t(17)$ | **Raw** $p$ | **Bonferroni** $p$ | **Sig** |
|---|---|---|---|---|---|---|---|
| No environment - Open environment | 18 | $-0.060$ | $-0.461$ | $-1.954$ | 0.067 | 0.202 | ns |
| No environment - Closed environment | 18 | $-0.137$ | $-1.077$ | $-4.570$ | $< 0.001$ | $< 0.001$ | *** |
| Open environment - Closed environment | 18 | $-0.076$ | $-1.111$ | $-4.715$ | $< 0.001$ | $< 0.001$ | *** |

Together, these results show that **the absence of a surrounding environment makes the illusory motion appear stronger**: the stereo effect was strongest in the no-environment condition, weaker in the open environment, and weakest in the closed environment in the PSE summaries, with significant participant-wise contrasts for both no environment versus closed environment and open environment versus closed environment. The explicit frame border did not produce reliable simple within-scene PSE differences.

### 5.3.5 Forward Model Interpretation

Experiment 3 held the direction of measurement parallax fixed (negative, matching Experiment 2) and varied only the scene surrounding the illusory display. Within the forward-model framework (Figure 1.1; Miall and Wolpert 1996, Wolpert and Kawato 1998), the core mechanism is unchanged: stereoscopic depth specifies structure, head movement generates an expectation of motion parallax, and the absence of matching parallax yields a prediction error that is experienced as illusory motion. What changes across contexts is not the sign of that mismatch but its *effective strength* in the comparison task, as summarized by the fitted curves in Figure 5.7 and the participant-wise PSE and slope summaries in Tables 5.4 and 5.6.

We interpret the ordering of contexts as a modulation of how strongly the visual system treats the stimulus as a stable, immersive three-dimensional layout, along the lines of *presence* versus more "picture-like" or locally isolated interpretations discussed in relation to VR [Troje, 2019]. In the **no-environment** condition, the scene offers little global layout; disparity-defined depth may dominate the forward model's estimate of expected parallax for the judged object, so the absence of physical parallax produces a large, relatively unambiguous prediction error and the illusory motion is judged stronger over a wide range of measurement gains. In the **closed-environment** and **open-environment** conditions, additional depth cues specify nearby references (room and table surfaces) and farther references (distant background structure). Those cues may stabilize perceived layout, redistribute attention across surfaces, or change how prediction error is attributed (e.g., motion of the target relative to a stable room vs. a freestanding object), thereby *reducing* the judged strength of the illusory motion compared with the same measurement parallax, consistent with PSEs closer to zero and steeper group psychometric slopes in the closed-environment case.

The relatively small simple PSE effects of frame visibility are compatible with this account: drawing or hiding a border around the same scene content leaves the major environmental structure and depth relations largely unchanged, and is therefore unlikely to materially alter the predicted parallax signal or the global priors that differentiate an isolated cube from a furnished room. In other words, context effects in Experiment 3 plausibly reflect how surrounding scene structure weights forward-model predictions and the resulting prediction error, while the explicit frame border appears to have a more limited role.

This contextual modulation complements Experiments 1 and 2: binocular disparity remains necessary for the illusion, and illusory motion remains aligned with negative parallax in direction; Experiment 3 shows that, given those constraints, the *magnitude* of the phenomenon is further shaped by whether the observer has close references, farther references, or no surrounding references in the virtual environment.

### 5.3.6 Summary Interpretation

Overall, Experiment 3 rejects the contextual prediction in Hypothesis 3 (Section 3.4.1). The hypothesis predicted that adding closer or more numerous visual references to the scene would strengthen the illusion, with nearby references in the closed environment expected to produce the strongest effect, farther references in the open environment an intermediate effect, and no surrounding reference the weakest effect. Instead, the Experiment 3 PSE summaries and repeated-measures ANOVA showed that environment was a reliable modulator in the opposite direction: the illusion was strongest when surrounding references were removed and weakest in the closed environment. Thus, the results support the broader claim that scene context matters, but they do not support the specific prediction that closer or more numerous references increase illusion strength.

This reversal of the directional prediction is interpreted within an attribution-based account in Chapter 6, where we argue that surrounding references shape the illusion not by acting as a more sensitive parallax ruler but by changing how the visual system explains the cue mismatch.

## 5.4 Experiment 4: Viewing Angle Manipulation

### 5.4.1 Design and Descriptive Results

Experiment 4 examined whether the perceived strength of illusory parallactic motion varies as a function of **viewing angle**. The angle manipulation was implemented by changing the vertical placement of the AlbertiFrame and its coupled scene geometry; the eight reported angles are therefore the angular offsets corresponding to the feasible vertical frame positions within the limits of the scene geometry. Both the illusory motion parallax display and the measurement display were presented at one of these eight angular offsets (illustrated schematically in Figure 4.8): 2.83°, 13.14°, 22.90°, 31.60°, 38.80°, 44.84°, 50.00°, and 54.23°. The measurement display used the same negative binocular gain set as the earlier magnitude-comparison experiments, $\{-0.0625, -0.125, -0.25, -0.5\}$, with six repetitions for each combination of angle and gain.

Figure 5.13 shows the group-average psychometric curves for the probability that the illusory (stereo) interval was judged stronger than the measurement interval as a function of measurement-display gain, separately for each viewing angle. Across all angles, the same broad psychometric pattern was observed as in earlier experiments: as the negative gain approached zero, participants increasingly judged the illusory interval as stronger. Crucially, the fitted curves are shifted as a function of angle. At steeper viewing angles, the probability of choosing the illusory interval as stronger rises earlier, and the fitted PSE moves leftward to more negative gains, indicating that the illusory motion remains dominant even when compared against relatively stronger negative parallax in the measurement display.

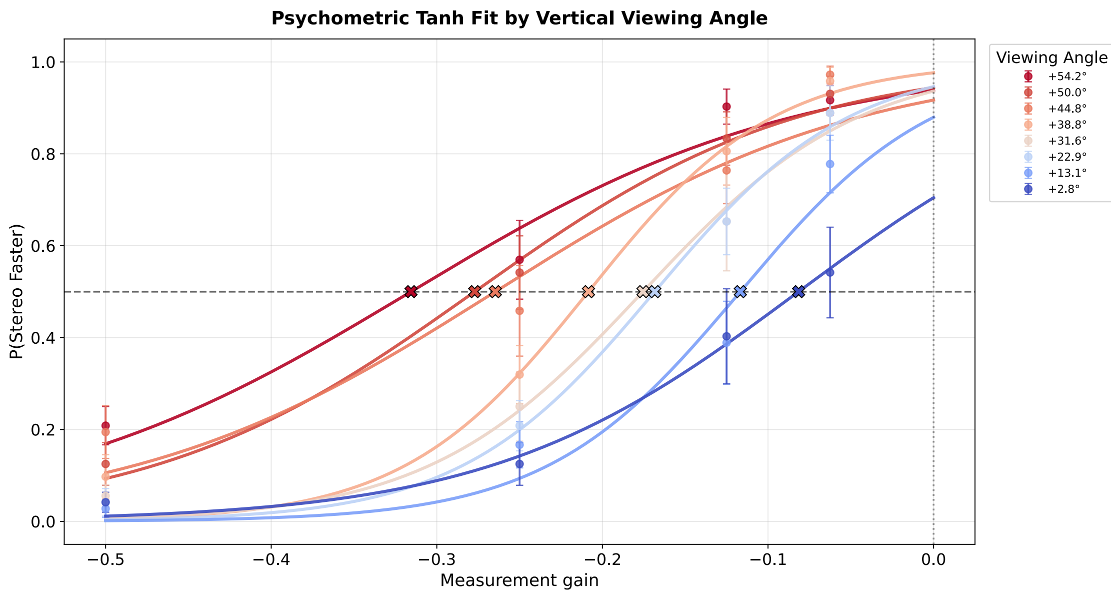


Figure 5.13: Tanh fits for Experiment 4, showing the group-average probability that the illusory (stereo) interval was judged stronger than the measurement interval as a function of measurement-display gain for each viewing-angle condition. The eight conditions correspond to viewing angles from 2.83° to 54.23°. Crosses mark the fitted PSEs. The systematic leftward displacement of the curves at steeper viewing angles indicates earlier and stronger dominance of the illusory interval under those conditions.

Figure 5.14 summarizes the same effect using the fitted PSE for each condition. When ordered by viewing angle, the shallowest conditions produced the least negative PSEs (2.83°: −0.082; 13.14°: −0.117), whereas the steepest conditions produced the most negative values (50.00°: −0.277; 54.23°: −0.316). The intermediate angles followed the same ordered shift (22.90°: −0.168; 31.60°: −0.175; 38.80°: −0.208; 44.84°: −0.265). This monotonic pattern shows that **higher viewing angles are associated with a stronger illusory-motion**

**percept relative to the matched measurement display**. Put differently, as the display is viewed from steeper angles, the illusory interval remains stronger over a broader range of negative measurement gains.

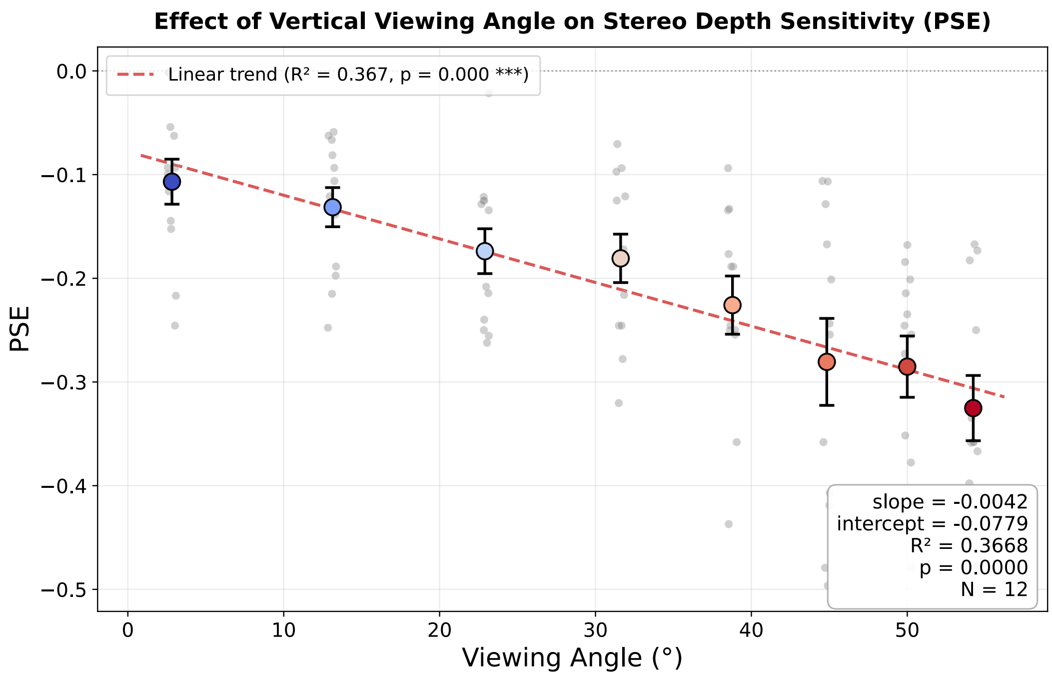


Figure 5.14: Experiment 4 PSE as a function of viewing angle. Points show condition means and error bars indicate uncertainty across participants; grey points show participant-level values. More negative PSE values indicate that the illusory interval remained stronger relative to the measurement interval over a broader range of negative gains. The ordered shift in PSEs indicates that steeper viewing angles were associated with stronger illusory-motion judgments overall.

Figure 5.14 makes the overall angle effect more explicit. The fitted PSE values show a consistent trend: as viewing angle increases, PSE becomes more negative. This means that participants required a stronger negative-parallax measurement display to match the apparent motion of the illusory interval at steeper viewing angles, indicating that the illusory motion remained perceptually stronger over a broader range of gains. Thus, the main effect of Experiment 4 is not simply a difference between isolated angle conditions, but an overall increase in illusory-motion strength as the observer's viewing angle becomes steeper.

### 5.4.2 PSE Trend Analysis

This trend was confirmed statistically in multiple complementary analyses. A linear mixed-effects model predicting participant-wise PSE from viewing angle showed a strong negative fixed effect of angle ($\beta = -0.00421$, $z = -11.95$, $p < 0.000001$), indicating that each increase in viewing angle was associated with a more negative PSE and therefore a stronger illusory-motion percept (Table 5.7).

Table 5.7: Linear mixed-effects analysis of Experiment 4 PSE values as a function of viewing angle. The fitted negative slope indicates that increasing viewing angle predicts progressively more negative PSEs, consistent with stronger illusory-motion judgments at steeper angles.

| **Term** | **Coef** | **SE** | **z/t** | **p** | **CI_low** | **CI_high** | **Sig** |
|---|---|---|---|---|---|---|---|
| Intercept | -0.077883 | 0.025747 | -3.0250 | 0.002487 | -0.128346 | -0.027421 | ** |
| Angle | -0.004206 | 0.000352 | -11.9539 | < 0.000001 | -0.004896 | -0.003517 | *** |

The effect was accompanied by substantial individual variability, but many observers nonetheless exhibited the same broad tendency toward stronger illusory-motion judgments and more negative PSEs at steeper viewing angles.

### 5.4.3 Summary Interpretation

This pattern is consistent with the interpretation that steeper viewing angles reduce the effective contribution of far-background motion parallax in the measurement display. If background structure contributes less usable parallax information at larger angles, the real motion in the measurement interval becomes less salient relative to the stereo-driven illusion, making the two more difficult to distinguish. On this account, the viewing-angle effect in Experiment 4 is compatible with the scene-context effect observed in Experiment 3: when broader environmental structure becomes less informative for the comparison, the illusory percept appears stronger. At the same time, we offer this explanation as a theoretically consistent interpretation of the present pattern rather than as a uniquely established mechanism.

## 5.5 Experiment 5: Observer Distance Manipulation

### 5.5.1 Design and Descriptive Results

Experiment 5 examined whether the perceived strength of illusory parallactic motion varies as a function of **observer distance from the frame**. The stereo-only illusory parallax display was compared against a parallax-only measurement display using only **positive gains** ({0.125, 0.25, 0.5, 1}), tested at three viewing distances: **0.4 m**, **0.8 m**, and **1.2 m**. As described in the Methodology, the physical size of the frame-and-scene arrangement was held constant across these distance conditions and no compensatory scaling was applied; therefore, changing observer distance also changed the visual angle subtended by the display. The parallax measurement display was shown under both **monocular** and **binocular** viewing. Nineteen participants completed the experiment; the PSE-based analyses below include **18 valid participant-level PSEs per distance-by-viewing-mode condition**.

Figure 5.15 summarizes the fitted PSEs for Experiment 5 across the three observer distances and the two parallax measurement-display viewing modes. The same distance pattern is visible in both viewing modes: PSEs are largest at the nearest distance (0.4 m), particularly in the monocular parallax condition, and move closer to zero at 1.2 m.

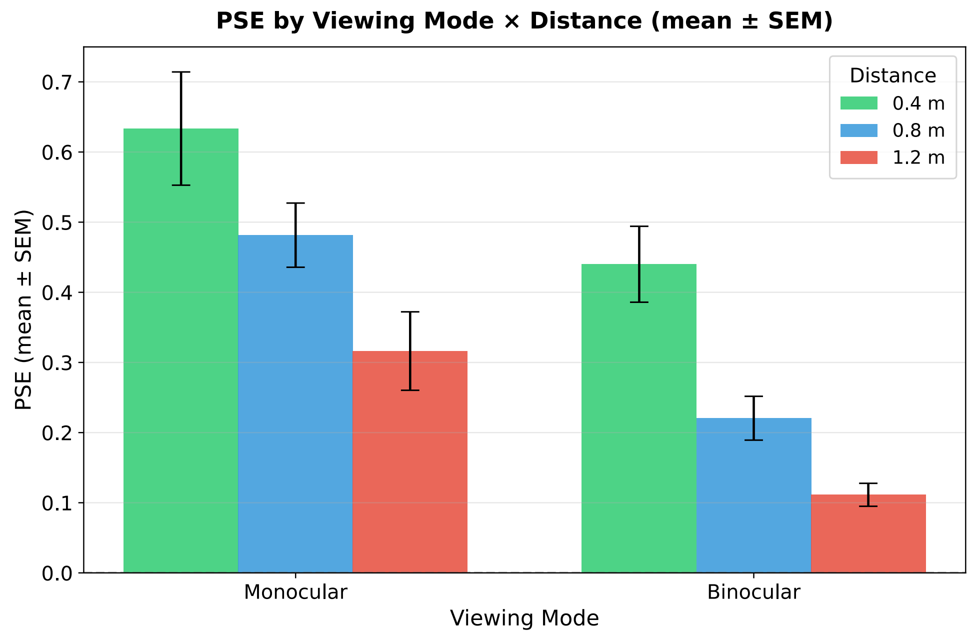


Figure 5.15: Mean fitted PSEs for Experiment 5 by measurement-display viewing mode and observer distance. Error bars indicate ±1 SEM across participants. PSEs decrease as observer distance increases, while monocular measurement viewing yields larger PSEs than binocular measurement viewing at each distance.

This distance-related shift is also shown in the fitted **Points of Subjective Equality (PSEs)** in Table 5.8. For the **monocular parallax measurement display**, mean PSE decreased from $M = 0.633$, SEM = 0.081 at 0.4 m to $M = 0.481$, SEM = 0.046 at 0.8 m and $M = 0.316$, SEM = 0.056 at 1.2 m. For the **binocular parallax measurement display**, the same pattern was present but at consistently smaller magnitudes: $M = 0.440$, SEM = 0.054 at 0.4 m, $M = 0.220$, SEM = 0.031 at 0.8 m, and $M = 0.111$, SEM = 0.016 at 1.2 m. Thus, in both viewing modes, **increasing observer distance from the frame reduced the gain required for the measurement display to match the illusory display**, indicating a weaker illusory-motion percept at larger distances.

Table 5.8: Summary of fitted PSEs for Experiment 5 across observer distance and parallax measurement-display viewing mode. Larger positive PSEs indicate that stronger positive parallax was required to match the apparent motion of the illusory display.

| Distance | Condition | N | Mean | SD | SEM | 95% CI |
|---|---|---|---|---|---|---|
| 0.4 m | BM+ | 18 | 0.633 | 0.342 | 0.081 | [0.475, 0.792] |
| 0.4 m | BB+ | 18 | 0.440 | 0.230 | 0.054 | [0.334, 0.546] |
| 0.8 m | BM+ | 18 | 0.481 | 0.194 | 0.046 | [0.392, 0.571] |
| 0.8 m | BB+ | 18 | 0.220 | 0.133 | 0.031 | [0.159, 0.282] |
| 1.2 m | BM+ | 18 | 0.316 | 0.237 | 0.056 | [0.207, 0.426] |
| 1.2 m | BB+ | 18 | 0.111 | 0.069 | 0.016 | [0.079, 0.143] |

### 5.5.2 Linear Mixed-Effects Model

A direct linear mixed-effects model provided inferential support for the distance trend. Participant-wise PSEs were modeled as a function of observer distance, measurement-display viewing mode, and their interaction, with participant included as a random intercept. In the monocular reference condition, PSE decreased significantly as distance increased ($\beta = -0.396$, $z = -5.31$, $p < 0.000001$). The binocular measurement-display condition also produced lower PSEs overall ($\beta = -0.208$, $z = -2.28$, $p = 0.022$), while the distance-by-viewing-mode interaction was not significant ($\beta = -0.014$, $z = -0.14$, $p = 0.892$). Thus, the distance effect was not merely descriptive: it was statistically reliable, and there was no evidence that the distance slope differed strongly between monocular and binocular measurement displays (Table 5.9).

Table 5.9: Linear mixed-effects analysis of Experiment 5 PSE values as a function of observer distance and measurement-display viewing mode. Distance was coded in meters, monocular measurement viewing served as the reference level, and participant was included as a random intercept.

| Term | Coef | SE | z | p | CI low | CI high | Sig |
|---|---|---|---|---|---|---|---|
| Intercept | 0.794 | 0.071 | 11.216 | $< 0.000001$ | 0.655 | 0.933 | *** |
| Binocular viewing | −0.208 | 0.091 | −2.283 | 0.022 | −0.387 | −0.029 | * |
| Distance | −0.396 | 0.075 | −5.311 | $< 0.000001$ | −0.543 | −0.250 | *** |
| Distance × binocular | −0.014 | 0.106 | −0.136 | 0.892 | −0.221 | 0.193 | ns |

### 5.5.3 Monocular Versus Binocular Measurement Viewing

Direct paired comparisons between monocular and binocular parallax measurement displays showed that the monocular condition produced larger PSEs at all three distances, and this difference remained statistically reliable after Bonferroni correction across the three paired contrasts. At 0.4 m, the mono versus bino contrast was significant ($t(17) = 4.18$, raw $p = 0.000624$, Bonferroni-corrected $p = 0.001872$, Cohen's $d = 0.986$). The same pattern remained at 0.8 m ($t(17) = 8.95$, raw and corrected $p < 0.000001$, $d = 2.110$) and 1.2 m ($t(17) = 3.56$, raw $p = 0.002386$, Bonferroni-corrected $p = 0.007158$, $d = 0.840$). The 1.2 m paired-difference distribution departed from normality, and the Bonferroni-corrected Wilcoxon signed-rank test again supported the same mono-versus-bino difference ($p = 0.000572$). These comparisons indicate that the parallax-only measurement display consistently required stronger gain under monocular than binocular viewing to match the illusory display.

Table 5.10: Paired comparisons of fitted PSEs between monocular and binocular parallax measurement-display conditions in Experiment 5, shown separately for each observer distance. Raw and Bonferroni-corrected $p$ values are reported, with significance based on the corrected values.

| Contrast | N | Mean A | Mean B | Mean Diff | SD Diff | Cohen's $d$ | $t$ | df | Raw $p$ | Bonferroni $p$ | Shapiro $p$ | Sig |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.4 m: BM+ vs BB+ | 18 | 0.633 | 0.440 | 0.193 | 0.196 | 0.986 | 4.183 | 17 | 0.000624 | 0.001872 | 0.1920 | ** |
| 0.8 m: BM+ vs BB+ | 18 | 0.481 | 0.220 | 0.261 | 0.124 | 2.110 | 8.953 | 17 | < 0.000001 | < 0.000001 | 0.5842 | *** |
| 1.2 m: BM+ vs BB+ | 18 | 0.316 | 0.111 | 0.205 | 0.244 | 0.840 | 3.564 | 17 | 0.002386 | 0.007158 | 0.0018 | ** |

### 5.5.4 Summary Interpretation

Taken together, these results indicate that **the illusory-motion percept becomes weaker as the observer moves farther from the frame under the constant-size display geometry used here**. The monotonic reduction in PSE across both monocular and binocular parallax measurement displays was supported directly by the mixed-effects distance model, not only by descriptive condition means. Because no compensatory scaling was applied, this distance effect should be interpreted as the effect of observer-frame distance together with its associated change in visual angle, rather than as a pure distance manipulation isolated from retinal image size. At the same time, monocular measurement viewing consistently required larger positive gains than binocular measurement viewing, showing that measurement-display eye state continued to modulate the gain needed to match the stereo-driven illusion across all tested distances.

# Chapter 6

# Discussion

This chapter summarizes the empirical pattern across the five experiments and then discusses what that pattern implies for illusory motion parallax. The central claim is that illusory motion parallax arises when stereoscopic depth leads the observer to expect head coupled motion parallax, but the display does not provide the corresponding physical parallax. The strength of the percept, however, is not determined only by this local cue mismatch. It is also shaped by the measurement display, the surrounding scene context, and the viewing geometry in which the mismatch is interpreted.

Table 6.1 provides a compact overview of the five experiments, their manipulations, key results, and the status of the associated hypotheses, before each result is unpacked in the sections that follow.

Table 6.1: Overview of the five experiments, their manipulations, key results, and hypothesis outcomes. Hypothesis labels refer to Section 3; "Confirmed" indicates that the empirical pattern matched the directional prediction, "Reversed" indicates that the empirical pattern was reliable but in the opposite direction to the prediction, and "Tested" indicates that the relevant geometric expectation was directly probed but was not formulated as a strict directional hypothesis in advance.

| **Exp.** | **Manipulation** | **Key result** | **Hypothesis status** |
|---|---|---|---|
| 1 | HMD eye state (binocular vs. monocular) of the illusory and measurement displays, across positive and negative gains. | Reliable non-zero PSE when the illusory display was binocular; PSE near zero when it was monocular. Measurement-display viewing mode shifted PSE but did not abolish the illusion. | H1 (necessity of binocular disparity): Confirmed. |
| 2 | Direction-judgment task with the illusory display compared against positive and negative measurement parallax. | Illusory motion aligned with *negative* measurement parallax (opposite to natural head-coupled parallax). | H2 (signed direction predicted by the forward model): Confirmed. |
| 3 | Scene context (closed, open, no environment) and frame visibility (visible vs. hidden). | Illusion strongest with no environment, weakest with the closed environment; frame visibility had no reliable effect. | H3 (closer/more references strengthen the illusion): Reversed, reducing environmental references strengthened the illusion. |
| 4 | Vertical viewing angle of the display (eight downward angles from 2.83° to 54.23°). | Steeper downward viewing angles produced systematically more negative PSEs, indicating a stronger illusion. | H3 (geometric modulators): Tested, viewing angle reliably modulated illusion strength. |
| 5 | Observer distance from the frame under constant-size geometry (0.4, 0.8, 1.2 m), crossed with monocular vs. binocular measurement-display viewing. | Increasing distance shifted the PSE toward zero (weaker illusion); monocular measurement viewing required larger gains at every distance. | H3 (geometric modulators): Tested, observer distance reliably modulated illusion strength. |

## 6.1 Summary of Empirical Findings

Experiment 1 established the measurement paradigm and showed that the stereo-only display produced a measurable illusion when binocular disparity was present. When binocular disparity was removed from the illusory display by occluding one eye (monocular viewing), the measured PSE returned toward zero. A PSE of zero corresponds to the baseline at which no motion needs to be added to the parallax-only measurement display in order to match the perceived motion of the stereo-only display, that is, the case in which no illusion is reported. This pattern directly supports Hypothesis 1 (Section 3.2.1): binocular disparity is necessary for the illusion, rather than the task simply measuring a generic tendency to report motion. It is also consistent with prior accounts that emphasize the dependence of stereoscopic motion illusions on binocular disparity [Shimono et al., 2002, Rogers and Graham, 1982, Benzeroual and Allison, 2013].

Experiment 1 also showed that the comparison stimulus itself shapes the estimated strength of the illusion. The measurement display modulated the PSE in two ways. First, positive and negative measurement display gains were not perceptually symmetric: a smaller absolute negative gain was sufficient to match the stereo-only illusion than the corresponding positive gain, indicating that negative direction parallax in the measurement display carries greater perceptual weight per unit of gain than positive direction parallax. Second, monocular viewing of the measurement display required a larger absolute gain than binocular viewing to produce the same match against the stereo-only illusion. The conclusion is therefore not that the illusion is ill defined, but that the PSE reported in this thesis is the matching gain between two specific displays under the specific viewing conditions used in each experiment; the same illusion would be expected to produce different PSE values under different measurement display configurations.

Experiment 2 showed that the illusion has a signed direction. Observers judged the stereo-only illusory display as moving in the direction associated with *negative* measurement display parallax, that is, the direction opposite to the parallax that would naturally accompany the observer's head movement given a stable scene. This directional result is important because it shows that the illusion is not merely an increase in apparent motion or uncertainty. The forward model account predicts that the visual system generates a predicted parallax signal from stereoscopic depth and head movement; when the display provides no measured parallax, the predicted signal is not cancelled by an opposing measured signal, and the residual is experienced as motion whose direction is opposite to the physical parallax that would have cancelled it. The observed sign of the illusion matches that prediction, which is why we describe it as the sign expected from an uncancelled prediction of motion parallax, and supports Hypothesis 2 (Section 3.3.1).

Experiment 3 showed that scene context modulates the magnitude of the illusion. The illusion was strongest in the no environment condition, weaker in the open environment, and weakest in the closed environment. In other words, the strength of the illusion decreased as more environmental references were added around the judged display, consistent with the interpretation that the additional references helped disambiguate the spatial position and behavior of the framed scene relative to the surrounding world and therefore reduced the proportion of the cue mismatch that was attributed to motion of the judged content. Frame visibility produced a much smaller effect than the broader environment manipulation. Thus, the important contextual factor was not simply whether the display had a visible border, but how the display was embedded within a larger scene: a stereo-only display embedded in a stable surrounding layout is more readily explained as a stable display within a stable scene, whereas the same display presented in a sparse environment is more readily explained as itself in motion.

Experiments 4 and 5 showed that viewing geometry also matters. In Experiment 4, steeper downward viewing angles produced systematically more negative PSEs, indicating a stronger illusion at steeper angles. In Experiment 5, increasing observer distance from the frame shifted the PSE toward zero under the constant size display geometry used in the experiment, indicating a weaker illusory percept farther from the frame. Together, these results show that the illusion depends not only on which cues are present, but also on where the display sits in the observer's visual field and how much of the surrounding environment is available as a reference.

## 6.2 Measurement Context in Experiment 1

Experiment 1 served two purposes. First, it validated the matching procedure: when the illusory display was viewed binocularly (Set A), the fitted PSEs were reliably non-zero, whereas the same procedure applied to a monocularly viewed illusory display (Set B) returned PSEs that clustered near zero, that is, near the baseline at which no motion needs to be added to the measurement display in order to match the perceived motion of the illusory display (the case in which no illusion is reported). This confirms that the measured PSE tracks the presence of binocular disparity in the illusory display rather than a generic response bias, and it supports the claim that binocular vision is necessary for the illusion to be experienced, consistent with prior work showing that the strength of stereoscopic motion illusions depends on the presence of binocular disparity [Shimono et al., 2002, Rogers and Graham, 1982, Benzeroual and Allison, 2013]. Second, within Set A the fitted PSEs were not a single number that summarized the illusion. They depended both on the polarity of the measurement display gain and on whether the measurement display was viewed binocularly or monocularly. Because the two intervals in a trial were viewed in sequence rather than simultaneously, this dependency does not imply that the real parallax of the measurement display biases or contaminates the percept of the illusory display. Rather, because the PSE is by construction the gain value at which the two intervals are reported as equivalent in perceived motion. The PSE therefore reflects how each display is perceived under its respective viewing conditions, not a fixed strength of the illusion measured in isolation.

### 6.2.1 Asymmetry Between Positive and Negative Measurement Gains

Within Set A, the absolute PSE required to match the illusory display was asymmetric between the two measurement display gain polarities. A smaller absolute negative gain in the measurement display was sufficient to match the perceived motion of the illusory display than the corresponding positive gain, and this held for both monocular and binocular measurement display viewing. The asymmetry is therefore a property of the fitted matching curves between the illusory and measurement displays, not of the gain axis itself.

One interpretation of this asymmetry is that negative parallax is less natural for the visual system than positive head coupled parallax. In everyday viewing, when an observer moves their head, the retinal image of a stationary scene moves in a predictable, head coupled direction; the visual system has learned to expect this lawful relationship between self-motion and retinal motion. Positive measurement display gains preserve that relationship, whereas negative gains invert it, producing retinal motion in the direction opposite to what self-motion alone would normally generate. The same absolute gain may therefore feel more salient and more motion like in the negative direction than in the positive direction, simply because it deviates further from the predicted retinal consequences of head movement. On this account, a smaller absolute negative gain is sufficient to match the perceived motion of the illusory display not because the illusion itself is weaker when compared against negative gains, but because each unit of negative measurement parallax carries greater perceptual weight per unit of gain.

A second interpretation is that negative parallax resembles a moving world event. In physical viewing, the only way for image motion to remain opposite to head movement in the way produced by negative gain is for the surrounding scene itself to be moving relative to the observer in the same direction as the head, that is, against the observer's motion, at a magnitude exceeding what would be cancelled by simply moving with them. The negative parallax measurement display may therefore be interpreted not only as an unnatural depth cue but also as a scene that is actively in motion. This moving world interpretation could make the displayed scene appear faster or less stable than a comparable positive gain display, producing a lower absolute PSE when matching against the illusory display with negative measurement display gains.

### 6.2.2 Monocular and Binocular Measurement Displays

Within Set A, the illusory display itself was always viewed binocularly, and only the HMD viewing mode of the *measurement* display was manipulated. Under that arrangement, the absolute PSE required to match the illusory display was smaller when the measurement display was viewed binocularly than when the measurement display was viewed monocularly, and this held for both gain polarities. In other words, a smaller absolute parallax gain was sufficient to match the perceived motion of the binocular illusory display when the measurement display was also viewed binocularly than when the measurement display was viewed monocularly.

One way to interpret this result is in terms of cue conflict inside and outside the measurement frame. In the monocular measurement condition, the main depth and motion cue available in the measurement display was motion parallax itself, because the rest of the HMD scene was also viewed monocularly. In the binocular measurement condition, the measurement display still presented the parallax-only stimulus inside the frame, but the surrounding HMD scene outside the frame was viewed binocularly and therefore contained both stereoscopic depth and motion parallax information. This created a richer cue structure around a measurement stimulus whose defining signal was parallax. The contrast between a parallax-only region inside the frame and a binocularly specified surrounding scene outside the frame may have made the motion inside the measurement display more salient, leading observers to match the illusory display at lower absolute measurement display gains.

This interpretation also clarifies how the measurement display manipulation shapes the PSE measured in the later experiments. The PSE is a comparison between the illusory display and a real parallax measurement display, so any cue that changes the perceived strength of the measurement display can shift the estimated PSE. The binocular measurement condition is therefore useful and ecologically relevant, but it should be understood as a measurement context with its own cue structure, not as a neutral ruler. More broadly, the fact that all of these factors (gain polarity and measurement display viewing mode) produced systematic, quantifiable shifts in the fitted PSE rather than abolishing the matching curve provides important validation of the paradigm: the procedure is sensitive enough to capture small, principled changes in display properties, and is therefore well suited to detecting the contextual and geometric modulators examined in Experiments 3–5.

## 6.3 Signed Direction and Forward Model Interpretation in Experiment 2

The results of Experiment 2 support a forward model interpretation of the illusory motion parallax. A forward model predicts the sensory consequences of self-motion from the currently inferred scene structure [Miall and Wolpert, 1996, Wolpert and Kawato, 1998]. When an observer moves their head while viewing a stereoscopically defined object, the visual system expects the retinal image to update in a way consistent with the depth of that object in the scene. If the display provides stereoscopic depth but no corresponding motion parallax, the predicted parallax is simply absent from the sensory input. This residual, unconsumed prediction is what the next paragraph formalizes and what the directional result in Experiment 2 tests empirically.

The perceived scene motion can be described as a comparison between the parallax actually delivered to the retina and this prediction. A simple linear form of that comparison can be written as

$$P \approx \beta\,\pi_{\text{measured}} - \alpha\,\pi_{\text{predicted}},$$

where the symbols denote the following quantities:

$P$ perceived scene motion attributed to the framed display (positive when in the natural head coupled direction; negative in the opposite direction).

$\pi_{\text{measured}}$ motion parallax actually delivered to the retina during the head movement.

$\pi_{\text{predicted}}$ motion parallax predicted by the forward model from the current head movement and the stereoscopically specified depth of the scene.

$\alpha, \beta \geq 0$ non-negative weights placed by the visual system on the predicted and measured parallax signals, respectively.

In natural viewing, the rendered scene matches its inferred three-dimensional structure, so $\pi_{\text{measured}} \approx \pi_{\text{predicted}}$ and the two terms approximately cancel; the scene is perceived as stable while the retinal changes are attributed to the observer's own head movement or self-motion through space. In the stereo-only illusory display, binocular disparity is present so $\pi_{\text{predicted}}$ remains non-zero, but no motion parallax is rendered, so $\pi_{\text{measured}} = 0$. The equation reduces to

$$P \;\approx\; -\alpha\,\pi_{\text{predicted}},$$

which is signed opposite to the parallax that the inferred scene geometry would naturally produce during head movement.

The important result from Experiment 2 is that this mismatch is signed in precisely the manner predicted by the forward model account. The illusory motion aligned with negative measurement display parallax, consistent with the idea that the percept reflects the residual of a predicted parallax signal that was not delivered by the display. In other words, the observer does not simply notice that something is inconsistent: the visual system interprets the missing update as scene motion in a specific direction, the direction opposite to the parallax that natural head coupled viewing would have produced.

This directional result also informs an analytical decision in the subsequent experiments. Because the illusion was perceived in the negative parallax direction, comparing it against a measurement display that moved in the *same* perceptual direction provides the most sensitive matching regime: the two intervals are then varying along a common perceptual dimension. Negative measurement display gains were therefore adopted as the reference direction in Experiments 3 and 4, so that the stereo-only display could be quantified against a real parallax display whose motion was already in the perceptual direction predicted by the forward model account.

## 6.4 Scene Context and Attribution in Experiment 3

Following from the conclusion of the previous section, Experiment 3 used the negative parallax measurement direction established in Experiment 2 as a sensitive baseline against which to compare the stereo-only illusion, while varying the scene surrounding the illusory display. Three scene layouts were tested: a closed environment with nearby walls and surfaces, an open environment that retained farther background structure, and a no environment condition that removed all external references. Within each scene, the border frame around the illusory display was either visible or hidden. The illusion was strongest in the no environment condition, intermediate in the open environment, and weakest in the closed environment, while frame visibility did not produce a reliable shift and showed no evidence of an interaction with the scene environment manipulation. The Experiment 3 findings therefore demonstrate a strong scene context effect on the magnitude of the illusion and a negligible frame effect.

### 6.4.1 Environmental Reference and the Attribution of the Cue Mismatch

The direction of the environment effect ran opposite to our initial predictions, which held that nearby or more numerous references would make the missing parallax easier to detect and therefore strengthen the illusion. The observed pattern reversed that expectation: reducing environmental context strengthened the illusion, while adding structured scene elements weakened it. This does not mean that environmental references were irrelevant. Rather, it suggests that references help to disambiguate the spatial behavior of the framed scene relative to the surrounding world, and in doing so they also change the perceptual explanation that the visual system assigns to the cue mismatch. In other words, the missing parallax is the same physical mismatch in every condition; what differs is how the visual system explains that mismatch, and this explanation depends on the surrounding scene structure available to anchor the framed display.

When the stereo-only display appears with little external context, the framed object becomes the main carrier of expected motion parallax. The missing update is therefore more directly assigned to that object, making the illusory motion more pronounced. When the same display is embedded in a structured environment, the visual system has more ways to stabilize the perception of object depth. The mismatch can be distributed across the surrounding layout, discounted as part of the display geometry, or interpreted as less diagnostic of object motion. In this sense, additional context weakens the apparent motion of the target even though it provides more spatial information.

### 6.4.2 Frame Visibility and Scene Continuity

The frame manipulation was intended to test whether the local border around the illusory display would behave like the larger scene structure. If the effect of scene context was driven mainly by the nearest visible border, frame visibility should have produced a substantial shift in PSE between the visible frame and hidden frame conditions. Instead, the broader environment shaped the illusion much more than the local frame, and the data showed no clear indication of an interaction between frame visibility and scene environment at all. A visible frame may mark the display as a bounded window, but it does not by itself determine whether the space inside the frame is experienced as continuous with the observer's surrounding environment or as a separate, divided region. The closed, open, and no environment layouts changed that global continuity, that is, whether the target stimulus space was perceived as part of the same scene as the observer or as a self-contained region behind the frame, and the cue mismatch was resolved within that global interpretation. The frame itself operated within that interpretation rather than as an independent driver of the illusion's perceived strength.

## 6.5 Viewing Angle in Experiment 4

Experiment 4 examined how the vertical viewing angle, specifically, downward tilt that placed the display in the lower half of the visual field, influenced the strength of illusory motion parallax. The frame and surrounding scene were presented at a range of downward angular offsets relative to the observer's eye level, and the measurement display used the established negative gain set. As the downward viewing angle became steeper, a more negative parallax was required in the measurement display to match the perceived motion of the stereo-only display, indicating that the illusion was perceived as stronger at steeper downward angles.

This finding can be discussed as an effect of display position and height within the visual field. At steeper downward viewing angles, the frame and object arrangement was viewed from a lower position relative to the observer's eye level and projected primarily onto the lower half of the retina, which is the region of the visual field most strongly engaged by near, ground related visual information. One possible interpretation is that this geometry made the target stimulus appear closer to the observer's feet and therefore more behaviourally or perceptually consequential, and more sensitive to head movement. The lower visual field is known to be particularly important for processing optic flow and motion parallax in support of postural control, locomotion, and balance: information from this region of the visual field contributes to stabilizing the body and to specifying near space layout. If the stereo-only display recruits these near space, lower visual field mechanisms more strongly at steeper downward angles, it would also imply a stronger expected parallax update from a forward model perspective, and the absence of physical parallax would then produce a more pronounced apparent motion.

A second interpretation focuses on the measurement display rather than the stereo-only display. The effect may not mean that the illusory display itself became stronger at steeper angles. Instead, the motion parallax measurement display may have become a weaker comparison stimulus. At steeper angles, the observer may have had less useful far background structure available for judging real parallax in the measurement interval. If the far background provides reference information for distinguishing near object motion from global scene motion, then losing that reference would reduce the perceived strength or reliability of the measurement display parallax. The stereo-only illusion would then appear stronger in comparison, even if the underlying illusion did not increase by the same amount.

These two explanations are not mutually exclusive. Steeper downward angles may both increase the expected parallax in the stereo-only display, via the recruitment of lower visual field near space mechanisms, and reduce the effectiveness of the measurement display as a comparison reference by removing useful far background structure. The observed angle effect therefore supports the broader claim that viewing geometry changes both sides of the matching task: the expected parallax generated by the stereo-only display and the perceived strength of real parallax in the measurement display.

## 6.6 Observer Distance in Experiment 5

Experiment 5 examined whether observer distance from the frame had an effect on the strength of illusory motion parallax under a constant size display geometry. As observer distance increased, less positive parallax in the measurement display was required to match the perceived motion of the stereo-only display, indicating that the illusion was perceived as weaker at larger distances. The same experiment also replicated the measurement display viewing pattern from Experiment 1: monocular viewing of the measurement display required larger positive gains than binocular viewing at every distance.

Several non-exclusive interpretations are consistent with this distance effect, and we list them briefly before examining each in turn: (i) a geometric, retinal size account in which both the retinal displacement produced by a fixed head translation and the effectiveness of motion parallax as a depth cue decrease with viewing distance; (ii) a stereoscopic reliability account in which the retinal disparity generated by a fixed depth separation also decreases with viewing distance, weakening the stereo derived expectation for parallax; and (iii) a scene context account that links the distance effect back to Experiment 3 by way of the changing proportion of the visual field that is occupied by the framed display.

The first interpretation is geometric. The retinal image motion produced by a given lateral head translation depends on the distance from the observer to the target: for a fixed head movement, both retinal displacement and the rate of retinal change scale approximately with the inverse of viewing distance, so the same physical head translation produces smaller image motion of the framed scene at larger distances. This same geometry also governs the diagnosticity of motion parallax as a depth cue: motion parallax is most informative for near targets and progressively less so at far distances, where small retinal displacements become harder to use to disambiguate depth [Rogers and Graham, 1979, Nawrot and Stroyan, 2009]. Under the constant size geometry used in Experiment 5, increasing observer distance therefore reduced both the visual angle subtended by the frame and the per distance effectiveness of any parallax signal that the visual system could derive from head movement. The illusion, which arises from a stereo driven expectation of parallax that is then not delivered, should accordingly be weaker at larger distances simply because the expected parallax signal would itself be smaller and less perceptually informative at those distances. This is consistent with the weaker illusion observed at larger distances.

A second interpretation, complementary to the first, concerns stereoscopic depth reliability. For a fixed scene geometry, the retinal disparity generated by a given depth separation also becomes smaller as viewing distance increases [Rogers and Graham, 1982, Howard and Rogers, 2012]. Because the illusion depends on stereoscopic depth generating an expectation for motion parallax, a weaker or less reliable disparity signal at farther distances could reduce the expected parallax update and therefore reduce the resulting illusory motion. This stereoscopic reliability account and the geometric retinal size account share a common consequence: both predict that the magnitudes that drive the illusion (expected disparity based depth and the resulting expected parallax) shrink with distance, and they can be regarded as two sides of the same distance related attenuation rather than as independent mechanisms. The present experiment did not isolate stereoscopic reliability from visual angle or scene scale, so the relative contribution of each cannot be quantified here; what the result does show is that the stereo-only cue conflict is strongest in the viewing regime where stereoscopic depth and motion parallax are both perceptually most effective.

A third interpretation connects Experiment 5 to the scene context result in Experiment 3. At close distances, the frame occupies more of the observer's visual field, leaving less surrounding environment visible outside the judged display. This makes the close distance condition resemble the reduced context case from Experiment 3, where the illusory motion was strongest. At farther distances, more of the surrounding scene is available as a reference, which may help stabilize the display and reduce the motion assigned to the stereo-only display. Thus, Experiment 5 supports the broader conclusion that reducing external context can make the illusory motion more dominant.

## 6.7 Synthesis: Local Prediction Error, Global Scene Interpretation

Across the experiments, illusory motion parallax is best understood as a local prediction error filtered through a global scene interpretation. The local mechanism begins with stereoscopic depth and head movement. Together, they lead the observer to expect a motion parallax update. When that update is absent, the remaining mismatch has a signed structure and can be experienced as motion. This explains why binocular disparity is necessary and why the illusion aligns with negative parallax.

The global interpretation determines how strongly that local mismatch is experienced. The measurement display changes the comparison scale; the surrounding environment changes whether the mismatch is assigned to the target object or absorbed into the scene; and viewing geometry changes both expected parallax and the reliability of the comparison stimulus. The illusion is therefore not a fixed property of the stereo-only display. It is the outcome of a perceptual explanation built from the target, the measurement display, the surrounding scene, and the observer's pose.

This synthesis also refines the forward model account. A forward model does not operate over an isolated object stripped of context. It operates over an interpreted scene. The same stereo-only object can therefore produce different measured strengths depending on whether it is surrounded by a stable room, presented with minimal context, viewed from a steep angle, or seen from a near or far distance. The prediction error may be similar in origin, but its perceptual attribution changes with the scene representation.

## 6.8 Implications for Virtual Reality Design

The practical implication for VR is that cue conflict should be evaluated at the level of the full viewing situation, not only at the level of the target object. A stereoscopic object without appropriate motion parallax may appear more or less unstable depending on the surrounding environment, the amount of visible context, the viewing angle, and the observer's distance from the display. Designers who want to reduce unintended illusory motion may benefit from stable surrounding structure and consistent motion parallax cues, while designers who want to amplify apparent motion may use reduced context or viewing geometries that make the mismatch more dominant.

The results also suggest that a single calibration point is not enough. The same display content can produce different perceptual consequences across observer positions. This matters for head-mounted displays because users naturally move through a range of distances and angles. Cue consistency should therefore be checked across the range of poses that users are likely to occupy, especially when stereoscopic depth is presented without a fully matching parallax update.

## 6.9 Limitations

Several limitations qualify the interpretation. First, the experiments measured perceived relative strength through comparison tasks. This approach is well suited for estimating PSEs, but it does not directly reveal the internal attribution process. The discussion therefore treats attribution, cue reliability, and forward model prediction as explanatory accounts of the behavioral pattern rather than as directly measured internal variables.

Second, the measurement display is itself part of the perceptual comparison. Experiment 1 shows that changing the measurement display viewing mode changes the resulting PSE. Later experiments used a consistent measurement display configuration to support comparison across conditions, but the PSE should still be interpreted as a relative matching measure rather than an absolute readout of the illusion.

Third, the distance manipulation in Experiment 5 changed observer distance while holding the physical frame and scene geometry constant. As a result, observer distance, visual angle, and the amount of surrounding visible context were not fully separated. This design captures an important practical viewing geometry, but it should not be interpreted as a pure distance manipulation. Similarly, the viewing angle manipulation in Experiment 4 changed the position of the display within the visual field, and the resulting effect may reflect both changes in expected parallax and changes in the effectiveness of the measurement display.

Fourth, the magnitude of the stereoscopic disparity in the illusory display was held fixed across all experiments. Specifically, the inter-ocular separation of the two embedded virtual cameras was set to the participant's measured inter-pupillary distance (IPD), as reported automatically by the Meta Quest 3 plugin, and no IPD-gain or supra/sub-IPD manipulation was applied. The forward-model account predicts that the magnitude of the illusion should scale with the expected parallax, which in turn depends on the stereoscopically specified depth. Holding disparity at the natural IPD provides a stable, ecologically meaningful reference but does not let us test that scaling relationship directly. Varying the IPD-gain or otherwise manipulating disparity magnitude would allow a direct test of whether the strength of the illusion tracks the magnitude of the stereoscopically specified depth, as the forward-model interpretation would predict.

Finally, the experiments used controlled scenes, constrained gain values, and a magnitude comparison task. More complex interactive environments, richer textures, larger head movements, or naturalistic tasks may reveal additional ways in which context and geometry shape the same underlying cue conflict.

## 6.10 Summary

The main interpretive shift is from a cue strength account to an attribution account. Illusory motion parallax begins with a signed mismatch between expected and absent motion parallax, but the visual system must still decide what that mismatch means. Experiment 1 shows that the measurement display shapes the comparison scale. Experiment 2 shows that the mismatch has a signed direction. Experiment 3 shows that surrounding scene context changes how strongly the mismatch is assigned to the object. Experiments 4 and 5 show that viewing geometry and visual field structure further modulate the effect. The illusion is therefore best understood as a forward model prediction error whose perceived strength depends on measurement context, scene context, and viewing geometry.

# Chapter 7

# Future Work

Building upon the findings of this study, several avenues for future research are proposed to further explore the complexities of depth perception and illusory motion in virtual reality (VR) environments:

## 7.1 Research Extensions

1. **IPD-gain and disparity-magnitude manipulation:** The present experiments held the stereoscopic disparity of the illusory display at the participant's natural inter-pupillary distance (Section 4). The forward-model account predicts that the strength of the illusion should scale with the magnitude of the expected parallax, which in turn depends on the stereoscopically specified depth. A direct test would manipulate IPD-gain or otherwise vary disparity magnitude while holding the rest of the geometry constant, and ask whether the illusion magnitude tracks expected parallax in the predicted way.

2. **Other depth-cue manipulations:** Extend the paradigm to other visual depth cues beyond stereopsis and parallax. Systematically removing or manipulating linear perspective, height-in-field, occlusion, and shadows would test whether the scene-context reversal observed in Experiment 3 generalizes across cue types, and would clarify whether the attribution mechanism that resolves the cue mismatch operates over the same global scene representation regardless of which cues define depth.

3. **Multisensory integration:** Investigate how non-visual cues, in particular auditory and haptic cues that specify a stable surrounding scene, interact with the scene-attribution

effect demonstrated here. If reducing visual scene context strengthens the illusion by removing alternative attributions of the cue mismatch, then adding non-visual scene structure that supports a stable world should weaken the illusion even when visual context is held constant.

4. **Neurophysiological correlates:** Use neuroimaging to identify the neural substrate of the prediction-error signal proposed in the forward-model account, and in particular to test whether activity associated with the prediction error is modulated by the scene-context manipulation used in Experiment 3. This would provide a biological complement to the behavioral attribution account.

5. **Adaptation and longitudinal studies:** Test whether prolonged exposure to stereo-only VR environments alters the strength of the illusion over time, and whether the forward-model prediction adapts to non-natural disparity-parallax relationships. The cue-context dependence observed here predicts that adaptation should depend on the scene structure under which exposure occurs, not only on the cue conflict itself.

# Chapter 8

# Conclusion

This thesis investigated illusory motion parallax as a cue-conflict phenomenon in virtual reality. By decoupling stereoscopic depth from the motion parallax normally produced by head movement, the experiments tested when the visual system predicts parallax, how the resulting mismatch is experienced, and which scene and viewing factors change the strength of the percept.

The main conclusion is that illusory parallactic motion is not a simple by-product of having many visual references in the scene, nor of having references close to the display. The illusion requires binocular disparity and has a reliable signed direction, supporting a forward-model account in which stereoscopic depth and self-motion generate an expectation of motion parallax. However, the magnitude of the illusion depends on how the whole scene is interpreted. Most importantly, Experiment 3 showed that the illusion was strongest when surrounding context was removed, revealing that closer or more numerous scene references do not necessarily strengthen the effect.

The broader contribution of the thesis is therefore a revised account of the illusion. Motion-parallax prediction remains central, but the perceived strength of the illusion is shaped by global scene attribution and viewing geometry. Close and far environmental references may stabilize the scene and reduce the amount of mismatch attributed to target motion, while steeper viewing angles and nearer viewing geometries can make the stereo-driven mismatch more perceptually consequential.

These findings matter for VR because they show that cue conflict cannot be evaluated only at the level of a single object or display frame. The same stereoscopic content can produce different apparent motion depending on the surrounding environment, viewing angle, and observer distance. More natural virtual environments will therefore require not only accurate depth cues, but also careful control over how those cues interact with scene context and user movement.

In final terms, this work shows that illusory motion parallax is best understood as the product of local cue conflict filtered through global scene interpretation. The visual system does not simply detect missing motion parallax; it decides how that mismatch should be explained within the scene. That attribution process is what determines whether the illusion remains weak, becomes measurable, or dominates the perceptual comparison.

# Bibliography


Robert S Allison, Brian J Rogers, and Mark F Bradshaw. Geometric and induced effects in binocular stereopsis and motion parallax. *Vision Research*, 43(17):1879–1893, 2003. doi: 10.1016/S0042-6989(03)00298-0.

Adam O Bebko and Nikolaus F Troje. Bmltux: Design and control of experiments in virtual reality and beyond. May 2020. doi: 10.31234/osf.io/arvkf. URL `psyarxiv.com/arvkf`.

Karim Benzeroual and Robert S Allison. Cyber (motion) sickness in active stereoscopic 3d gaming. In *2013 International Conference on 3D Imaging*, pages 1–7. IEEE, 2013.

Mark F Bradshaw and Brian J Rogers. The interaction of binocular disparity and motion parallax in the computation of depth. *Vision research*, 36(21):3457–3468, 1996.

Mark F Bradshaw, Andrew D Parton, and Richard A Eagle. The interaction of binocular disparity and motion parallax in determining perceived depth and perceived size. *Perception*, 27(11):1317–1331, 1998.

Dorita H. F. Chang, Laurence R. Harris, and Nikolaus F. Troje. Frames of reference for biological motion and face perception. *Journal of Vision*, 10(6):22, 2010. doi: 10.1167/10.6.22.

Andy Clark. Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204, 2013. doi: 10.1017/S0140525X12000477.

James E. Cutting and Peter M. Vishton. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In William Epstein and Sheena Rogers, editors, *Perception of Space and Motion*, pages 69–117. Academic Press, San Diego, CA, 1995.

Marc O. Ernst and Martin S. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415:429–433, 2002. doi: 10.1038/415429a.

Karl Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11:127–138, 2010. doi: 10.1038/nrn2787.

Barbara Gillam. Matching needed for stereopsis. *Nature*, 373:202–203, 1995. doi: 10.1038/373202a0.

Walter C Gogel. A theory of phenomenal geometry and its applications. *Perception & Psychophysics*, 48(2):105–123, 1990.

Walter C Gogel and Jerome Tietz. The effect of perceived distance on perceived movement. *Perception & Psychophysics*, 16(1):70–78, 1974. doi: 10.3758/BF03203253.

Richard L. Gregory. Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society B*, 352(1358):1121–1127, 1997. doi: 10.1098/rstb.1997.0095.

David M. Hoffman, Ahna R. Girshick, Kurt Akeley, and Martin S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3):33, 2008. doi: 10.1167/8.3.33.

Ian P Howard and Brian J Rogers. *Perceiving in depth, volume 2: Stereoscopic vision*. Number 29. OUP USA, 2012.

Daniel Kersten, Pascal Mamassian, and Alan Yuille. Object perception as bayesian inference. *Annual Review of Psychology*, 55:271–304, 2004. doi: 10.1146/annurev.psych.55.090902.142005.

David C. Knill and Alexandre Pouget. The bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719, 2004. doi: 10.1016/j.tins.2004.10.007.

Jan J. Koenderink and Andrea J. van Doorn. The structure of visual spaces. *Journal of Mathematical Imaging and Vision*, 31:171–187, 2008. doi: 10.1007/s10851-008-0076-3.

Jan J. Koenderink, Andrea J. van Doorn, and Johan Wagemans. Depth. *i-Perception*, 2(6):541–564, 2011. doi: 10.1068/i0438aap.

R Chris Miall and Daniel M Wolpert. Forward models for physiological motor control. *Neural networks*, 9(8):1265–1279, 1996. doi: 10.1016/S0893-6080(96)00035-4.

Mark Nawrot and Randolph Blake. Neural integration of information specifying structure from stereopsis and motion. *Science*, 244(4905):716–718, 1989.

Mark Nawrot and Keith Stroyan. The motion/pursuit law for visual depth perception from motion parallax. *Vision research*, 49(15):1969–1978, 2009. doi: 10.1016/j.visres.2009.05.008.

Hiroshi Ono and Hiroyasu Ujike. Motion parallax driven by head movements: Conditions for visual stability, perceived depth, and perceived concomitant motion. *Perception*, 34(4): 477–490, 2005.

Thomas V Papathomas and Lisa M Bono. Experiments with a hollow mask and a reverspective: Top-down influences in the inversion effect for 3-d stimuli. *Perception*, 33(9):1129–1138, 2004.

Mary A Peterson. Illusory concomitant motion in ambiguous stereograms: evidence for nonstimulus contributions to perceptual organization. *Journal of Experimental Psychology: Human perception and performance*, 12(1):50, 1986.

Rajesh P. N. Rao and Dana H. Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1): 79–87, 1999. doi: 10.1038/4580.

Brian Rogers and Maureen Graham. Motion parallax as an independent cue for depth perception. *Perception*, 8(2):125–134, 1979. doi: 10.1068/p080125.

Brian J. Rogers and Maureen E. Graham. Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, 22(2):261–270, 1982. doi: 10.1016/0042-6989(82)90126-2. URL `https://doi.org/10.1016/0042-6989(82)90126-2`.

K Shimono, WJ Tam, and H Ono. Apparent motion of monocular stimuli in different depth planes with lateral head movements. *Vision Research*, 47(8):1027–1035, 2007. doi: 10.1016/j.visres.2007.01.012.

Koichi Shimono, Wa James Tam, Lew Stelmach, and Evan Hildreth. Stereoillusory motion concomitant with lateral head movements. *Perception & psychophysics*, 64(8):1218–1226, 2002.

Nikolaus F. Troje. Reference frames for orientation anisotropies in face recognition and biological-motion perception. *Perception*, 32:201–210, 2003. doi: 10.1068/p3392.

Nikolaus F Troje. Reality check. *Perception*, 48(11):1033–1038, 2019. doi: 10.1177/0301006619879062.

Xiaoye Michael Wang and Nikolaus F. Troje. Relating visual and pictorial space: Binocular disparity for distance, motion parallax for direction. *Visual Cognition*, 31(2):107–125, 2023. doi: 10.1080/13506285.2023.2203528.

Xiaoye Michael Wang and Nikolaus F. Troje. Relating visual and pictorial space: Integration of binocular disparity and motion parallax. *Journal of Vision*, 24(13):7, 2024. doi: 10.1167/jov.24.13.7.

Daniel M Wolpert and Mitsuo Kawato. Multiple paired forward and inverse models for motor control. *Neural networks*, 11(7-8):1317–1329, 1998. doi: 10.1016/S0893-6080(98)00066-5.

# Appendix A

# Appendix

# A.1 Experiment 1 Participant-Level Psychometric Fits

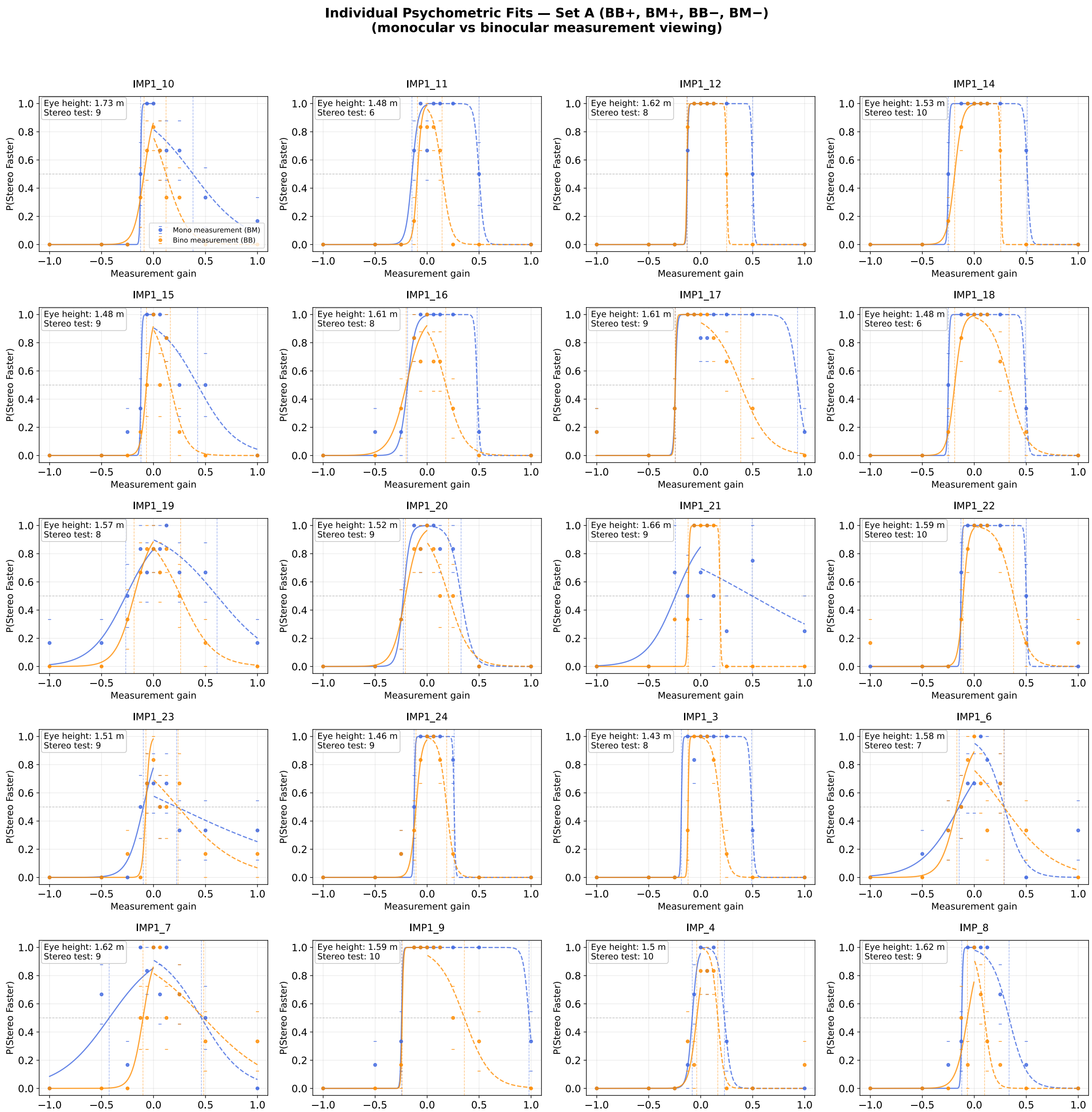


Figure A.1: Participant-wise psychometric fits for the Set A binocular-illusory-display conditions in Experiment 1. Each panel shows one participant, allowing visual comparison of the four stereo-illusory conditions at the individual level.

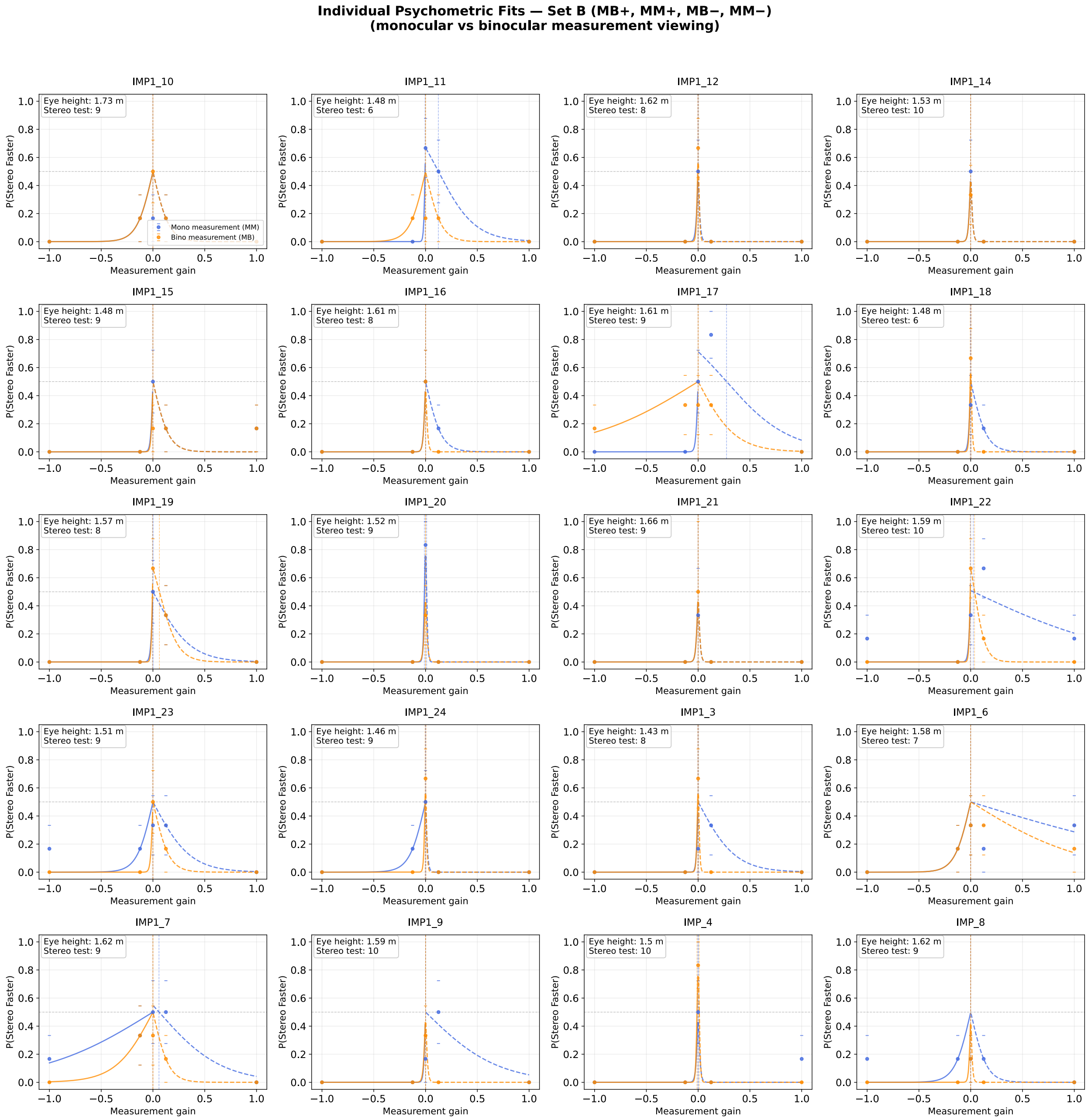


Figure A.2: Participant-wise psychometric fits for the Set B monocular-illusory-display conditions in Experiment 1. Each panel shows one participant, illustrating that responses in the monocular-illusory conditions remain comparatively weak and do not show the same clear illusion pattern as Set A.

## A.2 Experiment 1 All Eight Conditions Viewed Together

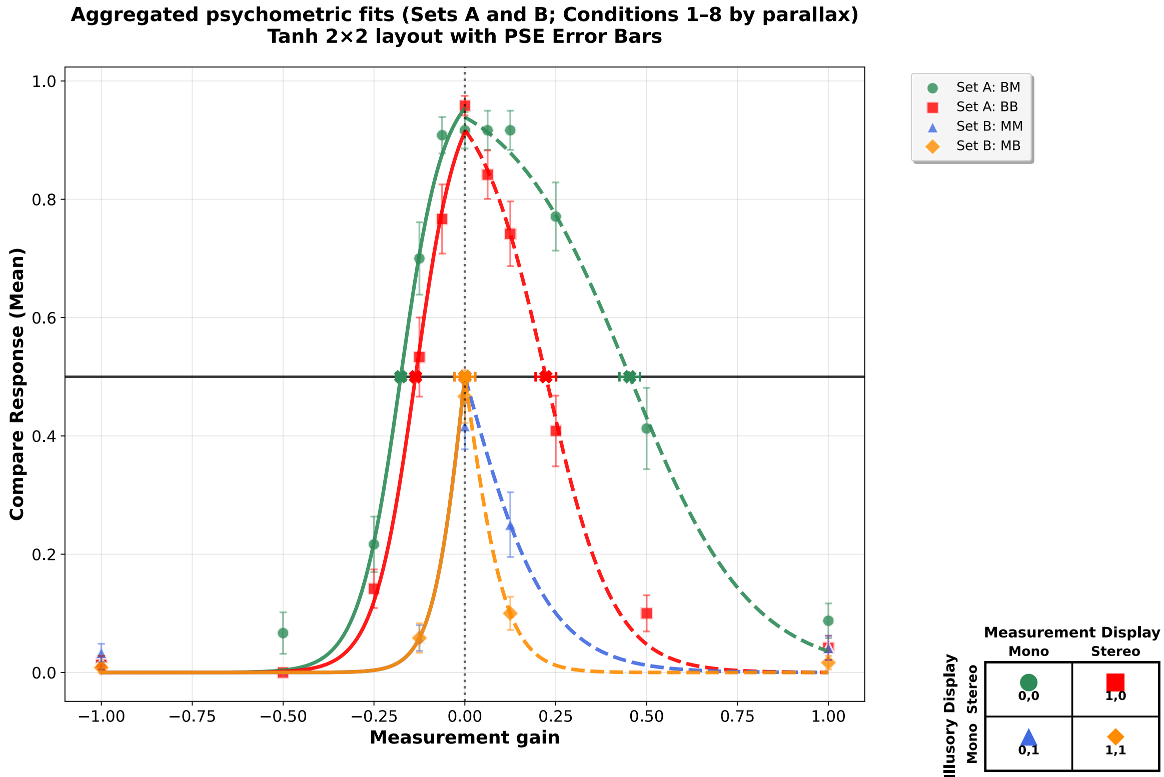


Figure A.3: Combined psychometric overview of all eight Experiment 1 conditions (N=20), essentially uniting Set A (binocular illusory display) and Set B (monocular illusory display) in a single plot so that all conditions can be viewed together. The horizontal axis represents the parallax gain values in the measurement display. The vertical axis shows participants' average responses on a 0–1 scale, where values closer to 1 indicate stronger perceived motion in the illusory motion parallax display relative to the measurement display. A value of 0.5 indicates equal perceived motion between the two. The Set A binocular-illusory conditions (BB+, BM+, BB−, BM−) show the standard illusion comparisons, whereas the Set B monocular-illusory conditions (MB+, MM+, MB−, MM−) show the disparity-removed comparisons. In the monocular-illusory conditions, responses remained below 0.5, confirming that the illusion requires stereoscopic disparity.

## A.3 Experiment 2 Participant-Level Plots

These supplementary figures show the participant-level directional-judgment patterns for Experiment 2 using the full participant-wise panel visualization and a consistency summary.

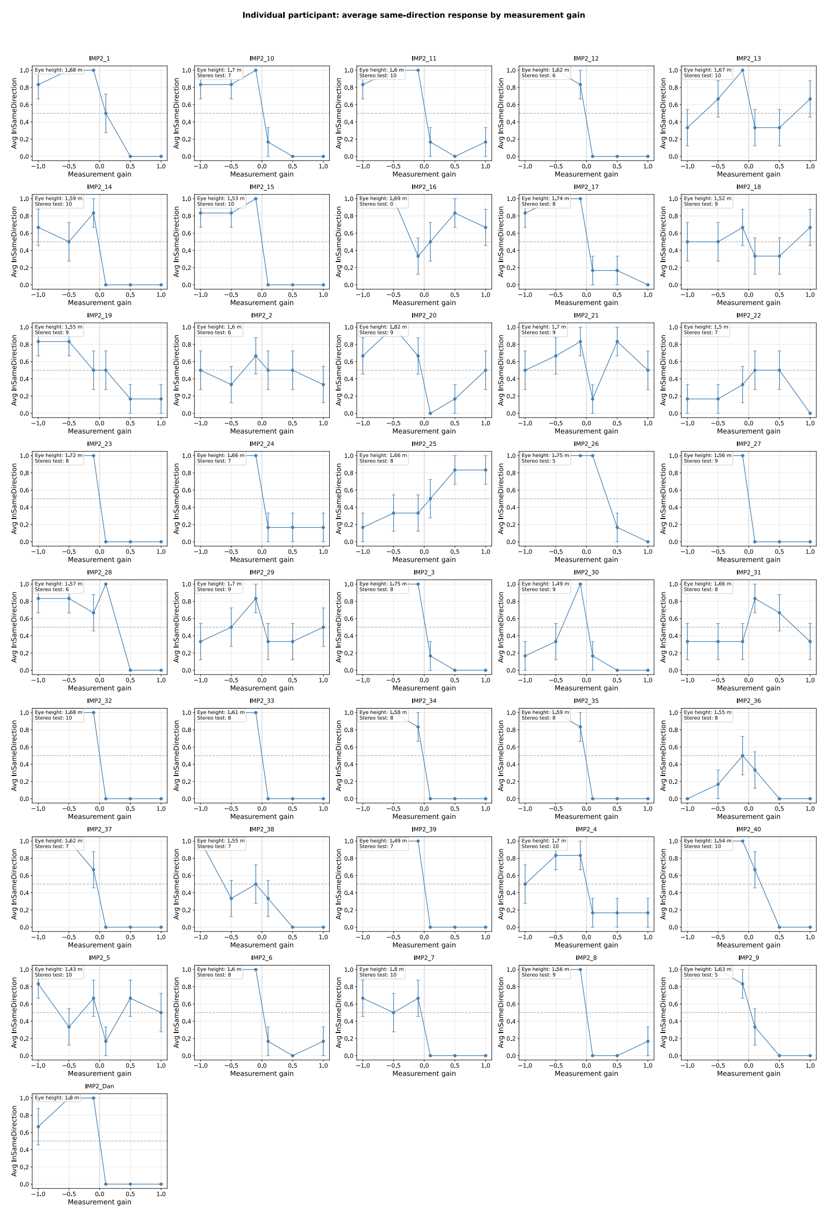


Figure A.4: Per-participant proportion of "same direction" responses as a function of measurement-display gain in Experiment 2. Each panel corresponds to one participant. The consistent drop from high values at negative gains to low values at positive gains indicates that the directional alignment of the illusion with negative parallax held at the individual level.

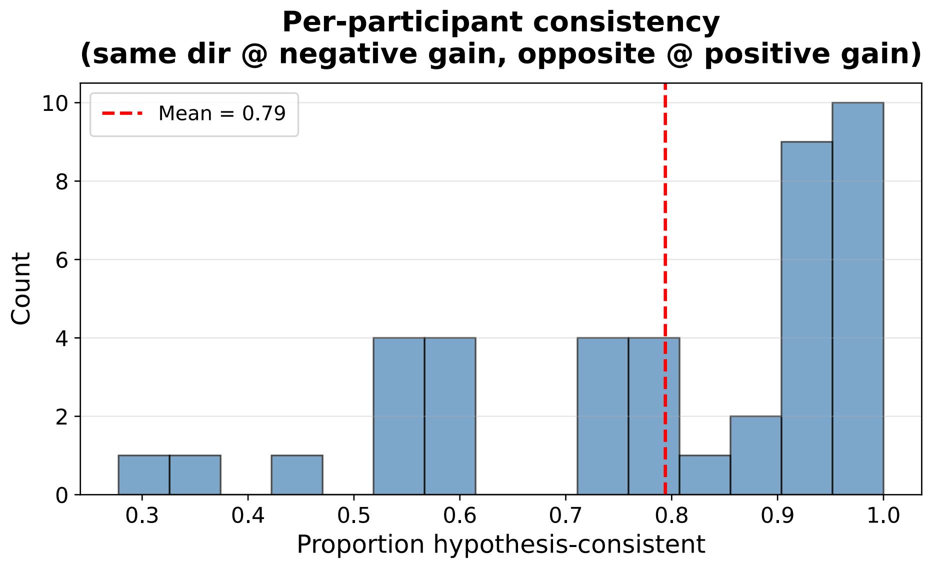


Figure A.5: Per-participant consistency with the directional Hypothesis in Experiment 2, defined as "same direction" responses for negative gains and "opposite direction" responses for positive gains. Most participants show high consistency, indicating that the group-level effect reflects broadly shared perceptual judgments.

## A.4 Experiment 3 Participant-Level Plots

This supplementary figure shows the participant-level psychometric visualization for Experiment 3.

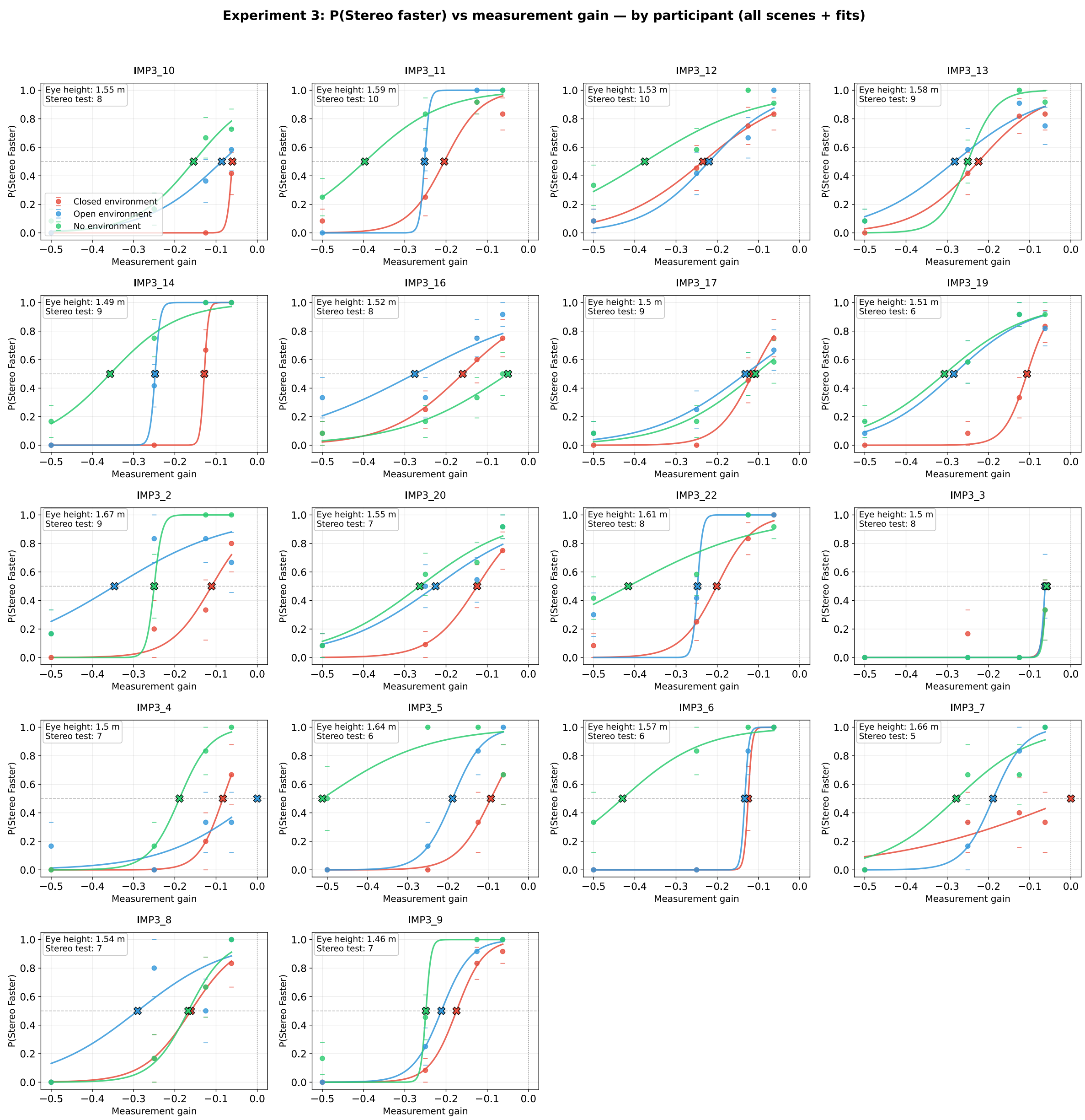


Figure A.6: Per-participant fitted psychometric curves for Experiment 3. Each panel shows the three scene contexts for one observer, with fitted PSEs marked where available. The figure illustrates both substantial individual variability and the recurring tendency for the closed-walls condition to yield less negative PSEs than the open-walls or no-context conditions.

# A.5 Experiment 4 Participant-Level Plots

This supplementary figure shows the participant-level fitted psychometric visualization for Experiment 4.

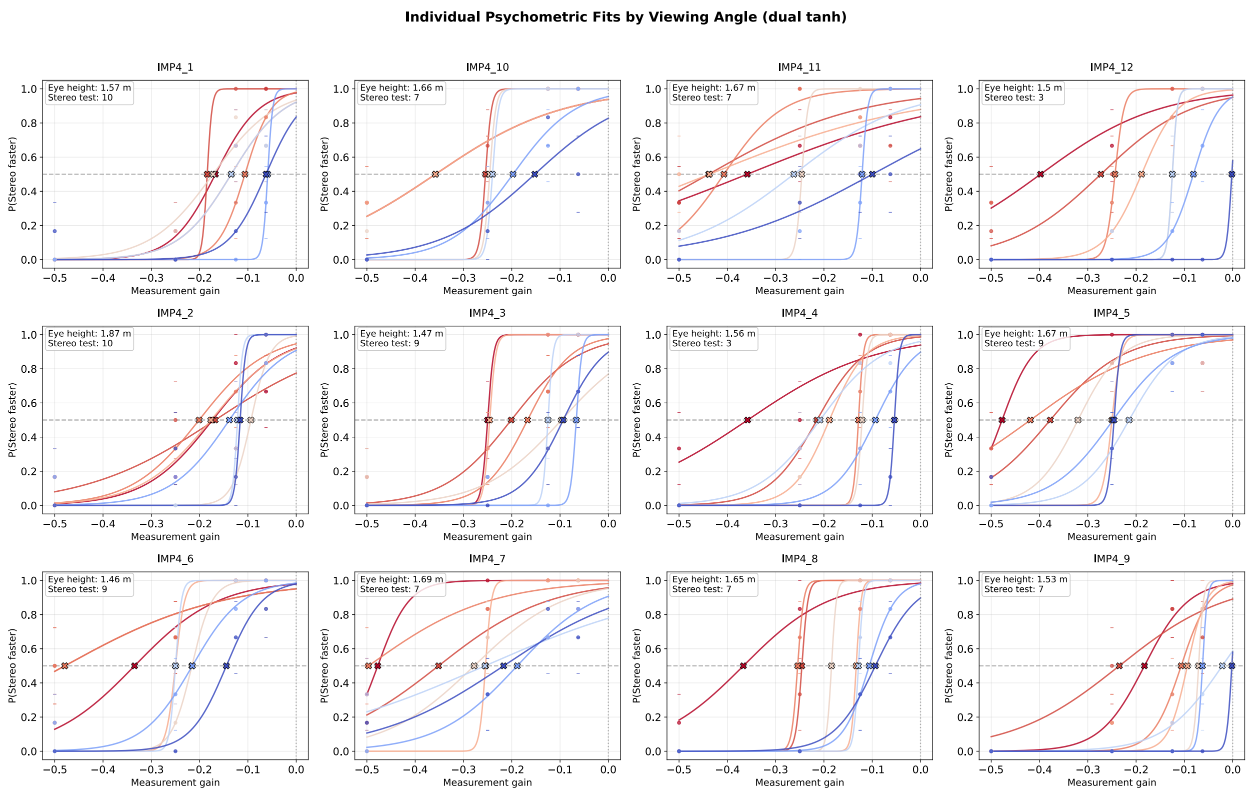


Figure A.7: Participant-wise fitted psychometric curves for Experiment 4. Each panel shows one participant, with fitted psychometric functions plotted across measurement-display gain separately for the different viewing-angle conditions. Although participants vary in overall noise and spread, many show systematic differences across viewing-angle conditions consistent with the group-level angle effect.